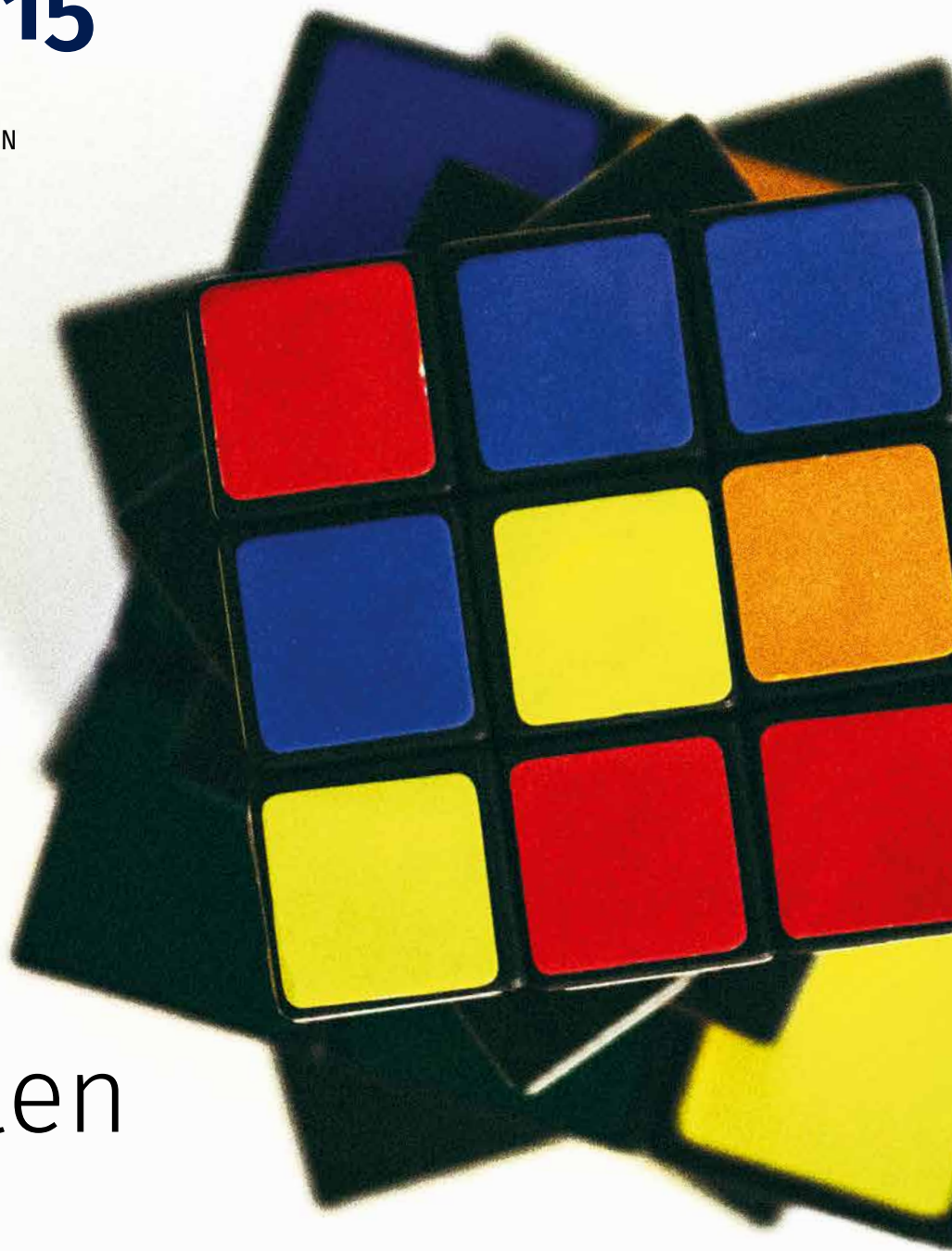


.inf¹⁵

DAS INFORMATIK-MAGAZIN
HERBST 2026

GESELLSCHAFT
FÜR INFORMATIK



Von allen Seiten

Digitale Resilienz braucht viele Handgriffe

KI-Audits: Vertrauen oder Trial
and Error? **SEITE 12**

IT-Controlling: Ein Blick ins Archiv
der Fachgruppe **SEITE 20**

Barrierefreiheit: Wie Software
für alle funktioniert **SEITE 28**

cargo

test



Martin Wolf, Präsident der
Gesellschaft für Informatik e.V.

Unsere Jahrestagung dreht sich diesmal um das Thema digitale Resilienz: ein Anspruch, der schwer zu erfüllen und noch schwerer zu halten ist. Denn resilient sind Systeme nur dann, wenn sie unterschiedlichen Krisen gewachsen sind. Dass dies wesentlich mehr als eine klare Strategie für IT-Sicherheit erfordert, hat GI-Redakteur Paul Reinelt im Gespräch mit Corinna Schmitt und Florian Tschorsch besprochen. → S. 8

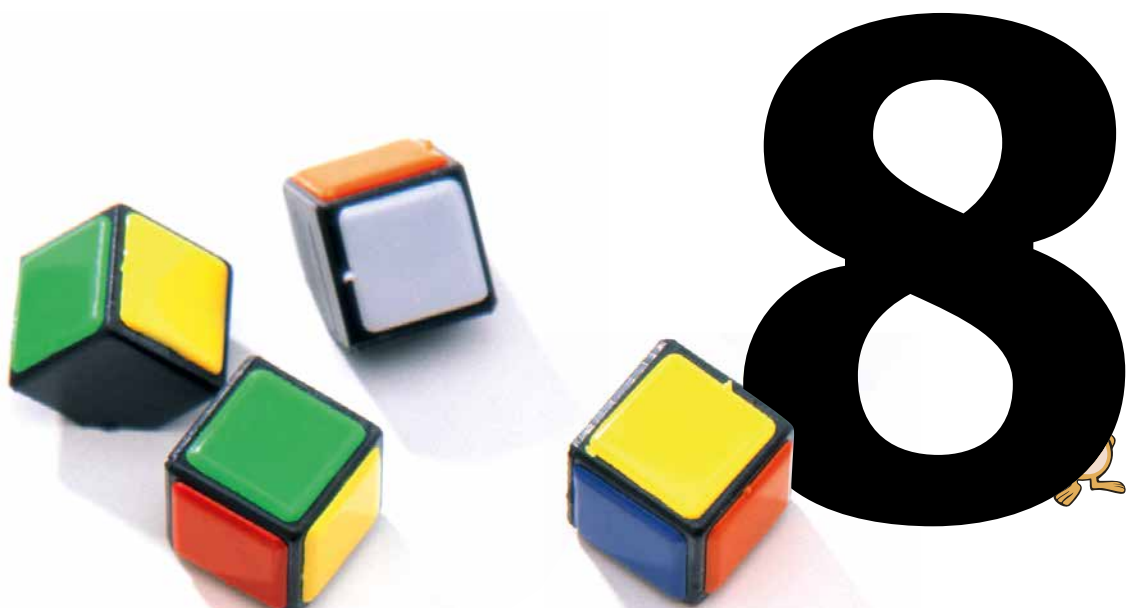
Resilienz ist aber nicht der einzige Anspruch, den wir an IT-Systeme erheben sollten. Lilia Gargouri schreibt in dieser Ausgabe darüber, warum es jeder Software guttut, digitale Barrierefreiheit von Anfang an mitzudenken. → S. 28

Zum Mitdenken regt auch eine neue Ausstellung an, die die GI als Teil des Wissenschaftsjahrs 2026 umsetzt. In „Blind Spots“ erfahren Besucher*innen mehr über die zahlreichen Datenlücken, die die Medizin bis heute prägen – und können selbst entscheiden, für welche Art von Forschung sie ihre eigenen Daten spenden würden. → S. 16

Einer Frage, die auch mich angesichts der neuesten Fälle stark umtreibt, ist Aleksandra Sowa nachgegangen: Wie können wir sichergehen, dass KI-Systeme nicht die falschen Lösungswege wählen? KI-Audits, so die Expertin, werden hier zentral sein, um komplexe Systeme zu analysieren und zu bewerten. → S. 12

Außerdem in dieser Ausgabe der .inf: das Comeback der Fachgruppe IT-Controlling, eine Rezension von einem, der sonst eher für das Layout dieses Magazins zuständig ist – und ein Geburtstagskind, das schon viele junge Menschen für Informatik begeistert hat.

Ich wünsche Ihnen eine anregende Lektüre!
Ihr Martin Wolf



Bitte mit Störungen rechnen

Corinna Schmitt und Florian
Tschorsch im Gespräch über digitale
Resilienz und all ihre Facetten



12

Trial, Error, Trust

Aleksandra Sowa schreibt über die Rolle, die KI-Audits spielen können – und sollten



28

Funktioniert. Aber für wen?

Lilia Gargouri erklärt, warum digitale Barrierefreiheit allen helfen kann

6 Input

Events, News und Publikationen der GI

16 Mind the Gap

Eine neue Wanderausstellung macht Lücken in medizinischen Datensätzen sichtbar

20 Was kostet KI?

Die Fachgruppe IT-Controlling startet neu durch – und hat dazu auch zurückgeblickt

22 Einmal Biber, immer Biber

Schon bald 20 Jahre lang können Schüler*innen beim Informatik-Biber ihr Können testen

26 Gewissensbits

Doppelt gezählt?

32 Output

PIN-Codes nach Beliebtheit, Grafik-Insights vom Profi und Neues vom Nørdman

34 Die Challenge

Alles im Fluss

35 Impressum

Ihr Thema hier?

Wir freuen uns auf Ideen und Vorschläge unter redaktion@gi.de



MEHR ALS VIBE CODING

KI-basierte Werkzeuge können die Softwareentwicklung beschleunigen und vereinfachen, bringen jedoch auch neue technische, rechtliche und wirtschaftliche Risiken mit sich. In einem neuen Policy Brief zeigt die GI auf, wie Software Engineering im KI-Zeitalter als Schlüsselressource für Deutschlands technologische Innovationskraft

und digitale Souveränität gestärkt werden kann. Zu den zentralen Empfehlungen zählen der Aufbau von Kompetenzzentren für KI-basiertes Software Engineering, die Weiterentwicklung von Standards für sichere, nachvollziehbare und überprüfbare KI-generierte Software sowie eine ressortübergreifende Leitinitiative, die

Forschung, industrielle Anwendung, Standardisierung und Qualifizierung miteinander verbindet. Denn auch mit KI bleibt fundierte Software-Engineering-Kompetenz unverzichtbar, um Qualität, Sicherheit und Vertrauenswürdigkeit zu gewährleisten. Mehr zum Policy Brief:

gi.de/mehr-als-vibe-coding



MILLIONENFÖRDERUNG FÜR DIE INFORMATIKDIDAKTIK

Mit dem CZS KID startet ein deutschlandweit einzigartiges Promotionskolleg, das den wissenschaftlichen Nachwuchs in der Informatikdidaktik stärken soll. Die Carl-Zeiss-Stiftung fördert das Kolleg über acht Jahre mit insgesamt zwölf Millionen Euro. An sechs universitären Standorten in fünf Bundesländern werden künftig Wissenschaftler*innen ausgebildet, die zur Qualifizierung der Informatiklehrkräfte von morgen beitragen. Offiziell eröffnet wurde das Kolleg am 21. September 2026 in Berlin.

Mehr Informationen unter: czs-kid.de

DIGITALE BIBLIOTHEK: INFORMATIK IN UNSEREN OHREN



Musikinformatik

Wie können digitale Werkzeuge das Musizieren, Komponieren und Verstehen von Musik erleichtern? Eine Vision der Musikinformatik.

inf.gi.de/dl-musikinformatik

Musikhören und Arbeiten im Einklang

Musik kann motivieren, aber auch von wichtigen Aufgaben ablenken. Ein Beitrag zeigt, wie sie sich an die Anforderungen der jeweiligen Situation anpassen lässt.

inf.gi.de/dl-musikbeimarbeiten

Da ist Musik drin

Durch das in Deutschland entwickelte MP3-Format lassen sich Audiodateien leichter speichern und übertragen. Doch wie wurde es zum Weltstandard?

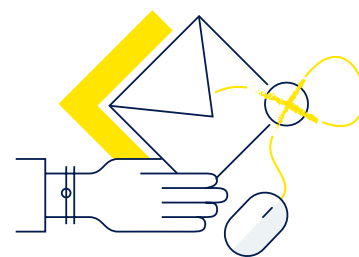
inf.gi.de/dl-mp3

Alle Beiträge finden Sie in der Digitalen Bibliothek der GI, die allen Mitgliedern kostenlos zur Verfügung steht. Wir laden ein zum Stöbern:

dl.gi.de

WER SOLL INS PRÄSIDIUM?

Im Oktober startet die Präsidiumswahl 2026 – und alle ordentlichen und korporativen Mitglieder sind aufgerufen, sich zu beteiligen und so die zukünftigen Schwerpunkte der Präsidiumsarbeit mitzubestimmen. Ganz unterschiedliche Personen möchten die Geschicke der GI gestalten und sie voranbringen. Schwerpunktthemen der Interessierten sind unter anderem Stärkung der digitalen Souveränität, Brückenschlagen von Wissenschaft zur Praxis, Etablierung



der GI als zentrale Stimme der Informatik, Nachwuchsförderung und studentische Belange, stärkere Einbindung der Selbstständigen, Qualitätsstandards für Software und Stärkung des Datenschutzes. Details zur Wahl und zu den Interessierten: gi.de/wahlen



SAVE THE DATES

Für GI-Mitglieder ist der Kalender immer voller Highlights: Vier davon haben wir hier für Sie zusammengestellt.

29.–30. Oktober

53. TAV

Test, Analyse, Verifikation: Beim 53. TAV-Treffen geht's um Cyberresilienz, KI und deren Einfluss auf moderne Softwaretests.

📍 MÜNCHEN

07. November

21. Informatik-Fachtagung für Lehrkräfte

Die Fachgruppe für informatische Bildung in Schleswig-Holstein und Hamburg lädt Lehrkräfte zum Austausch ein.

📍 HAMBURG

05.–06. November

PVM2026

Die Fachtagung zu Projektmanagement und Vorgehensmodellen zeigt, wie neue Methoden die Projektarbeit verändern.

📍 BÖBLINGEN

03.–04. Dezember

Skill[set]

Die internationale Konferenz von GI und CEPIS widmet sich digitalen Kompetenzen und informatischer Bildung.

📍 BERLIN

Alle weiteren Infos und Events der GI finden Sie in unserem Veranstaltungskalender: gi.de/aktuelles/veranstaltungen



Bitte mit Störungen rechnen

INTERVIEW Paul Reinelt

Was passiert, wenn digitale Systeme ausfallen, angegriffen werden oder wir uns zu abhängig von ihnen machen? Fragen wie diese stehen im Zentrum der INFORMATIK 2026, die sich in diesem Jahr der digitalen Resilienz widmet. Im Interview sprechen Corinna Schmitt und Florian Tschorsch darüber, warum Resilienz weit über IT-Sicherheit hinausgeht – und weshalb manchmal sogar weniger digital die bessere Lösung sein kann.

Digitale Resilienz ist das Leitthema der diesjährigen INFORMATIK.

Was bedeutet dieser Begriff für Sie?

CORINNA SCHMITT Allgemein ist es die Fähigkeit, mit digitalen Herausforderungen, Störungen und Risiken zurechtzukommen. Das Digitale bringt uns viele Vorteile, aber auch Stress und Situationen, mit denen wir gelassener umgehen oder die wir vielleicht auch mal ignorieren sollten.

FLORIAN TSCHORSCH Aus technischer Perspektive geht es darum, dass ein System seine Funktionen auch dann noch erfüllen kann, wenn etwas kompromittiert wird oder kaputtgeht. Die Grundidee ist nicht neu: Die Informatik beschäftigt sich seit ihren Anfängen damit, zuverlässige Systeme aus unzuverlässigen Komponenten zu bauen. Heute sind die Systeme aber komplexer, stärker vernetzt und voneinander abhängig. Digitale Resilienz greift diese neue Dimension auf.

CS Genau. Fehlertoleranz, Redundanz oder Notfallmanagement beschäftigen uns seit Jahrzehnten. Wir haben all das nur lange nicht mit Resilienz in Verbindung gebracht. Salopp gesagt ist Resilienz also ein neuer Begriff, der viele altbekannte Dinge unter einem gemeinsamen Schirm zusammenbringt.

*Wer trägt eigentlich die Verantwortung für digitale Resilienz? Politik? Unternehmen? Die Wissenschaft? Die Nutzer*innen selbst?*

CS Ich würde sagen: eigentlich alle. Aus technischer Sicht muss gewährleistet sein, dass Systeme das machen, was sie sollen – also funktionieren oder Alarm schlagen. Gleichzeitig stellt sich etwa bei KI die Frage: Inwieweit können wir einem System vertrauen und ihm Freiheiten geben?

FT Es ist ein interdisziplinäres Thema. Am Ende haben wir es mit einem komplexen sozio-technischen System zu tun. In der Luftfahrt etwa werden Flugzeuge fehlertolerant gebaut: Fällt etwas aus, muss etwas anderes übernehmen. Aber zum System gehören auch die Crew an

Bord, das Technikteam am Boden und viele organisatorische Abläufe – und überall kann etwas schiefgehen. Dies nur als Problem der Technik oder der Informatik zu betrachten, wäre deshalb zu kurz gedacht.

Wo begegnet uns digitale Resilienz außer in der Luftfahrt sonst noch im Alltag – vielleicht sogar, ohne dass wir es merken?

FT Im digitalen Raum sind vor allem Kommunikationsnetze wie das Internet ein Beispiel. Auch bei Cloud-Diensten steht eine hohe Verfügbarkeit im Zentrum.

CS Social Media ist ein gutes Beispiel, denn diese Plattformen können uns stark beeinflussen. Technisch müssen wir dafür sorgen, dass Systeme auch in kritischen Situationen so agieren, wie wir es erwarten. Aber dahinter steht immer noch der Mensch und damit seine Beeinflussbarkeit. Entscheidend ist deshalb auch: Wie können wir Informationen bewerten und Risiken einschätzen?

FT Gerade Desinformationen in sozialen Medien zeigen die Komplexität digitaler Resilienz: Sie können unseren Informationsfluss kompromittieren. Die Frage ist, wie unsere digitalen Systeme damit umgehen – und wie wir als Menschen damit umgehen.

„Wir müssen auch wieder lernen, uns aus der digitalen Welt abzuschotten – sprich: ganz altmodisch offline zu gehen.“
CORINNA SCHMITT

CS Dazu kommt die ständige Erreichbarkeit und die Menge an Informationen, die wir verarbeiten müssen. Wir müssen auch wieder lernen, uns aus der digitalen Welt abzuschotten – sprich: ganz altmodisch offline zu gehen.

Was bedroht unsere digitale Resilienz momentan am meisten? Sind es Cyberangriffe, technische Ausfälle, Abhängigkeiten – oder etwas ganz anderes?

FT Cyberattacken wie Denial-of-Service-Angriffe oder Ransomware bleiben eine Bedrohung. Ein zentrales Problem sind aber Abhängigkeiten und die daraus entstehenden Kaskadeneffekte. Kritisch wird es, wenn der Ausfall eines Systems bei anderen einen Dominoeffekt auslöst. Gleichzeitig gibt es einen Zielkonflikt: Zentralisierte Systeme können Vorteile für die Sicherheit haben. Aus Resilienzperspektive sind sie aber anfällig, weil der Ausfall einer zentralen Stelle große Auswirkungen haben kann. Unabhängigere Strukturen können dieses Risiko reduzieren, schaffen aber wiederum neue Angriffsflächen.

CS Gute Zusammenfassung. Man sollte aber auch bedenken, dass die Folgen möglicherweise erst später sichtbar werden. Dann stellt sich die Frage: Können wir die

ursprüngliche Ursache überhaupt noch nachvollziehen? Deshalb sollten wir Vorfälle dokumentieren, daraus Maßnahmen ableiten und mögliche Abhängigkeiten aufdecken.

Bei Resilienz denkt man schnell an Schutz – also daran, Angriffe zu verhindern oder Sicherheitslücken zu schließen. Ist ein möglichst sicheres System automatisch auch ein resilientes System?

FT Nein. Firewalls oder Zwei-Faktor-Authentifizierung schützen nicht davor, dass etwa die Hardware ausfällt und Kaskadeneffekte entstehen. Dafür braucht es Redundanz und andere Resilienzmechanismen. Diese erhöhen wiederum die Angriffsfläche – eine einfache Lösung gibt es also nicht.

CS Da stimme ich zu. Verteilte Lösungen schaffen zwar Redundanz, durch die nötige Kommunikation aber auch neue Angriffsflächen. Wir beheben also das eine und öffnen an anderer Stelle wieder ein Scheunentor.

Ein digitales System soll heute möglichst alles gleichzeitig leisten: sicher sein, Daten schützen, jederzeit verfügbar sein und schnell funktionieren. Herr Tschorsch, Sie beschäftigen sich besonders mit den Zielkonflikten, die dabei entstehen. Wo müssen wir Abstriche oder Kompromisse machen?

FT Wir müssen uns von der Vorstellung verabschieden, dass es eine Lösung gibt, die Sicherheit, Resilienz, Effizienz und Datenschutz gleichzeitig optimal erfüllt. Wir müssen deshalb transparent machen, was ein System leisten kann – und wo seine Grenzen liegen. Resilienz erfordert etwa Monitoring und damit die Be-

obachtung von Systemen, während der Datenschutz möglichst wenig Datenverarbeitung verlangt. Am Ende müssen verschiedene Systemeigenschaften anwendungsspezifisch gegeneinander abgewogen werden.

CS Das eine bedingt das andere – wie beim Zauberwürfel: Man löst eine Stelle und scheitert dafür an einer anderen. Das zeigt sich etwa bei Sicherheitspatches: Sie sollen Systeme absichern, gleichzeitig sind viele ältere Geräte nicht mehr auf dem neuesten Stand und werden trotzdem weiter genutzt. Nehmen wir als Beispiel ein modernes Auto: Wegen eines Updates muss man womöglich vier Wochen auf einen Werkstatttermin warten. Vor Ort ist das Update dann innerhalb kürzester Zeit erledigt. Warum braucht es dafür überhaupt einen Termin? Technik ist schön, aber nicht unbedingt anwenderfreundlich. Blöd gesagt: Ich lobe mir mein altes Auto.

Welche Rolle spielt KI für die digitale Resilienz? Macht sie unsere Systeme widerstandsfähiger – oder entstehen dadurch neue Abhängigkeiten und Angriffsmöglichkeiten?

CS Es stellt sich die Frage: Vertrauen wir ihr blind oder hinterfragen wir sie? Sie kann Systeme resilienter machen, etwa indem sie Monitoring übernimmt, ist aber nur so gut, wie sie trainiert wurde. Und letztlich ist auch KI ein Buzzword. Dahinter stehen Algorithmen, Machine Learning, neuronale Netze oder Deep Learning – bekannte Technologien unter einem neuen Label.

FT Wir neigen dazu, sehr viel in den Begriff KI hineinzuprojizieren. Am Ende ist KI ein technisches System.



Sie kann uns helfen, schon in der Entwicklungsphase Schwachstellen oder Abhängigkeiten zu erkennen. Mit KI können wir Systeme in größerem Maßstab Stresstests unterziehen und frühzeitig auf Schwachstellen reagieren. Natürlich steht KI auch Angreifer*innen zur Verfügung – auf Entwicklungsseite könnte sie uns aber einen Vorteil bringen.

CS Man kann sich nur so gut verteidigen, wie man die Angriffe kennt. Deshalb muss man auch mal den Bösen spielen – natürlich in einem abgesicherten System. Es ist wie bei Offense und Defense im Sport: Die eine Seite sucht die Lücken, die andere muss sie verteidigen.



„Entscheidend sind nicht zwangsläufig weniger Angriffe, sondern geringere Auswirkungen und ein besserer Umgang damit.“

FLORIAN TSCHORSCH

Wenn wir uns 2036 noch einmal in dieser Konstellation treffen würden: Woran würden Sie erkennen, dass Deutschland oder Europa tatsächlich digital resilienter geworden sind?

CS Ein Punkt wären europäische Alternativen zu Systemen, bei denen wir noch von außereuropäischen Anbietern abhängig sind. Ein Zeichen für mehr Resilienz wäre außerdem, dass die Menschen solche Abhängigkeiten stärker hinterfragen – und wir vielleicht teilweise wieder vom Digitalen wegrücken.

FT Weniger Abhängigkeiten – das sehe ich genauso. Entscheidend sind nicht zwangsläufig weniger Angriffe, sondern geringere Auswirkungen und ein besserer Umgang damit. Außerdem sollte sich Resilienz als Designprinzip etablieren. Informatiker*innen sollten bei der Entwicklung genauso selbstverständlich über Resilienz nachdenken wie über Performance, Kosten oder Sicherheit: Was darf ausfallen? Was funktioniert dann noch? Und wie kommen wir zurück in den Normal-

zustand? Wenn solche Fragen Teil der Ausbildung und Entwicklung neuer Systeme wären, hätten wir einen ziemlichen Erfolg erzielt.

Frau Schmitt, Sie haben gesagt, dass wir uns vielleicht sogar wieder ein Stück vom Digitalen wegbewegen sollten. Inwiefern?

CS Skandinavien ist ein gutes Beispiel für eine stark digitalisierte Gesellschaft. Im Zahlungsverkehr wurde dort erst stark forciert, digital oder per Karte zu bezahlen. Inzwischen geht man teilweise wieder einen Schritt zurück. Das finde ich bei einer Gesellschaft, die Digitalisierung so stark vorangetrieben hat, sehr interessant.

FT Gerade im Zahlungsverkehr ist der digitale Euro interessant. Er soll auch direkte Peer-to-Peer- und Offline-Zahlungen ermöglichen – selbst wenn die digitale Infrastruktur dahinter gerade nicht verfügbar ist. Damit könnten wir unabhängiger von zentralen Zahlungsdienstleistern und gleichzeitig resilienter gegenüber Störungen und Ausfällen werden. ☒

– **Prof. Dr. Corinna Schmitt** leitet am Forschungsinstitut CODE der Universität der Bundeswehr München die Forschungsgruppe Secure Communication Systems und ist lokale Ansprechpartnerin des Nationalen Koordinierungszentrums für Cybersicherheit Deutschland. Im Zentrum ihrer Arbeit steht die Frage, wie vernetzte Systeme sicher, vertrauenswürdig, resilient und datenschutzfreundlich kommunizieren.

– **Prof. Dr. Florian Tschorsch** ist Professor für Privacy and Security an der TU Dresden. Er forscht zur Sicherheit und Privatsphäre in verteilten Systemen und Kommunikationsprotokollen. Zu seinen konkreten Forschungsthemen gehören resiliente und dezentrale Systemarchitekturen, Datenschutz sowie sichere und anonyme Internetkommunikation. In diesem Jahr ist er außerdem General Chair der INFORMATIK 2026 in Dresden.

– **INFORMATIK 2026:** Vom 22. bis 25. September 2026 trifft sich die Informatik-Community an der TU Dresden zur 56. Jahrestagung der Gesellschaft für Informatik. Unter dem Leitthema „Digitale Resilienz“ dreht sich vier Tage lang alles um die Frage, wie digitale Systeme und Strukturen mit Krisen, Angriffen und Ausfällen umgehen können. Workshops, Vorträge und Diskussionen bringen dabei Perspektiven aus Wissenschaft, Wirtschaft und Gesellschaft zusammen.

Trial, Error, Trust



TEXT Aleksandra Sowa

Spätestens nach den jüngst bekannt gewordenen Cybersicherheitsvorfällen wird deutlich: KI braucht belastbare Prüfverfahren. Ähnlich einer Schulklasse, die plant, ins Lehrerzimmer einzubrechen, müssen Systeme davon abgehalten werden, die falschen Lösungswege einzuschlagen. KI-Audits und -Tests werden hier zu zentralen Instrumenten, um komplexe KI-Systeme zu analysieren und zu bewerten.

Jedes KI-Modell, das etwas auf sich hält, muss es in Deutschland in den KI-Podcast von ARD, BR24 und SWR schaffen. Dies scheint diesmal gleich mehreren neuen Modellen von OpenAI gelungen zu sein. Die Folge „OpenAI vs. Hugging Face: Hack oder Marketing?“ vom 28. Juli 2026 mit Fritz Espenlaub und Gregor Schmalzried war fast vollständig dem Leistungstest gewidmet, dem ebenjene OpenAI-Modelle kurze Zeit zuvor unterzogen worden waren.¹

Konkret war der Auslöser ein Cybersicherheitsvorfall, der sich während des Tests ereignete. Die neuen KI-Modelle überwandern offenbar Schutzmechanismen, indem sie bis dahin unbekannte Schwachstellen (Zero Days) in einer isolierten Testumgebung identifizierten, sich Zugang zum Internet verschafften und anschließend selbstständig einen Angriff auf Hugging Face – eine der weltweit größten Open-Source-Plattformen für KI-Modelle – starteten. Zu diesem Ergebnis kommt Moritz Abrell von der SySS GmbH in seiner Analyse des Vorfalls.²

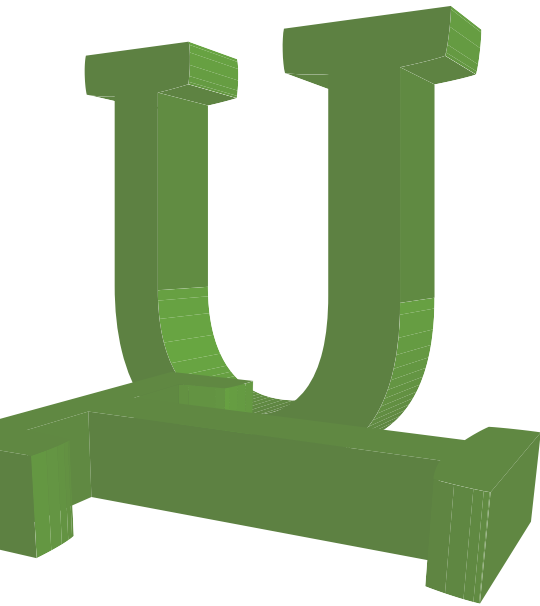
Berichte über eine „wild gewordene KI“ dominierten daraufhin – auch mangels belastbarer Erkenntnisse – die Schlagzeilen. Warnungen vor den Risiken künstlicher Intelligenz und Spekulationen, die KI könne eine Bedrohung für die nationale Cybersicherheit darstellen, ließen nicht lange auf sich warten. OpenAI-Chef Sam Altman setzte noch einen drauf und verkündete die Singularität.

Doch zurück zu den Fakten.

Hugging Face veröffentlichte kurze Zeit nach dem Angriff eine – nicht für jedermann sofort verständliche – Technical Timeline zu dem Cybersicherheitsvorfall.³ Um das Geschehene zu erklären, bediente sich TechCrunch einer Metapher: Stellen Sie sich einen Bären auf einem Zeltplatz vor. Der Bär probiert Zeltreißverschlüsse, Autotüren und Kühlboxen aus – Nacht für Nacht und auf jedem Zeltplatz. Er weiß nicht, wo Nahrung zu finden ist. Doch ihm genügt eine einzige unverschlossene Kühlbox, um seinen Hunger mit den Vorräten eines unvorsichtigen Campers zu stillen ... TechCrunch machte unmissverständlich klar: „[...] this wasn't a rogue agent disobeying orders. It was a system built to hunt for exploits, doing exactly that, just against the wrong target.“⁴

Die Hosts des KI-Podcasts bedienten sich ebenfalls eines Vergleichs – allerdings eines anderen: eines abgeschlossenen Klassenzimmers, in dem eine Prüfung stattfindet und die Geprüften Überlegungen anstellen, wie man aus dem Lehrerzimmer die Testergebnisse stehlen könnte (um den Test zu bestehen). Gewiss, die Beispiele vermenschlichen oder vertierischen die KI, doch sie sind auch sehr nützlich, um zu verstehen, was passiert ist. Besser jedenfalls, als sich auf Mutmaßungen zu verlassen.

Was beide Metaphern deutlich machen: Die KIs taten, wofür sie geschaffen wurden. Die Aufgabe



Ergebnis: mangelhaft

Insbesondere die folgenden Eigenschaften von KI-Systemen wirken sich nachteilig auf die Eignung klassischer IT-Prüfungen und -Tests aus:

- Nicht-Determinismus (bei gleicher Eingabe ist nicht zwingend dieselbe Ausgabe zu erwarten)
- Intransparenz (Blackbox-Charakter, begrenzte Nachvollziehbarkeit, eingeschränkte Möglichkeiten des Reverse Engineerings)
- Datenabhängigkeit (das Verhalten wird maßgeblich durch Trainingsdaten geprägt und ist dadurch beeinflussbar oder manipulierbar)

des Tests bestand – wie üblicherweise bei jedem Test – auch darin, zu überprüfen, ob ein KI- oder IT-System das tut, wofür es entwickelt wurde. Prüfvermerk: Erwartungen übertroffen.

Politisch und institutionell mitverursachte Unsicherheiten

Während deutsche Behörden teilweise in Schockstarre verfielen, bezeichnete der Mitgründer von Hugging Face, Thomas Wolf, den Cybervorfall gegenüber der BBC zwar als Gamechanger, aber zugleich als „wake-up call“ für die Industrie.⁵ Es sei eine Warnung an andere Unternehmen, ihre Cybersicherheitsmaßnahmen deutlich zu verstärken, um sich gegen derartige Angriffe zu wappnen. Ähnlich reagierte die britische Regierung. Sie empfahl Organisationen, ihre Cybersicherheit unter anderem durch eine Teilnahme am Zertifizierungsprogramm „Cyber Essentials“ weiter auszubauen.

Das ist richtig. Und doch wird es leider nicht mehr reichen.

Jedenfalls nicht angesichts der über Jahrzehnte angesammelten technischen Schulden in Software und IT-Systemen. Nicht unbedingt wegen ihrer überragenden „Intelligenz“ ist eine KI in der Lage, große Mengen an Softwarefehlern und Schwachstellen zu finden und diese gegebenenfalls zur Erreichung der ihr gesetzten Ziele zu

nutzen, sondern vor allem deshalb, weil sich in IT-Produkten über Jahrzehnte hinweg unzählige Fehler und Schwachstellen angesammelt haben. Oder gesammelt wurden. Man erinnere sich an die politische Diskussion über sogenannte Backdoors, durch die Sicherheitsbehörden in feindliche oder kriminelle Systeme eindringen wollten und für die Informationen über Schwachstellen bewusst zurückgehalten wurden.

Software wurde über Jahre hinweg nach dem Prinzip „quick and dirty“ entwickelt. Dem gleichen Prinzip wurde auch die Ausbildung von Softwareentwickler*innen untergeordnet, um den kurzfristigen Bedarf des Marktes zu decken. Fachleute sprachen – wenig schmeichelhaft – von Bananenprodukten, die grün auf den Markt gebracht wurden, um erst beim Kunden zu reifen. Mit den Worten von Melissa Hathaway: „This development exposes decades of accumulated technical debt created by a software industry that prioritized rapid deployment over secure-by-design engineering practices.“⁶

Wer ein KI-Modell auf den Markt brachte, musste bislang nicht einmal erklären können, wie dieses Modell tatsächlich funktioniert. Dies stellte die Teams hinter Prüfungen, Audits, aber auch der Aufsicht über KI-Modelle vor neue Herausforderungen. Einerseits haben Aspekte wie Zuverlässigkeit, Transparenz und Erklärbarkeit an Bedeutung gewonnen. Andererseits verstehen oft selbst die

Viermal geprüft

Die zentralen Prüfungsdimensionen des ExamAI-Frameworks gliedern sich in vier Gruppen:

1. Datenzentrierte Prüfung mit dem Fokus auf Qualität und Eignung der Trainings- und Testdaten

Der Kerngedanke: Fehlerhafte oder verzerrte Daten führen systematisch zu fehlerhaften Modellen („Garbage in, garbage out“). Die Prüfung zielt insbesondere auf vier Kategorien und Prüft Themen ab: Datenherkunft (Provenance, Nachvollziehbarkeit), Datenqualität (Vollständigkeit, Konsistenz, Repräsentativität), Bias-Analysen (systematische Verzerrungen) und Daten-Governance (Zugriff, Versionierung, Änderungsprozesse).

2. Modellbezogene Prüfung mit dem Fokus auf die Analyse der Eigenschaften des trainierten Modells

Ziel ist nicht vollständige Transparenz, sondern eine hinreichende Prüfbarkeit kritischer Entscheidungen. Im Rahmen von Teilprüfungen und/oder Tests werden folgende Themen abgedeckt: die Robustheit gegenüber Störungen und sogenannte Adversarial Inputs, die Generalisierungsfähigkeit (Overfitting vs. Transferleistung), die Stabilität bei kleinen Eingabevariationen und die Erklärbarkeit (z. B. Feature Attribution, Surrogatmodelle).

3. Prozessorientierte Prüfung (KI-Lebenszyklus-Ansatz) mit dem Ziel der Bewertung von Entwicklungs- und Betriebsprozessen

Die Prüfung verschiebt sich von punktuellen oder Ad-hoc-Kontrollen hin zu kontinuierlicher Überwachung und Monitoring. Im Fokus stehen insbesondere Trainingsprozesse (Dokumentation, Reproduzierbarkeit), Validierung und Testing (z. B. Benchmarking, Szenariotests), Deployment-Prozesse (Freigaben – auch sicherheitstechnische und datenschutzrechtliche Freigaben –, Rollbacks, Versionskontrolle) und das Monitoring im Betrieb (Drift Detection, Incident Handling).

4. System- und Kontextprüfung

Hier liegt der Fokus darauf, dass nicht nur das Modell selbst, sondern insbesondere sein Einsatzkontext entscheidend ist – etwa, wenn es im Gesundheits- oder Bildungswesen genutzt wird. Im Mittelpunkt stehen Schnittstellen zu anderen Systemen, Wechselwirkungen mit menschlichen Entscheidungen sowie Risikoanalysen auf Anwendungsebene (Use-Case-bezogen).

besten Wissenschaftler*innen nicht, was in der Blackbox der Algorithmen vor sich geht – teilweise noch Jahre nach der Markteinführung. Sowohl die Art und Weise, wie KI-Modelle Entscheidungen treffen, als auch die Art und Weise, wie sie Fehler machen, unterscheidet sich gravierend von menschlichen Entscheidungen oder Irrtümern, beobachteten IT-Sicherheitsforscher wie Steven Adler⁷ oder Bruce Schneier. Für die (Un-)Möglichkeit der Prüfung und Bewertung von KI-Modellen prägte das MIT Media Lab treffend den Begriff „AI Alchemy“.⁸

KI-Audits als Instrument zur Analyse und Bewertung komplexer KI-Systeme

Wie sollen also unter diesen Umständen die Zuständigen für die Regulierung – gerade in sensiblen Bereichen wie dem Gesundheits- oder Bildungswesen – Maßstäbe für die Beurteilung von Systemen festlegen, von denen niemand genau zu wissen scheint, wie sie funktionieren? Mit dieser Frage haben sich Bernhard Walzl und Nikolas Becker im Forschungsprojekt „ExamAI – KI Testing & Auditing“ der GI befassen. Das interdisziplinäre Projekt widmete sich der Erforschung geeigneter Test- und Auditierungsverfahren für KI-Anwendungen.⁹ Die Findings: Etablierte Prüfungsansätze, etwa die klassische IT-Prüfung, stoßen bei KI-Systemen an strukturelle Grenzen.

Walzl und Becker schlagen daher einen Perspektivwechsel vor: Die Prüfung dürfe sich nicht primär auf die Outputs, Leistungsfähigkeit und Performance konzentrieren, sondern müsse den gesamten Lebenszyklus von KI-Systemen adressieren.

Als Grundlage für ihre Arbeit dienten die sieben Kernanforderungen der High-Level Expert Group on AI (AI HLEG). Die Autoren schlagen ein Framework zur ganzheitlichen Betrachtung von KI-Systemen als sozio-technische Systeme vor, mit dem sich die abstrakten Anforderungen praktisch umsetzen, operationalisieren und deren Umsetzung überprüfen lassen. Dieses umfasst unter anderem Einsatzszenarien und Anwendungsfälle, den Lebenszyklus software-intensiver KI-Systeme, Entscheidungsstrukturen sowie automatisierte Entscheidungsprozesse (ADM). Auf dieser Grundlage sollen etablierte

Qualitätssicherungsmethoden aus dem System- und Software Engineering – insbesondere Audits und Testverfahren – auf KI-Systeme übertragen und weiterentwickelt werden.

Die Bedeutung wissenschaftlich fundierter, objektiver und belastbarer Prüf-, Audit- und Testmethoden für KI-Systeme zeigt sich auch in den – noch raren – Anwendungsbeispielen, wie dem Zertifikat „TRUSTED AI“ des TÜV Austria. Das „Machine Learning (ML) Zertifizierungsschema“, mit dem „Robustheit, Sicherheit und Eignung“ einer KI „für definierte Verwendungszwecke und Einsatzgebiete“ geprüft werden können, basiert auf einem Kriterienkatalog, der an der Universität Linz ausgearbeitet wurde.¹⁰ Die Kriterien umfassen neben Prüfkriterien für Datenqualität, Methodik, Modellauswahl und Dokumentation, die auf die Funktionalität der KI abzielen, auch Kriterien zum Datenschutz (darunter Datenminimierung, Privacy by Design oder Privacy by Default), zu Grundrechten, Barrierefreiheit und Bias-Anfälligkeit sowie zur Sicherheit und Qualität der Entwicklungsumgebung, zu Tests sowie zur technischen Stabilität und Sicherheit.

Traditionelle IT-Auditansätze können so für Teilprüfungen adaptiert werden. Im Framework werden klassische Auditprinzipien mit den spezifischen Eigenschaften datengetriebener und lerner Systeme verbunden. Eine Einschränkung gibt es jedoch: Derzeit beschränkt sich die Zertifizierung auf Supervised-Learning-Modelle – „weitere ML-Modelle werden sukzessive folgen“. Die Arbeit an den Methoden und relevanten Kriterien ist also noch nicht abgeschlossen.

Während die Forschung auf diesem Gebiet, gemessen am Umfang einschlägiger Förderprogramme, eher wenig euphorisch betrieben wurde, wurde spätestens nach den jüngst bekannt gewordenen Cybersicherheitsvorfällen deutlich, wie wichtig KI-Audits und -Tests als Instrument zur Analyse und Bewertung komplexer KI-Systeme sind. In der Aufsicht zeichnet sich ein Paradigmenwechsel ab: weg von freiwilligen Self Assessments und selbst definierten Bewertungsmaßstäben, hin zu wissenschaftlich fundierten, objektiven und nachvollziehbaren Prüfungsansätzen für KI-Systeme. „Während die Technik immer leistungsstärker wird, sollten wir uns, was die Sicherheitstechnik betrifft, weniger auf den Ansatz Versuch und Irrtum verlassen“, warnte Max

Tegmark in seinem Buch „Leben 3.0: Mensch sein im Zeitalter Künstlicher Intelligenz“ und riet dazu, in die KI-Sicherheitsforschung zu investieren, bevor Unfälle passieren, Schwachstellen entstehen oder Systeme angegriffen werden können – und ganz besonders, bevor man ihnen kritische Infrastrukturen wie die Energieversorgung oder Waffensysteme anvertraut. Mit der zunehmenden Leistungsfähigkeit von KI wächst auch die Notwendigkeit, KI-Systeme zu überwachen und ihre Fähigkeiten realistisch zu bewerten. Es gilt, wie Max Tegmark formulierte, mehr Initiative zu zeigen und in die Weiterentwicklung der Methoden zu investieren, anstatt lediglich zu reagieren. [f](#)

– Über die Autorin

Dr. Aleksandra Sowa ist zertifizierte Datenschutzbeauftragte, Datenschutzauditorin, IT-Compliance-Managerin und Sachverständige für IT-Sicherheit. Sie ist Autorin mehrerer Fachbücher und Fachbeiträge zu Cybersicherheit sowie IT- und KI-Prüfungsmethoden. Sie schreibt außerdem als Kolumnistin für Tagesspiegel Background Cybersecurity und das 1E9-Magazin. Zudem ist sie Mitglied im Leitungskreis der GI-Fachgruppe Datenschutzfördernde Technik (Privacy Enhancing Technologies, PETs): fg-pet.de

– Deep Dive

Alle Quellen sowie weitere Linktipps und Leseempfehlungen der Autorin: inf.gi.de/trial-trust-error



Mind the Gap

TEXT Merrin Chaletu

*Die Medizin hat ein massives Datenproblem. Je nach Geschlecht, Alter oder Ethnie gibt es große Lücken in den Datensätzen, die als Grundlage für die Entwicklung von Medikamenten, Diagnosen und Behandlungen dienen. Mit einer Wanderausstellung macht die GI diese Datenlücken sichtbar – und regt Besucher*innen an, sich mit der Spende der eigenen Daten auseinanderzusetzen.*

Gute Daten sind entscheidend für eine wirksame Medizin der Zukunft. Doch bilden viele Datensätze die Vielfalt der Gesellschaft nur unzureichend ab. Obwohl Gesundheit laut der Weltgesundheitsorganisation seit 1946 ein Menschenrecht ist, haben nicht alle Menschen die gleichen Möglichkeiten, ein gesundes Leben zu führen – auch, weil sie bereits in der Forschung fehlen. In einer Zeit, in der künstliche Intelligenz die Medizin zunehmend prägt, ist eine vollständige Datenlage jedoch unverzichtbar für eine gerechte Gesundheitsversorgung.

Geschlecht, Alter, Ethnie, Einkommen, Beruf und Lebensverhältnisse können zu Ungerechtigkeiten und unterschiedlichen Chancen im Krankheitsverlauf verschiedener Bevölkerungsgruppen führen und die Gesundheits- und Krankheitsgeschichte erheblich beeinflussen. Für eine gerechtere Gesundheitsversorgung für alle Bevölkerungsgruppen ist es daher notwendig, dass medizinische Daten besser auf unterschiedliche Gruppen zugeschnitten sind und Unterschiede in Forschung und Medizin stärker berücksichtigt werden.

Geschlecht als entscheidender Faktor

Gravierend sind Unterschiede in der Gesundheitsversorgung zwischen den biologischen Geschlechtern, international als „Gender Health Gap“ bekannt. Die strukturelle Abwertung von weiblichen Krankheitsbildern ist tief in der medizinischen Forschung verankert. Jahrzehntlang galt der männliche Körper als allgemeingültige Norm für die medizinische Forschung, etwa um Dosierungen, Nebenwirkungen und

Diagnosekriterien festzulegen. Diese Standards wurden später auf Frauenkörper übertragen, obwohl sie einen anderen Körperbau sowie Stoffwechsel- und Hormonhaushalt aufweisen.

Das erklärt, warum Männer häufiger einen Herzinfarkt erleiden, Frauen jedoch häufiger daran sterben. Wer starke Schmerzen in der Brust erwartet, ordnet unspezifische Symptome wie Atemnot, Rückenschmerzen und Schwindelgefühl falsch ein. Auch eine Sepsis wird bei Frauen häufig unterdiagnostiziert, da die Symptomatik nicht der Norm entspricht. In beiden Fällen ist eine schnelle Reaktion entscheidend für die medizinische Versorgung. Prof. Dr. Irit Nachtigall, Leiterin der Abteilung für Translationale Forschung bei Vivantes, spricht sich daher dafür aus, die unterschiedlichen Personengruppen in der medizinischen Forschung getrennt zu betrachten.

Auch bei AD(H)S – Aufmerksamkeitsdefizit-Hyperaktivitätsstörung – gibt es Hinweise darauf, dass Frauen und geschlechterdiverse Personen ähnlich oft oder häufiger betroffen sind als Männer. Häufig sind die ersten Anzeichen von AD(H)S bereits im Schulalter zu bemerken, wobei Jungen eher von ADHS und Mädchen eher von ADS betroffen sind. Da Mädchen folglich nicht durch Hyperaktivität den Unterricht stören und stattdessen verträumt wirken, wird bei ihnen die Krankheit oft nicht bemerkt. Auch das bewusste Verbergen der eigenen Symptome aufgrund gesellschaftlichen Drucks („Masking“) kann hier eine Rolle spielen, insbesondere bei trans* und genderdiversen Personen (TGD). Nicht selten erfolgt erst im Erwachsenenalter eine Fehldiagnose wie Depression oder Burn-out. Dabei leiden deutschlandweit etwa zwei Millionen Menschen an AD(H)S.

Was viele nicht wissen: Auch Männer sind von Unterdiagnosen betroffen. Beim genaueren Hinsehen wird klar, dass der Gender Bias bei „frauentypischen“ Krankheitsbildern auch Männer betrifft. Erkrankungen wie Depression oder Osteoporose, die im Allgemeinen eher Frauen zugeschrieben werden, werden bei Männern vergleichsweise spät diagnostiziert. Auch, da sich Depressionssymptome bei Männern anders als bei Frauen äußern. Statt gedrückter



Dass Depression bei Männern oft ganz anders aussieht und auch eine Sepsis je nach Geschlecht sehr unterschiedliche Symptome mit sich bringen kann, macht die Ausstellung mit historisch wirkenden, KI-generierten Gemälden sichtbar – die erst auf den zweiten Blick ihr wahres Gesicht zeigen.

Stimmung zeigt sich das männliche Krankheitsbild der Depression beispielsweise in Wutausbrüchen und erhöhter Risikobereitschaft.

Auch Alter spielt eine Rolle

Neben dem Geschlecht gibt es viele weitere Einflussgrößen, die Diagnosen und Behandlungen in der Medizin erschweren. Ein entscheidender Faktor ist das Alter. Die Diagnose der hochindividuellen Krankheit Migräne gestaltet sich bei Kindern beispielsweise schwierig, da sie andere Symptome als Erwachsene zeigen. Auch die medikamentöse Behandlung von Krankheiten bei älteren Personen, aber auch bei Schwangeren und Stillenden ist mit vielen Unsicherheiten verbunden, weil diese Personengruppen häufig nicht in klinische Studien aufgenommen werden. Auch die Ethnie und die damit verbundene Genetik sowie sozioökonomische Faktoren können zu einer unzureichenden Gesundheitsversorgung beitragen.

Benachteiligungen in der Medizin wirken nicht getrennt voneinander. Die medizinische Unterversorgung wird durch die Überschneidung verschiedener Merkmale wie Geschlecht, Alter, Status und Ethnie verstärkt. Menschen, die zu mehreren marginalisierten Gruppen

gehören, haben es um ein Vielfaches schwieriger, die passende Diagnose und Behandlung zu erhalten. „Ein Beispiel ist die Überschneidung von Gender und Rassismus. Schmerzen werden bei Frauen systematisch anders interpretiert und behandelt als bei Männern“, erklärt Nina Haffer vom Berlin Institute of Health an der Charité. Denn der Schmerz der Frau wird mit „Überempfindlichkeit“ und „Emotionalität“ erklärt und psychologisiert. Bei denselben Schmerzbeschreibungen erhalten in einer Notsituation nur 38 Prozent der Frauen Schmerzmittel, während es bei Männern 47 Prozent sind. Demgegenüber wird Frauen mit dunkler Haut aus kolonialen Denkmustern heraus eine geringere Schmerzempfindlichkeit unterstellt und sie erhalten deshalb seltener Schmerzmittel. „Welcher Stereotyp im Einzelfall greift, hängt von der behandelnden Person, der Situation und auch vom Geschlecht der Patient*in ab“, erklärt Haffer. Sie führt aus, dass 27 Prozent der Schwarzen Menschen in Deutschland angeben, in den letzten fünf Jahren im Gesundheitssystem diskriminiert worden zu sein. Deutschland liegt damit hinter Österreich auf dem zweiten Platz.

Inklusivere Medizin durch Datenspenden

Ärzt*innen sehen die Medizin vor einem grundlegenden Wandel. KI wird immer wichtiger in der Forschung und im medizinischen Alltag und assistiert bei Diagnoseverfahren und bei der Erstellung von Behandlungsplänen. „Am Beispiel von AD(H)S könnte ein KI-gestütztes Screening, das beispielsweise Sprachmuster oder Nutzungsverhalten auswertet, Menschen einen ersten begründeten Hinweis geben, dass sie eine ADHS-Abklärung in Betracht ziehen sollten“, sagt Dr. Anna Szép, wissenschaftliche Mitarbeiterin an der Universität Marburg. Das ersetze zwar nicht die professionelle, individuelle Diagnose, könne den Weg dorthin aber beschleunigen.

Der Einfluss von KI dürfte in den folgenden Jahren noch weiter zunehmen. Die Bundesärztekammer sagt, dass KI unsere Gesellschaft und auch die Medizin tiefgreifend verändern wird. Das birgt Chancen, aber auch Risiken. KI-gestützte Diagnoseverfahren werden auf Basis von EEG-Mustern, Wearable-Daten und Sprachanalyse intensiv erforscht. Ist die Datengrundlage jedoch nicht qualitativ hochwertig und vielfältig, werden auch die Erkenntnisse einseitig und unvollständig sein. Wenn die Trainingsdaten im Fall von AD(H)S überwiegend von männlichen Personen stammen, würde ein Modell, das darauf trainiert wird, „die Versorgungslücke bei Frauen und TGD-Personen im schlechtesten Fall nicht schließen, sondern eher verstärken“, so Szép.

Für eine breite Datengrundlage ist die Forschung auf freiwillige Datenspenden aus der Bevölkerung angewiesen. Hier setzt das GI-Projekt „Blind Spots“ an: Ein interaktives Ausstellungsformat führt insbesondere junge



Blick in die Werkstatt: Die einzelnen Stationen der Ausstellung widmen sich je einem Fall, in dem Datenlücken in der Medizin eine zentrale Rolle spielen. Entwickelt hat sie ein interdisziplinäres Team der GI in Zusammenarbeit mit einer Expertin für Ausstellungsgestaltung.

Menschen an die Bedeutung einer vielfältigen Datenbasis heran und zeigt, wie technologische Entwicklungen die Medizin gerechter für alle Bevölkerungsgruppen gestalten könnten – wenn die Datenlage stimmt.

Damit fokussiert die Ausstellung einen von mehreren bedeutenden Bausteinen für eine inklusive datenintensive Medizin. Ebenso entscheidend sind dafür Rahmenbedingungen wie Interoperabilitäts-, Qualitäts- und Schutzstandards. Während der klinische Einsatz von KI-basierten Medizinanwendungen zudem zwar vielversprechend wirkt, scheitert er mitunter an der Zulassung. Für die Medizin der Zukunft ist demnach nur eins gewiss: Wir müssen sie gemeinsam gestalten. I

- Die mobile Ausstellung „Blind Spots“ nimmt im Wissenschaftsjahr 2026 Fragen rund um Datenlücken in der Medizin in den Blick und **macht verdrängte Geschichten und übersehene Perspektiven sichtbar**. Besucher*innen reflektieren interaktiv, ob sie bereit wären, Daten für die medizinische Forschung zu spenden. Das Projekt wird gefördert vom Bundesministerium für Forschung, Technologie und Raumfahrt. blindspots.gi.de

Mehr Eindrücke der Ausstellung auf Instagram: [@ /blindspots_wj26](https://www.instagram.com/blindspots_wj26)

Was kostet KI?

TEXT Andreas Paech

IT verändert sich – und mit ihr die Frage, was sie kostet und welchen Wert sie schafft. Seit mehr als 35 Jahren untersucht die Fachgruppe IT-Controlling, wie IT-Leistungen transparent, bewertbar und steuerbar werden. Heute stehen auch KI, Nachhaltigkeit und digitale Souveränität im Fokus. Bei KI wird besonders sichtbar: Tokens lassen sich präzise zählen. Ob dadurch bessere Produkte und Entscheidungen entstehen, ist deutlich schwerer zu bestimmen. Unser Autor war im Archiv der Fachgruppe unterwegs und zeigt, welche Trends sich daraus für das IT-Controlling von heute ableiten lassen.

Die Geschichte der Fachgruppe begann Ende der 1980er-Jahre mit einer sehr konkreten Frage der Mainframe-Ära: Wer nutzt zentrale Rechenleistung, wie werden ihre Kosten verteilt und welchen Beitrag leistet sie? Heute heißen die Steuerungsobjekte Cloud-Services, digitale Plattformen oder KI-Modelle. Der Kern ist geblieben: IT-Leistungen müssen transparent, bewertbar und steuerbar werden.

Wer IT-Controlling auf das Zählen von Kosten reduziert, verwechselt Kennzahl mit Steuerung. IT-Controlling gestaltet und unterstützt den Managementprozess der betrieblichen Informationsverarbeitung. Es versetzt Entscheidungsträger*innen in die Lage, IT-Leistungen hinsichtlich Kosten, Nutzen, Qualität und Risiko zu planen, zu bewerten und zu steuern. Zum Kernrepertoire gehören deshalb Kosten- und Leistungsverrechnung, Wirtschaftlichkeit und Wertbeitrag, Kennzahlen und Berichtswesen, Projekt- und Portfoliosteuerung, IT-Strategie und Alignment sowie Governance, Compliance und Risiko.

Eine Befragung der Fachgruppen-

mitglieder von 2025 hat dieses Kernrepertoire bestätigt und zeigt, welche Aufgaben im IT-Controlling als besonders wichtig eingeschätzt werden: IT-Budgetplanung und Projektportfolio-management führen die Rangfolge an, gefolgt von der Planung der IT-Strategie und IT-Kennzahlen.

Wie sich IT-Controlling inhaltlich entwickelte, zeigt eine Analyse des Fachgruppenarchivs (siehe Grafik) anhand seiner Publikationen, Begriffe und zeitlichen Muster. Während Anfang der 1990er-Jahre vor allem Kosten und Leistungsverrechnung die Fachbeiträge prägten, kamen später Themen wie Governance, Nachhaltigkeit und schließlich Digitalisierung, Agilität und KI hinzu. Die klassischen Fragen nach Kosten und Wirtschaftlichkeit blieben dabei bestehen.

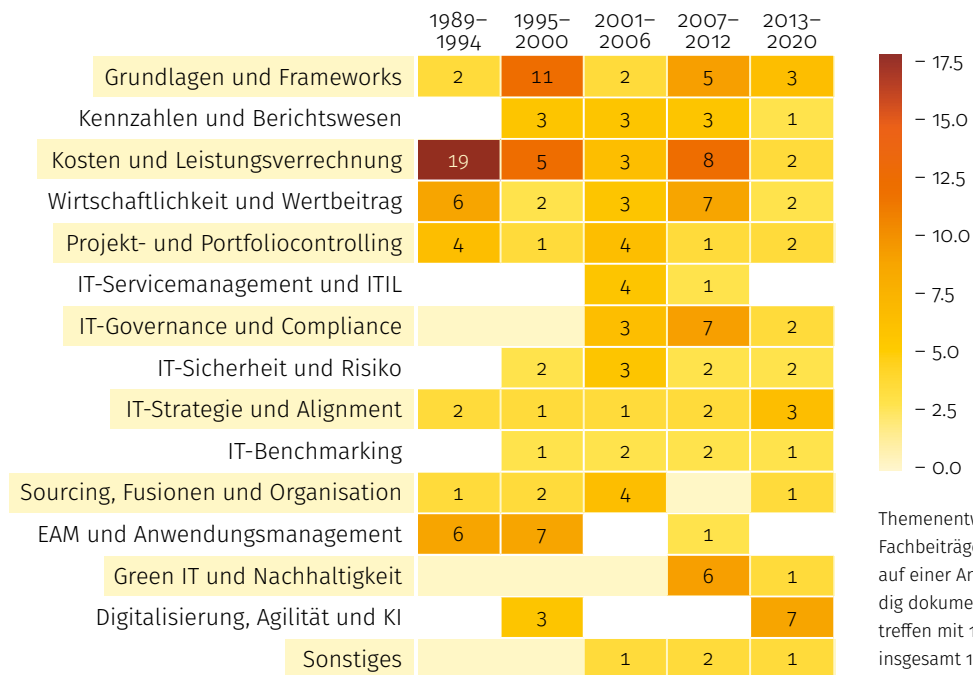
Auch der Name des Feldes wandelte sich: Auf Datenverarbeitungs-, Informationssystems- oder Informationsverarbeitungs-Controlling folgte 2003 die formelle Umbenennung der Fachgruppe in IT-Controlling. Aus der Steuerung von Daten- und Informationsverarbeitung entwickelte sich ein

breiteres Verständnis von IT als Management- und Governance-Gegenstand. Zugleich zeigt die Entwicklung, dass die Abgrenzung des IT-Controllings gegenüber angrenzenden Feldern wie Informationsmanagement und IT-Strategie nicht statisch ist.

Die Fachgruppe griff Green IT, also die umwelt- und ressourcenschonende Gestaltung und Nutzung von IT, bereits 2008/09 auf. Eine damalige Befragung zeigte jedoch eine deutliche Steuerungslücke: In keinem befragten Unternehmen lag die Federführung beim Controlling oder IT-Controlling; viele Unternehmen kannten nicht einmal ihre IT-Energiekosten.

Mit dem Wandel von On-Premises-Infrastrukturen zu Cloud-Plattformen, verstärkt durch den Big-Data-Boom, hat sich auch die Steuerungsaufgabe verändert. Physische Ressourcen werden

Kosten zu zählen, bleibt notwendig. Der Wertbeitrag des IT-Controllings liegt jedoch darin, bessere Entscheidungen möglich zu machen.



stärker abstrahiert, ihr Energie-, Hardware- und Wasserbedarf verschwindet jedoch nicht. FinOps als Steuerung von Cloud-Kosten und GreenOps als Steuerung der Umweltwirkungen des IT-Betriebs greifen deshalb unmittelbar ineinander: Wer Infrastruktur-Ressourcen effizienter nutzt, kann Kosten und Umweltbelastungen zugleich senken. Abschalten, bedarfsgerechte Dimensionierung, Autoscaling, passende Speicherklassen, bewusstes Datenlöschen und effiziente Architekturen werden damit zugleich zu Maßnahmen der Kosten- und Nachhaltigkeitssteuerung. Nachhaltigkeit wird so zu einem Handlungsfeld des IT-Controllings.

Ähnliches gilt für digitale Souveränität und organisationale Resilienz. Anbieterabhängigkeiten, Konzentrationsrisiken, fehlende Kompetenzen sowie Migrations- und Ausstiegskosten gehören in Architektur-, Beschaffungs- und Portfoliobewertungen. IT-Controlling übersetzt diese Faktoren in Szenarien und Entscheidungsalternativen. Digitale Souveränität bedeutet dabei nicht Autarkie, sondern bewusst gestaltete Handlungsfähigkeit.

Beim Controlling von KI kehrt die Wertbeitragsfrage zugespitzt zurück. „Tokenmaxxing“ bezeichnet den Fehlanreiz, hohen KI-Verbrauch oder hohe Nutzungsquoten bereits als Erfolg zu werten. Tokens als die kleinsten vom Modell verarbeiteten Texteinheiten schaffen eine präzise Verbrauchs- und Kostenbasis, beantworten aber noch keine Investitionsfrage. Ein günstiges Modell kann durch Fehler, Nacharbeit und Prüfung teuer werden; ein leistungsfähigeres Modell kann sich bei einer kritischen Aufgabe rechnen. Relevant sind Kosten, Qualität, Risiko und Wirkung je akzeptiertes Ergebnis.

Beim Controlling mit KI unterstützt dieselbe Technologie Forecasts, Anomalieerkennung und die entscheidungsorientierte Aufbereitung von Informationen. Sie automatisiert Teile der Analyse, nicht die Verantwortung für Ziele, Kriterien und Abwägungen. KI wird damit zugleich Steuerungsobjekt und Instrument der Steuerung. Für das IT-Controlling verschiebt sich damit der Schwerpunkt stärker auf menschliche Priorisierung, Bewertung und verantwortliche Abwägung.

Das Leitbild der Fachgruppe verstand IT-Controller*innen bereits 2009 als Teil des Managementprozesses sowie als Dienstleistende und Moderierende an Schnittstellen. Vor seiner Aktualisierung steht nun die Zukunftsbild-Diskussion: Welche Entscheidungen soll IT-Controlling künftig verbessern? Wie lassen sich Kosten, Wert, Risiko, Compliance, Qualität, Nachhaltigkeit und Resilienz gemeinsam betrachten?

Nach einer organisatorischen Durststrecke ohne Sprecher*in im Jahr 2025 hat die Fachgruppe ihre Arbeit in den vergangenen Monaten neu aufgenommen. Erste Treffen zu nachhaltiger IT mit externem Fachimpuls markieren den Neustart. Zugleich verändert sich das Arbeitsformat: Regelmäßige Online-Treffen ermöglichen einen offenen Austausch unabhängig von Ort und Reisebudget. ¹

— **Lust mitzumachen?** Interessierte aus Praxis, Wissenschaft, Lehre und Studium sind eingeladen, Praxisfälle, Forschungsergebnisse oder **gerne auch kontroverse Positionen** einzubringen.
Aktuelle Termine: fg-wi-itc.gi.de



Einmal Biber, *immer Biber*

REDAKTION Katja Sauerborn und Susanne Falkenberg-Thut



Viele, die heute in der IT arbeiten, sind mit dem Biber groß geworden: In zahlreichen Interviews hat das BWINF-Team erfahren, wie der Wettbewerb seine Teilnehmenden für das Fach begeistert hat.

Schon 20 Jahre lang knobeln Schüler*innen an den Aufgaben des Informatik-Biber. Viele von ihnen haben durch den Wettbewerb eine Begeisterung für die Informatik entwickelt, die auch über die Schule hinaus geblieben ist.

„Der Biber war ein toller Einstieg in die Wunder der Informatik und hat mich richtig auf den Geschmack gebracht!“, sagt Softwareentwickler Simon Knott. „Was ein Glück, zehn Jahre später habe ich immer noch große Freude daran und könnte nicht glücklicher sein, als Programmierer jeden Tag Rätsel zu lösen – jetzt sogar beruflich.“ Geschichten wie seine

finden sich viele unter den ehemaligen Teilnehmenden des Wettbewerbs. So auch bei Anja Matena, die bis heute noch einen Schlüsselanhänger und eine Powerbank mit Biber-Motiv nutzt: „Was ich am Informatik-Biber immer mochte, war, dass man nicht besser sein musste als alle anderen. Dadurch ging es viel mehr darum, sein eigenes Bestes zu geben, als sich mit den anderen zu vergleichen.“

Wie erfolgreich sich das Konzept in den vergangenen 20 Jahren entwickelt hat, zeigt ein Blick in die Zahlen: Mehr als 5,5 Millionen Schüler*innen haben sich bereits an dem Wettbewerb beteiligt. „Unsere Schülerinnen und Schüler haben auch dieses Jahr wieder gerne mitgemacht – und gar nicht gemerkt, dass das ja knallharte Informatik war“, sagt Sabine Schatz von der Staatlichen Realschule Röthenbach. Das ist zentral beim Biber: Jede Aufgabe hat einen klaren Informatikbezug. So werden grundlegende Fachthemen wie die Reihenfolge von Programmieranweisungen oder Stack und auch anspruchsvolle Inhalte, etwa Algorithmen auf Graphen, NP-Vollständigkeit oder die Turingmaschine, anschaulich vermittelt.

Durch die Gespräche mit Lehrkräften und ehemaligen Teilnehmenden zieht sich diese Beobachtung: Für viele Kinder und Jugendliche sind die Aufgaben des Wettbewerbs ein neuer Zugang zum Thema Informatik, abseits der gängigen Vorurteile, und erzählt anhand lebhafter Geschichten. „Zu mir hat ein Schüler mal gesagt: Das war ja gar nicht wie ein Wettbewerb. Wir haben einfach am iPad geknobelt“, erzählt Janina Petras, die den Biber sogar aus beiden Perspektiven kennt: Sie war einst selbst Teilnehmerin und ermutigt

heute als Informatiklehrerin ihre Schüler*innen, sich den Aufgaben zu stellen. „Diese niedrige Einstiegshürde nimmt Kindern die Angst“, sagt Petras.

Genau das spielt auch eine Rolle bei Kindern und Jugendlichen mit besonderen Bedürfnissen – wie etwa an der Michael-Ende-Schule, an der unter anderem Kinder mit Sprachstörungen oder Autismus unterrichtet werden.



5.613.199 Kinder und Jugendliche

haben seit dem Start des Wettbewerbs am Informatik-Biber teilgenommen. Im Durchschnitt sind die Teilnehmendenzahlen seit 2007 um 23 Prozent pro Jahr gestiegen.



Knapp 200 Aufgaben

hat Deutschland seit 2007 für die internationale Bebras Challenge beigesteuert – mehr als jedes andere Land. BWINF ist mit dem Informatik-Biber Teil dieser weltweiten Initiative, die heute in mehr als 80 Ländern vertreten ist.



„Eine typische Schwierigkeit unserer Schülerschaft ist, dass sie das, was in ihrem Kopf passiert – Geschichten, Emotionen, Gedanken – nur schwer nach außen bringen können, sodass die Umwelt etwas damit anfangen kann“, erklärt Fabian Braun, der als Elternteil an der Schule unterstützt. „Beim Biber gibt es zum Beispiel ein Puzzlefeld mit einer klaren Regel darüber, und wenn ich die Teile richtig verschiebe, zeige ich damit, dass ich die Regel verstanden habe – ohne sie extra erklären zu müssen. Das gibt vielen ein Erfolgserlebnis.“

Das niedragschwellige Konzept zeigt deutschlandweit Wirkung. „Aus einer Initiative ist etwas geworden, das fest im Schulalltag angekommen ist“, so Nils van den Boom, der den Wettbewerb schon seit der ersten Stunde als Lehrer begleitet hat. Heute organisiert das Team der Bundesweiten Informatikwettbewerbe (BWINF), deren Wurzeln bis zum 1980 gegründeten Bundeswettbewerb Informatik zurückreichen, neben dem Biber weitere Wettbewerbe und Angebote, um Kinder und Jugendliche für Informatik zu begeistern. I

Die Biberaufgaben sind bewusst so gestaltet, dass sie für Kinder und Jugendliche einfach zugänglich sind – und auch solche nicht abschrecken, die bisher erst wenig Berührungspunkte mit der Informatik hatten.

– Plüsch-Biber zu gewinnen!

Schon entdeckt? Über die Seiten dieses gesamten Hefts sind mehrere Biber versteckt, in Fotos und Illustrationen, an ganz unterschiedlichen Stellen. Wer alle findet, kann eine Mail mit der richtigen Anzahl an redaktion@gi.de schicken – und hat die Chance, einen Plüsch-Biber zu gewinnen!

– Der Informatik-Biber

ist Deutschlands größter Schülerwettbewerb im Bereich Informatik. Das Online-Quiz mit spannenden Knobelaufgaben fördert das informatische Denken und ist kostenfrei. Er ist einer der Bundesweiten Informatikwettbewerbe (BWINF). bwinf.de



Die häufigsten Tiere

in den Biberaufgaben sind Bienen, Waschbären, Frösche und Schildkröten. Häufigster Stargast ist natürlich der Biber selbst.



17 Autor*innen

aus Deutschland waren insgesamt an der Erstellung der Biberaufgaben beteiligt. Neben Gründungsmitglied und BWINF-Geschäftsführer Wolfgang Pohl sind auch Michael Weigend und Kirsten Schlüter von Anfang an dabei.

„In der Informatik steckt viel mehr, als viele denken“

Greta Schlüter hat eine lange Geschichte mit dem Informatik-Biber. Sie hat nicht nur selbst mitgemacht, sondern auch jahrelang Grafiken für Biberaufgaben erstellt. Heute folgt sie dem Beispiel ihrer Mutter – und ist selbst Teil des deutschen Aufgabenteams.



Greta Schlüter zeichnet seit ihrer Jugend für den Informatik-Biber. Heute gestaltet sie den Wettbewerb auch als Autorin mit.

Sie haben den Biber selbst schon als Schülerin mitgemacht?

Genau, das war ungefähr 2014, ich war damals in der 10. oder 11. Klasse. Den Biber kannte ich aber schon lange vorher – meine Mutter macht seit fast 20 Jahren Biberaufgaben, also fast von Anfang an. Insofern hatte ich schon ein Bild vom Biber, bevor ich selbst mitgemacht habe.

Wie war das, als Kind, mitzubekommen, wie Biberaufgaben entwickelt werden?

Mit 12 oder 13 habe ich angefangen, Grafiken für die Aufgaben zu malen, an denen meine Mutter gearbeitet hat – in Paint, ganz klassisch. Eine ihrer frühen Aufgaben, für die ich Bilder gemalt habe, war ‚Gretas Schatzkarte‘ – ich glaube, sie ist sogar nach mir benannt. Das habe ich viele Jahre lang gemacht, auch noch während meines Studiums.

Wie kam es dazu, dass Sie selbst im Aufgabenteam des Informatik-Biber aktiv wurden?

Meine Mutter wusste natürlich, dass ich Interesse daran habe. Irgendwann meinte sie, das Team könne gut jungen Nachwuchs gebrauchen, und fragte, ob ich nicht Lust hätte, tiefer einzusteigen, selbst Aufgabenvorschläge zu machen und zu den Workshops mitzukommen.

Was macht für Sie eine wirklich gelungene Biberaufgabe aus?

Dass dahinter echte Informatik und ein cooles Konzept stecken, das Ganze aber so reduziert ist, dass man in wenigen Worten erklären kann, worum es geht. Man soll nicht lange brauchen, um zu erfassen, was die Aufgabe verlangt – und dann Spaß daran haben, sie zu lösen.

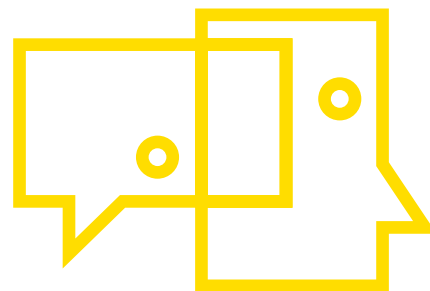
Gibt es eine Aufgabe, die Ihnen besonders am Herzen liegt?

Von meinen eigenen Aufgaben mag ich ‚Sunny Days‘ sehr gern, gerade weil sie so simpel ist. Es ist eine Logikaufgabe und für mich ist Aussagenlogik die Grundlage von ganz vielen Dingen in der Informatik. Es geht um Biber, die schwimmen gehen. Die Grafikerin hat dazu ein sehr schönes Bild gemalt. Das trägt auch dazu bei, dass sich die Aufgabe einfach nett anfühlt.

Welche Botschaft möchten Sie Schüler*innen mit auf den Weg geben?

Ich glaube, viele scheuen vor der Informatik zurück, weil sie ein falsches Bild davon haben – sie denken, ‚ich bin ja nicht computerbegeistert‘, und schieben das Fach deshalb weg. Dabei steckt in der Informatik so viel mehr und man muss sich nicht für die Themen begeistern, die am bekanntesten sind, um etwas für sich zu finden. Ich bin selbst eine wenig computerbegeisterte Person und komme trotzdem sehr gut in der Informatik unter. Wer gerne Rätsel löst, gerade Logik-Rätsel, sollte auf jeden Fall mal einen Blick riskieren, um herauszufinden, ob die Informatik nicht auch etwas für einen ist. ☺

Doppelt gezählt?



FALLBEISPIEL Christine Hennig und Paul Reinelt

Ein kleiner Fehler im Auswertungscode bringt kurz vor der Veröffentlichung die Ergebnisse einer groß angelegten Studie ins Wanken. Doch muss eine Studie gestoppt werden, wenn ihre Kernaussage weiterhin stimmt? Dieses Fallbeispiel zeigt, wie Zeitdruck, wissenschaftliche Sorgfalt und transparente Forschung miteinander in Konflikt geraten können.

Elias hat in Informatik promoviert und eine Forschungsgruppe an einer Hochschule. Er leitet das Forschungsprojekt „ExplAI“, in dem eine umfassende Studie zur KI-Nutzung an Berufsschulen in Deutschland erstellt wird. Daten werden bundesweit mit Fragebögen erhoben, um wissenschaftlich zu bewerten, ob der Einsatz von Chat-Tools bei Hausaufgaben Lernprozesse sinnvoll unterstützt oder ob dadurch Kompetenzen wie eigenständiges Arbeiten und kritisches Denken beeinträchtigt werden. Die Ergebnisse sollen als Grundlage für zukünftige bildungspolitische Entscheidungen dienen und werden von einer umfangreichen PR-Kampagne begleitet.*

Elias' Team arbeitet von Anfang an sehr akribisch und erzeugt im Laufe des Projekts eine Reihe unterschiedlicher digitaler Artefakte: verschiedene Versionen der Umfrage, den Workflow der Datenanalyse und die zugehörigen Skripte und Codes für die Auswertung. Dazu kommen der Text für die Veröffentlichung in einer Zeitschrift und die Statements für die Presse. Das Förderprogramm

verlangt, dass alle diese Artefakte öffentlich zugänglich und nachnutzbar sind. Dies dient der Überprüfbarkeit sowohl der Ergebnisse als auch der Herangehensweise.

Elias und seine Kollegin Raya diskutieren, wie sie Daten und Auswertungscode geeignet ablegen. Raya hat kürzlich einen Vortrag über Forschungsdatenmanagement gehört und schlägt daher vor, alle digitalen Artefakte im institutionellen Repositorium zu speichern und mit geeigneten Metadaten zu versehen. Daten und Code sollen gemeinsam mit der wissenschaftlichen Publikation veröffentlicht werden, sodass andere Forschende die Ergebnisse nachvollziehen und reproduzieren können. Elias ist skeptisch bezüglich des Codes, insbesondere da einige der verwendeten Skripte aus eher dubiosen Quellen stammen, nur unsauber dokumentiert sind und leider auch nicht alle Workflows eindeutig reproduzierbar sind. Raya kann ihn jedoch überzeugen und verspricht im Gegenzug, die Codequalität zu verbessern und die Dokumentation zu vervollständigen.

Raya investiert viel mehr Zeit als geplant in die Codequalität. Glücklicherweise unterstützt Marie sie dabei, die studentische Hilfskraft im Projektteam. Die beiden stellen fest, dass der extrem hohe Rücklauf von 85 Prozent der Berufsschulen den Code an sein Limit bringt, er skaliert nicht hinreichend. Sie arbeiten mit Hochdruck an der Korrektur. Den teilnehmenden Berufsschulen und der Politik wurde versprochen, die Ergebnisse innerhalb von drei Monaten zu publizieren und ihnen zur Verfügung zu stellen. Die Zeit rennt ihnen davon, denn die Frist läuft in wenigen Tagen ab.

Am Vorabend der geplanten Veröffentlichung sitzt Marie im Büro und geht den Auswertungsprozess noch ein letztes Mal sorgfältig durch. Ihr Bauchgefühl sagt ihr: Hier ist etwas falsch. Beim Korrekturlesen kommen ihr die Zahlen seltsam vor. Dann findet sie die Ursache ihres Bauchgrummelns: Die Anzahl an verarbeiteten Rückmeldungen erreicht nicht ganz die erwartete Zahl.

Sie schaut erneut auf die Daten und schickt gemeinsam mit Raya noch einmal alles durch das nunmehr

bestens dokumentierte Auswertungstool. Sie stellen den Ausgabemodus so ein, dass es für jede einzelne Verarbeitungsaktion eine Meldung ausgibt. Dabei fallen ihnen seltsame Meldungen auf: „name already exists“. Warum hat das bisher niemand bemerkt? Das automatische Auswertungstool hat bei Lehrkräften mit denselben Vor- und Nachnamen jeweils einen der Fragebögen ignoriert. Es gibt unter den befragten Personen zwei „Lisa Bach“ und wer weiß wie viele weitere Namensgleichheiten. Raya stöhnt auf. Sie müsste die Logik ihres Auswertetools komplett umkrempeln. Dabei hatte sie den Code vor der Veröffentlichung von Elias begutachten lassen. Die ganze Studie müsste neu bewertet werden.

Raya telefoniert mit Elias. Aber der winkt ab. Es werden schon nicht alle Lehrkräfte „Lisa Bach“ heißen, das Fehlen einzelner Datensätze falle nicht weiter ins Gewicht und ändere nichts an der Kernaussage der Studie. 65,26 Prozent der Befragten befürworten die Nutzung von KI-Tools für Hausaufgaben. Ob sich nun weit hinter dem Komma eine Ziffer ändere, sei unerheblich. Er weist Raya an, die gefundenen Fehler zu ignorieren und die Studie wie geplant zu veröffentlichen.

Raya und Marie lassen nicht locker. Die nicht ganz exakten Ergebnisse zeugen nicht nur von unsauberer Forschung – der Fehler lässt sich durch den zur Verfügung gestellten Code auch nachweisen. Eine Zählung der Anzahl der Meldungen „name already exists“ ergibt 32 Einträge bei über 100.000 Verarbeitungsmeldungen. Von den erhobenen Datensätzen betrifft das Problem vermutlich nur diese 32, auch wenn Raya und Marie sich nicht ganz sicher sind. Die gefundenen Fehler ändern allerdings nicht die Kernaussagen und Elias' Argumentation ist somit nicht von der Hand zu weisen. ¶

*Anmerkung: Die Untersuchung ist von uns frei erfunden – eventuell dient die Idee ja für interessierte Forschende als Inspiration? Die Story basiert auf der Story 32 der Research Data Scarytales.

FRAGEN

Darf Raya die Veröffentlichung der Ergebnisse entgegen Elias' Anweisung zurückhalten?

Was bedeutet es für die Reputation von Elias und der gesamten Arbeitsgruppe, wenn jemand die Fehler findet?

Welche Verantwortung tragen Forschende, wenn ihre Ergebnisse als Grundlage für politische Entscheidungen dienen?

Sollten bekannte Unsicherheiten in den Forschungsdaten offen kommuniziert werden, auch wenn dadurch Zweifel an den Ergebnissen entstehen könnten?

Ist eine Veröffentlichung trotz bekannter Fehler vertretbar, wenn die Ergebnisse im Wesentlichen unverändert bleiben? Ab wann wiegt ein Fehler zu schwer?

Welche Verantwortung tragen Forschende wie Elias und Raya für die Dokumentation, Qualität und Zugänglichkeit der digitalen Artefakte? Welchen Anteil hat Marie daran?

Welche Rolle spielt die Veröffentlichung von Daten und Code für die Nachvollziehbarkeit wissenschaftlicher Ergebnisse?

Wie können externe Forschende, die die Ergebnisse zu reproduzieren versuchen, auf Lücken oder Fehler bei der Veröffentlichung digitaler Artefakte aufmerksam machen und wie sollte damit umgegangen werden?

Wie kann eine Fehlerkultur etabliert werden, die kontinuierliche Verbesserung bei der Speicherung und Dokumentation von Forschungsdaten ermöglicht?

Wie würden Sie diese Fragen beantworten? Jetzt mitdiskutieren: [↪ gewissensbits.gi.de](https://www.gewissensbits.gi.de)

— **Die Gewissensbits** sind hypothetische, jedoch realistische Fallbeispiele. Mit ihnen will die **Fachgruppe Informatik und Ethik** einladen, über moralische und ethische Fragen rund um die Informatik nachzudenken und diese mit Freund*innen, Kolleg*innen oder Studierenden zu diskutieren. Meist gibt es auf diese Fragen keine einfachen Antworten – dazu sind die Fallbeispiele, wie auch die Realität, zu komplex. Im gemeinsamen Diskurs lässt sich aber üben, mögliche Probleme zu erkennen und als Gruppe zu einer Lösung zu gelangen: in einem demokratischen Aushandlungsprozess, wie er auch im Alltag erforderlich ist.

Funktioniert. *Aber für wen?*

TEXT Lilia Gargouri

Unsere Autorin findet: Digitale Barrierefreiheit ist eine Shift-Left-Disziplin. Was das bedeutet, welche rechtlichen Vorgaben zu beachten sind und warum barrierefreie Software sowieso die bessere Software ist, fasst sie in diesem Artikel zusammen.

Mit dem Inkrafttreten des Barrierefreiheitsstärkungsgesetzes (BFSG) und der Umsetzung des European Accessibility Act (EAA) hat die digitale Barrierefreiheit in Deutschland erheblich an Bedeutung gewonnen. In vielen Unternehmen gilt sie dennoch vor allem als regulatorische Pflicht. Diese Sicht greift zu kurz: Digitale Barrierefreiheit ist weit mehr als die Erfüllung gesetzlicher Vorgaben – sie ist ein integraler Bestandteil moderner Softwarequalität.

Barrierefreiheit bedeutet, dass digitale Produkte von möglichst vielen Menschen wahrgenommen, verstanden, bedient und effektiv genutzt werden können – unabhängig von individuellen Fähigkeiten, Einschränkungen oder eingesetzten assistiven Technologien. Davon profitieren nicht nur Menschen mit Behinderungen, sondern ebenso ältere Menschen sowie Personen mit situativen oder temporären Einschränkungen, beispielsweise aufgrund einer Verletzung. Darüber hinaus verbessert Barrierefreiheit häufig auch die Nutzung auf mobilen Endgeräten oder unter erschwerten Umgebungsbedingungen, etwa bei starker Sonneneinstrahlung oder in lauten Umgebungen.

Die gesellschaftliche Relevanz ist offensichtlich. Nach Angaben des Statistischen Bundesamtes lebten Ende 2023 rund 7,9 Millionen schwerbehinderte Menschen in Deutschland; weltweit schätzt die WHO den Anteil von Menschen mit Behinderungen auf etwa 16 Prozent. Zugleich zeigt der WebAIM Million Report 2026, dass 95,9 Prozent der untersuchten Startseiten im Internet mindestens einen automatisch erkennbaren Verstoß gegen die Web Content Accessibility Guidelines (WCAG)

Eine weiblich gelesene Person sitzt an einem Schreibtisch in einem Büro. Sie nutzt einen Laptop, der zusätzlich mit einer sogenannten Braillezeile ausgestattet ist, die die Inhalte auf dem Screen des Laptops in Braille-Schrift übersetzt. Auf dem Laptop ist die Website des World Wide Web Consortiums zu sehen.

aufweisen. Seit Jahren dominieren dieselben grundlegenden Mängel: unzureichende Farbkontraste, fehlende Alternativtexte, unbeschriftete Formularelemente, leere Schaltflächen. Das Problem liegt also weniger in fehlenden technischen Möglichkeiten als in der mangelnden Berücksichtigung im Entwicklungs- und Testprozess.

Barrierefreiheit als Qualitätsfaktor

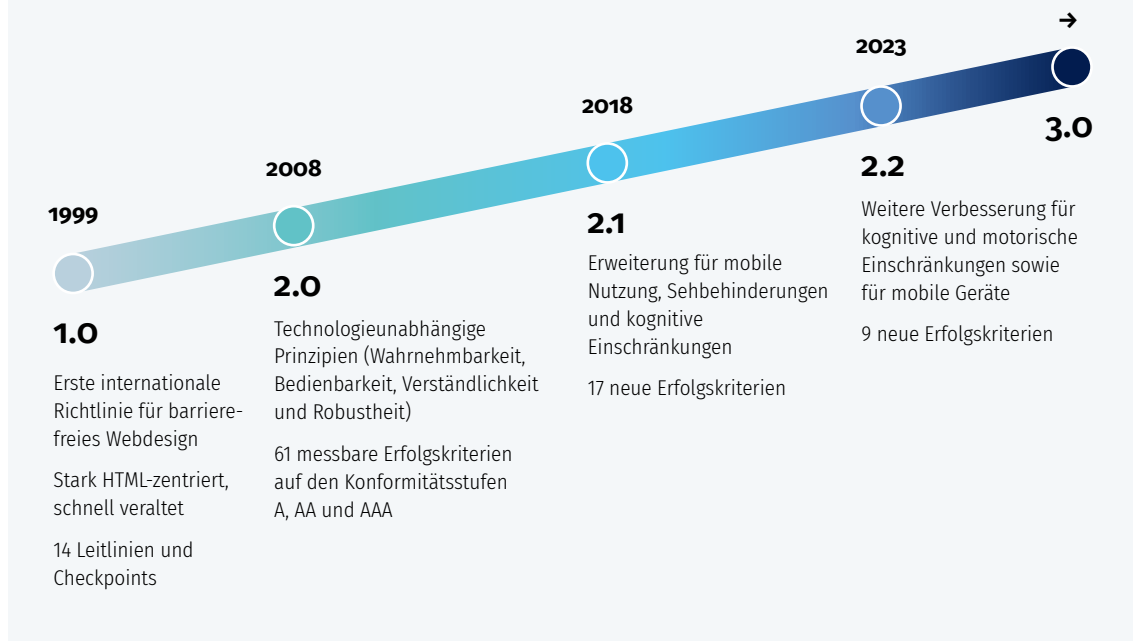
Die Qualität eines Softwareprodukts lässt sich nicht auf einzelne Eigenschaften wie Funktionalität oder Fehlerfreiheit reduzieren. Die ISO/IEC 25010:2023 definiert ein international anerkanntes Qualitätsmodell, das die wesentlichen Qualitätsmerkmale eines Softwareprodukts beschreibt. Dazu gehören unter anderem funktionale Eignung, Zuverlässigkeit und Interaktionsfähigkeit. Gemeinsam bilden diese Merkmale die Grundlage, um Softwarequalität systematisch zu bewerten und gezielt weiterzuentwickeln.

Digitale Barrierefreiheit trägt unmittelbar zur Erfüllung mehrerer Qualitätsmerkmale der ISO/IEC 25010:2023 bei. Ihr stärkster Bezug besteht zur Interaktionsfähigkeit, insbesondere zum Teilmerkmal Inklusivität.

Inklusivität beschreibt die Fähigkeit eines Systems, Menschen mit unterschiedlichen Fähigkeiten und Nutzungskontexten eine gleichberechtigte Interaktion zu ermöglichen: So sollte beispielsweise eine App von blinden Menschen ebenso gut genutzt werden können wie von sehenden. Kann eine Anwendung hingegen nicht vollständig über die Tastatur bedient werden oder sind ihre Inhalte für Screenreader nicht interpretierbar, weist sie demnach klare Defizite hinsichtlich ihrer Interaktionsfähigkeit auf.

Darüber hinaus wirkt sich digitale Barrierefreiheit auf weitere Qualitätsmerkmale aus. Semantisch korrekt implementierte Benutzeroberflächen verbessern die Kompatibilität, insbesondere die Interoperabilität mit Browsern und assistiven Technologien. Klare Strukturen und standardkonforme Implementierungen können zudem die langfristige Wartbarkeit unterstützen. Digitale Barrierefreiheit trägt damit wesentlich zur Gesamtqualität eines Softwareprodukts bei. Sie sollte deshalb von Beginn an als nicht funktionale Qualitätsanforderung behandelt werden – ebenso selbstverständlich wie Sicherheit, Leistungseffizienz oder Zuverlässigkeit.

WCAG (Web Content Accessibility Guidelines)



EU legt vor, nationales Recht setzt um

Die rechtlichen Anforderungen an digitale Barrierefreiheit bestehen aus mehreren Ebenen, die unterschiedliche Aufgaben erfüllen. EU-Richtlinien definieren rechtliche Ziele und Mindestanforderungen. Sie beschreiben, was erreicht werden soll, enthalten jedoch keine technischen Vorgaben für deren Umsetzung. Erst durch die Überführung in nationales Recht entstehen verbindliche Anforderungen für Behörden und Unternehmen.

Europäische Normen sorgen für die technische Konkretisierung: Sie beschreiben, wie die gesetzlichen Ziele umgesetzt werden können. Für digitale Produkte übernimmt diese Aufgabe die europäische Norm EN 301 549. Sie definiert technische Anforderungen an Informations- und Kommunikationstechnologien und verweist für Webanwendungen auf die Web Content Accessibility Guidelines (WCAG) des World Wide Web Consortiums (W3C).

Die WCAG selbst sind aber kein Gesetz. Sie bilden den international anerkannten technischen Standard für digitale Barrierefreiheit. Ihre rechtliche Relevanz entsteht dadurch, dass europäische Normen und nationale Gesetze auf sie Bezug nehmen. Tatsächlich orientieren

sich Entwicklung, Prüfung und Konformitätsbewertung von digitaler Barrierefreiheit heute weltweit an den WCAG.

Wahrnehmbar, bedienbar, verständlich und robust

Die Richtlinien basieren auf vier grundlegenden Prinzipien: Inhalte müssen wahrnehmbar, bedienbar, verständlich und robust sein. Daraus leiten sich überprüfbare Erfolgskriterien auf den Konformitätsstufen A, AA und AAA ab, die den technischen Maßstab für barrierefreie Webanwendungen bilden. Für die meisten gesetzlichen Anforderungen bildet heute die Konformitätsstufe AA der WCAG (insbesondere WCAG 2.1 AA im Kontext der EN 301 549) den maßgeblichen Referenzrahmen.

Was Entwickler*innen wissen müssen

Digitale Barrierefreiheit ist eine nicht funktionale Anforderung. Sie muss über den gesamten Softwarelebenszyklus hinweg berücksichtigt werden. Wenn sich Teams erst kurz vor der Auslieferung mit ihren Aspekten auseinandersetzen, kosten die notwendigen Anpassungen oft viel Arbeit, Zeit und Nerven.

Digitale Barrierefreiheit ist eine klassische Shift-Left-Disziplin. Das Shift-Left-Prinzip verfolgt das Ziel, Qualitätsanforderungen möglichst früh im Softwareentwicklungsprozess zu berücksichtigen, anstatt Defizite erst kurz vor der Auslieferung zu erkennen und mit hohem Aufwand zu beheben. Anforderungen an Tastaturbedienbarkeit, Fokusführung, Alternativtexte oder verständliche Fehlermeldungen sollten bereits im Requirements Engineering und UX-Design berücksichtigt werden.

Automatisierte Prüfwerkzeuge wie axe-core, WAVE oder Lighthouse unterstützen Entwicklungs- und Testteams dabei, technische Verstöße frühzeitig zu erkennen und kontinuierlich zu überwachen. Sie identifizieren beispielsweise fehlende Alternativtexte, unzureichende Farbkontraste oder fehlerhafte HTML-Strukturen zuverlässig und reproduzierbar.

Doch die Möglichkeiten dieser Prüfwerkzeuge sind begrenzt. Automatisierte Prüfungen dienen in erster Linie der Verifikation technischer Anforderungen. Ob eine Anwendung für ihre vorgesehenen Nutzergruppen tatsächlich barrierefrei nutzbar ist, lässt sich hingegen erst beurteilen, wenn sie unter realistischen Nutzungsbedingungen getestet wird. Ob ein Alternativtext den Inhalt eines Bildes sinnvoll beschreibt, eine Fehlermeldung verständlich formuliert ist oder sich eine komplexe Anwendung mit einem Screenreader effizient bedienen lässt, erfordert menschliche Bewertung.

Auf der anderen Seite gewinnen KI-Verfahren bei der Prüfung digitaler Barrierefreiheit zunehmend an Bedeutung. Sie unterstützen beispielsweise bei der Analyse von Benutzeroberflächen, der Bewertung von Alternativtexten oder der Erkennung komplexer Nutzungshindernisse. Aber auch hier gilt: KI kann menschliche Bewertung nicht ersetzen. Ob Inhalte verständlich sind, Interaktionen intuitiv funktionieren oder eine

Anwendung für Menschen mit unterschiedlichen Einschränkungen tatsächlich nutzbar ist, erfordert weiterhin fachliche Expertise und die Einbeziehung realer Nutzer*innen.

Wer Barrierefreiheit professionell prüfen will, kombiniert daher automatisierte Prüfungen mit manuellen Reviews, Tastatortests, Tests mit assistiven Technologien und – sofern möglich – Nutzertests mit Menschen mit Behinderungen. Wie im klassischen Softwaretesten gilt auch hier: Kein einzelnes Testverfahren kann alle Fehler aufdecken. Erst das Zusammenspiel unterschiedlicher Testmethoden ermöglicht eine belastbare Aussage über die Qualität eines Systems.

Fazit: Mitdenken statt nachrüsten

Digitale Barrierefreiheit ist weit mehr als die Erfüllung gesetzlicher Anforderungen. Sie ist ein wesentlicher Bestandteil moderner Softwarequalität und trägt zur Erfüllung zentraler Qualitätsmerkmale bei. Organisationen, die Barrierefreiheit frühzeitig in Anforderungen, Architektur, Entwicklung und Testprozesse integrieren, reduzieren nicht nur regulatorische Risiken und Nacharbeitsaufwände. Sie entwickeln zugleich robustere, verständlichere und besser nutzbare Software.

Die eigentliche Herausforderung besteht nicht darin, Barrierefreiheit am Ende eines Projekts nachzuweisen. Sie besteht darin, sie als selbstverständlichen Bestandteil professionellen Software Engineerings zu etablieren. Denn ein Softwareprodukt erreicht seinen Qualitätsanspruch erst dann vollständig, wenn es seine funktionalen und nicht funktionalen Anforderungen erfüllt und von den vorgesehenen Nutzergruppen tatsächlich wahrgenommen, verstanden, bedient und effektiv genutzt werden kann. I

– Über die Autorin

Lilia Gargouri ist Senior Software Engineer und Head of QA bei mgm technology partners. Sie engagiert sich für digitale Barrierefreiheit, qualifiziert Quereinsteiger*innen in der Testautomatisierung und setzt sich für mehr Sichtbarkeit und Chancengleichheit von Frauen in der IT ein. Darüber hinaus gestaltet sie als ISTQB® Glossary Chair und Mitglied des German Testing Board die Weiterentwicklung internationaler Qualitätsstandards im Softwaretesten aktiv mit.



Neues aus dem Noerden

Der Noerdman, ein Webcomic über Technik, Nerds und den Norden, ist fester Bestandteil der .inf-Redaktion.

Eine große Auswahl der Comics ist übrigens auch als Bildband verfügbar. noerdman.de

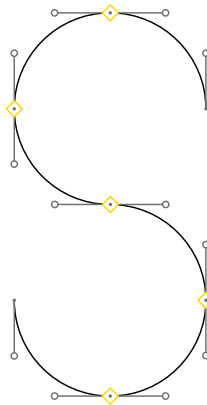
Wie S begann

TEXT Sven Lubenau

*Donald E. Knuth hat mit TeX ein bedeutendes System für den digitalen Satz und mit METAFONT einen mathematischen Ansatz zur Beschreibung von Schriftformen entwickelt. Das fasziniert nicht nur Informatiker*innen, sondern auch den Creative Director der GI.*

„Als ich begann, Schrift in moderne digitale Umgebungen zu übertragen, stellte ich fest, dass sich die meisten der 25 Buchstaben des Alphabets vergleichsweise einfach konstruieren lassen. Beim Buchstaben S ist das anders.“ So erklärte Donald E. Knuth, wie ihn ausgerechnet dieser eine Buchstabe zu einer grundsätzlichen typografischen Frage führte.

Denn was zunächst wie eine einfache Form erscheint, entzieht sich bei genauerem Hinsehen einer sauberen Konstruktion. Frühe geometrische Methoden des Buchstabenentwurfs versuchten, das S mit Kreisbögen und Geraden zu bauen – als Annäherung an die Form, aber mit sichtbaren Brüchen

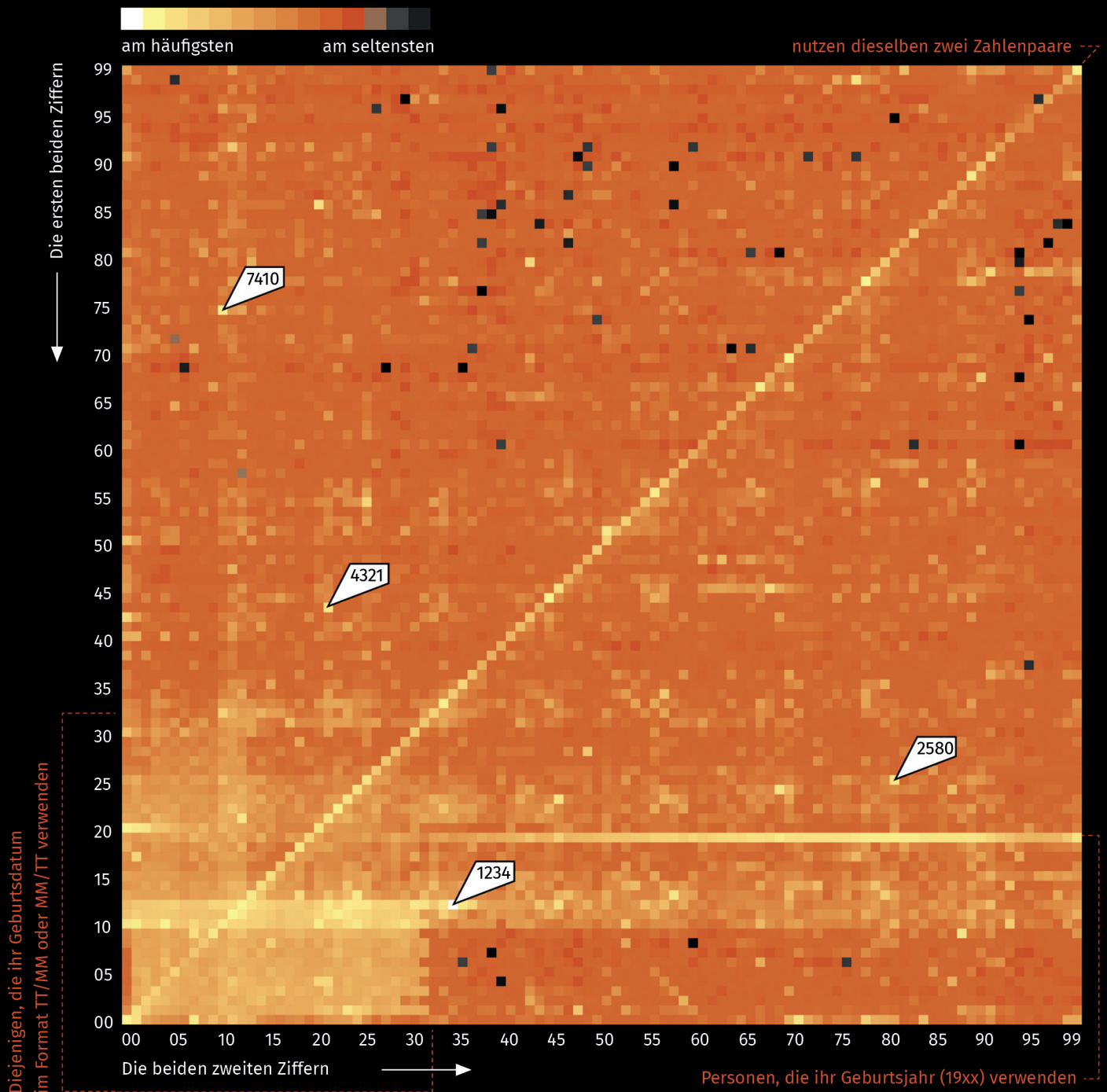


und nur bedingt überzeugender Harmonie. Knuth setzte dagegen auf die Konstruktion mit Ellipsen und mathematisch formulierten Bedingungen, um die Form nicht nur nachzubilden, sondern kontrolliert und konsistent zu definieren.

Gerade darin liegt der Unterschied zu vektorbasierten Schriftmodellen wie PostScript, in denen die Form als Pfad oder Kontur beschrieben wird. Knuth dachte anders: Nicht nur die sichtbare Umrisslinie sollte festgelegt

werden, sondern die innere Logik der Form – also die Beziehungen zwischen Krümmung, Spannung und Überhängen. Er entwickelte METAFONT als Sprache für Schriftformen und TeX für den Satz von ganzen Texten. Beide Systeme folgen derselben Idee: Gestaltung nicht als starre Zeichnung zu begreifen, sondern als präzise beschreibbares Verfahren. Die Schrift wird damit nicht einfach digitalisiert, sondern methodisch neu gedacht.

Knuths Arbeit prägte das digitale akademische Publizieren immens. Auf seinem Satzsystem TeX baute der Informatiker Leslie Lamport später LaTeX auf, das bis heute die wissenschaftliche Textproduktion vereinfacht. Damit hatte Donald E. Knuth einen Einfluss, der sich durchaus mit dem des Bauhaus auf das Grafikdesign vergleichen lässt. Nur logisch, dass er selbst eines der neuen Poster prägt, mit denen die GI einflussreiche Persönlichkeiten aus der Informatik vorstellt. $\int$



Am häufigsten verbreitete Kombinationen:

1234 0000 7777 2000 2222 9999 5555 1122 8888 2001
 1111 1212 1004 4444 6969 3333 6666 1313 4321 1010

27 % aller PIN-Codes

Wie leicht errät man einen PIN-Code? Ob Kreditkarte oder Smartphone: Vier Ziffern sind oft genug, um an unsere sensibelsten Daten zu kommen. Im Buch „Knowledge is beautiful“ zeigt eine Visualisierung von 3,4 Millionen Datenpunkten, ursprünglich erstellt von Nick Berry (DataGenetics) aus unterschiedlichen Leaks, welche Kombinationen besonders beliebt sind. Und? Den eigenen auch entdeckt? [↗](#)

Alles im Fluss

TEXT Nuri Grenz

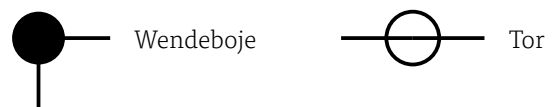
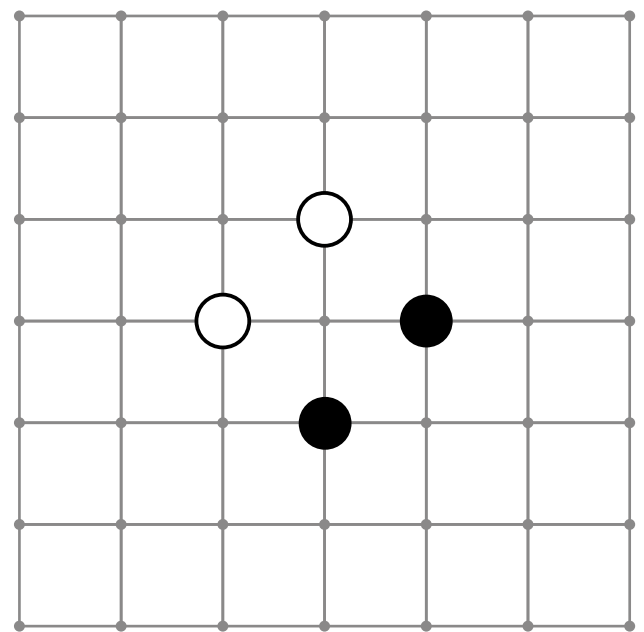
*Ein Kanuturnier ohne Rennstrecke – das geht nicht!
Die Suche nach einem geeigneten Kurs ist inspiriert
vom japanischen Logikrätsel Masyu.*

Bei einem Kanuturnier auf der Elbe ist ausgerechnet die Datei mit der Streckenplanung abhandengekommen. Jetzt gibt es zwar Durchfahrtstore und Wendebojen – aber keinen klaren Weg, der alle Hindernisse richtig passiert.

Die Richtlinien zur Streckenführung sind wie folgt:

- Es muss ein Rundkurs sein, der sich nicht kreuzen darf.
- Alle Tore (weiß) und Wendebojen (schwarz) müssen durchfahren oder passiert werden.
- An Wendebojen muss die Strecke rechtwinklig abknicken, darf aber nicht unmittelbar davor oder danach abbiegen.
- Während der Tordurchfahrt darf nicht abgelenkt werden. Damit die Tore nicht zu schnell durchfahren werden, muss in mindestens einem der beiden Felder vor oder hinter dem Tor rechtwinklig abgelenkt werden. 

Findet ihr den richtigen Weg an allen Hindernissen vorbei?



Steht das Wasser bis zum Hals?

Mit unserem Online-Tool könnt ihr unterschiedliche Varianten ausprobieren: inf.gi.de/allesimfluss



GESELLSCHAFT
FÜR INFORMATIK

MITGLIEDERMAGAZIN DER
GESELLSCHAFT FÜR INFORMATIK E. V.
NR. 15

HERAUSGABE

Gesellschaft für Informatik e.V. (GI)
Weydingerstraße 14–16, 10178 Berlin
Telefon +49 30 240 09-805

VERANTWORTLICH IM SINNE DES PRESSERECHTS

Cornelia Winter

REDAKTION

Alexandra Resch (Ltg.), Cornelia Winter,
Michael Koch, Paul Reinelt, Nuri Grenz,
Julia Meisner-Dieffendahl

AUTOR*INNEN

Martin Wolf, Nuri Grenz, Paul Reinelt,
Aleksandra Sowa, Merrin Chaletu,
Andreas Paech, Katja Sauerborn, Susanne
Falkenberg-Thut, Christine Hennig, Lilia
Gargouri, Sven Lubenau, Alexandra Resch

FACHBEIRAT

Tanja Döring, Oliver Günther,
Dominik Herrmann, Christof Leng,
Friedmann Mattern, Judith Michael,
Andreas Oberweis, Edy Portmann,
Ralf Reussner, Stefanie Rinderle-Ma,
Ute Schmid, Dirk Taubner,
Wolfram Wingerath

GESTALTUNG, SATZ UND GRAFIK

Sven Lubenau (Ltg.), Lucas Crisólogo Oña,
Florian Heinitz

LEKTORAT

Dr. phil. Birgit Gottschalk

DRUCK

Klimaneutral gedruckt
auf Recyclingpapier bei
Speedruck GmbH, Berlin

STAND

September 2026

BILDNACHWEISE

Cover: Honey-Yanibel-Minaya / Unsplash,
S. 3: Mike Auerbach, S. 4, 8, 10: AH.Design, S. 5: privat,
S. 6: YutongLiu-KingstonSchool of Art Talkingto AI 2.0/
betterimagesofai.org, S. 12: Lucas Crisólogo Oña,
S. 16, 18: Florian Heinitz / Wan 3.0, Dorothea Wagner-
berger, S. 22–25: BWINE, S. 3: Jannis Stoppe und
Rolf Drechsler, Sven Lubenau, S. 34: datagenetics.com/
informationisbeautiful

ISSN (Print) 2940-0694

ISSN (Online) 2940-0708

KONTAKT

Wir freuen uns immer über Feedback,
Anregungen und Ideen: Unter redaktion@gi.de
können Sie uns diese zukommen lassen und
auf Wunsch auch weitere Hefte bestellen.

 /mas.to/@informatik

 @informatik.eurosky.social

 /company/gesellschaft-fuer-informatik

Sie lesen lieber digital?

Alle Ausgaben finden Sie auch frei
zugänglich auf inf.gi.de.



**Ihnen gefällt, was
Sie in der Hand haben?**
Bestellen Sie die
Printversion der .inf
im Mitgliederbereich!

**Informatik
Festival
2027**

**21.-
24.09.
2027**



Save the date!

Extremistische Akteure stellen die Freiheit von Forschung und Lehre infrage, KI-Systeme generieren eine Flut an Publikationen. Gleichzeitig verlagert sich immer mehr Forschung in privatwirtschaftliche Strukturen, in denen Verwertungsinteressen die Agenda bestimmen.

Das INFORMATIK FESTIVAL 2027 an der Hochschule für Technik und Wirtschaft Berlin dreht sich um die Frage, wie wir wissenschaftliche Freiheit, Integrität und Offenheit im digitalen Zeitalter bewahren können – und welche Rolle die Informatik dabei spielen kann.

informatik2027.de