Hochschulforum Digitalisierung

Arbeitspapier Nr. 94 / September 2026

Wachsamkeit & Verifizierung

Digitale Resilienz in der Hochschule als Faktor für gelebte Demokratie

Simon D. Vonlanthen

Christoph Koitka-Fieke

Arbeitspapier Nr. 94 / September 2026

Wachsamkeit & Verifizierung

Digitale Resilienz in der Hochschule als Faktor für gelebte Demokratie

Autor:innen

Simon D. Vonlanthen, Hochschulforum Digitalisierung

Christoph Koitka-Fieke, Hochschulforum Digitalisierung

1. Inhalt

Das Hochschulforum Digitalisierung

Das Hochschulforum Digitalisierung (HFD) unterstützt Hochschulen deutschlandweit dabei, die digitale Transformation gemeinsam zu gestalten. Es stärkt die Kompetenzen von Hochschulangehörigen, stößt Entwicklungsprozesse und Veränderungen an und trägt dazu bei, dass konkrete Lösungen wirksam im Alltag von Studium und Lehre ankommen.

Das Hochschulforum Digitalisierung wurde 2014 gegründet. Das transferorientierte Kompetenzzentrum ist eine gemeinsame Initiative von Stifterverband, CHE Centrum für Hochschulentwicklung und Hochschulrektorenkonferenz. Gefördert wird es vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

STIFTERVERBAND

CHE Centrum für Hochschulentwicklung

HRK Hochschulrektorenkonferenz

GEFÖRDERT VOM Bundesministerium für Bildung und Forschung

Die Autoren

Simon D. Vonlanthen ist Referent für politische Analyse im Hochschulforum Digitalisierung. In seiner Arbeit setzt er sich nebst dem Fokus auf (Hochschul-)Politik unter anderem mit der Resilienz des Hochschulsystems auseinander.

Christoph Koitka-Fieke ist Referent für IT & Technologie im Hochschulforum Digitalisierung. Der kritische Umgang mit Desinformation hat ihn bereits in vorherigen Tätigkeiten in Medien und Erwachsenenbildung beschäftigt.

1 Einführung

Die maßgebliche Währung von Wissenschaft und Hochschulen ist das Vertrauen: in saubere Arbeit, in belastbare Ergebnisse und in gut informierte Diskursräume. Dieses Vertrauen ist keine Selbstverständlichkeit, sondern wird durch kontinuierliche Bemühungen immer wieder aufgebaut und erhalten. Genauso gibt es seit jeher Angriffe auf die Integrität von Hochschule und Wissenschaft. Die Digitalisierung hat diese, wie viele andere Felder auch, dynamisiert und vorangetrieben. Es ist heute leichter, Quellen zu fälschen oder Belege zu fabrizieren, als es das noch vor wenigen Jahren der Fall war. Insbesondere der Einsatz von Künstlicher Intelligenz (KI) hat hier die Entwicklung massiv vorangetrieben.

Die Hochschulen sind diesen Angriffen aber nicht schutzlos ausgeliefert. Die Antworten halten in den meisten Fällen technisch mit. Es geht aber nicht darum, in ein digitales Wettrüsten zwischen falschen und wahren Informationen einzutreten – die Verifikation ist nur eine Säule unseres Ansatzes.

In dieser Handreichung stellen wir die häufigsten Probleme vor, die im Zusammenhang mit sogenannten *Fehl-* und *Desinformationen* das Hochschulsystem belasten. Wir unterscheiden dabei zwischen Fehl- bzw. *Falschinformation*, bei der der Inhalt falsch ist, *Malinformation*, die - wahr oder falsch – mit Schädigungsabsicht verbreitet wird und Desinformation, die sowohl inhaltlich falsch ist als auch absichtlich deswegen verbreitet wird (s. Abb. 1).[1] Nebst der Darstellung gängiger Probleme mit Fehl- und Desinformation möchten wir vor allem zeigen, wie man ihnen auf individueller Ebene effektiv begegnet. Hierzu möchten wir uns an eine besonders vulnerable Gruppe im akademischen Bereich richten: Studierende und Promovierende.

Die Betonung der individuellen Ebene ist vor allem deshalb wichtig, weil der Umgang mit Fehl- und Desinformation (kurz: „irreführende Informationen“) eine wichtige Grundlage zur mündigen Teilnahme des Individuums an einer Demokratie ist.[2] Fake News, also bewusst manipulierte Informationen mit der Anmutung journalistischer Expertise, werden oft vor allem im medial-politischen Umfeld verortet. Die Legitimationskrise betrifft aber nicht nur die Medien, sondern ist eher eine Krise der Wissensvermittlung allgemein. Damit stehen auch Hochschulen im Fokus. Durch ihre Funktion, methodisch kontrollierte und kritisch geprüfte Informations- und Wissensangebote zu machen, berührt das Themenfeld den Kern ihres gesellschaftlichen Auftrages unmittelbar.[3]

[1] S. Weatherall (2024), S. 2f.
[2] Vgl. bspw. https://digital-strategy.ec.europa.eu/en/policies/media-literacy.
[3] Vgl. Antos & Ballod (2019), S. 31; s. auch Schroll et al. (2019), S. 4.

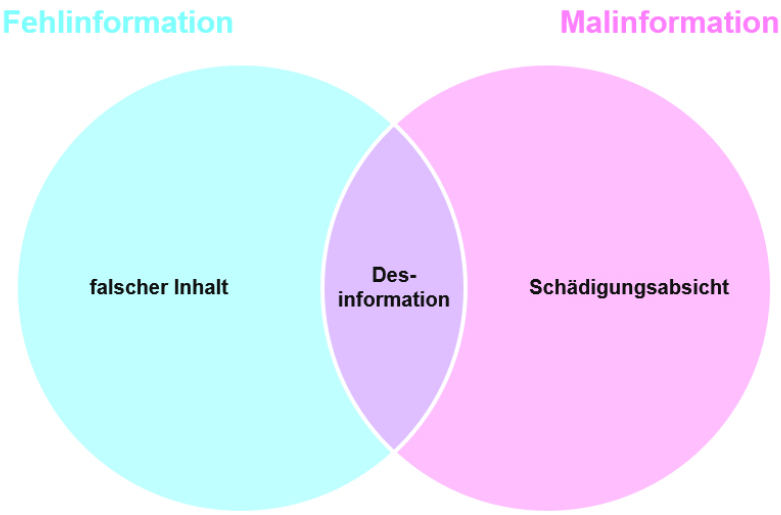


Abbildung 1: Die Begriffe Mal-, Fehl- und Desinformation und deren logische Beziehungen

Wegen unseres Fokus auf Studierende und Promovierende haben wir diese Publikation möglichst praxisnah ausgerichtet und nach dem Grundsatz „Wachsamkeit und Verifikation" gestaltet. Wachsamkeit und Verifikation sind zwei der wirksamsten Mittel gegen irreführende Informationen. Wachsamkeit bedeutet, Inhalte nicht unhinterfragt zu übernehmen, sondern kurz innezuhalten. Wer ist die Quelle? Wirkt die Aufmachung seriös, oder soll sie es nur? Diese Sekunden der Prüfung reichen oft aus, um offensichtliche Fälschungen zu erkennen. Verifikation geht einen Schritt weiter. Sie prüft aktiv nach: Existiert die zitierte Studie tatsächlich? Sagt die genannte Person das wirklich? Lässt sich das Bild einer Rückwärtssuche unterziehen? Beide Hebel wirken nicht isoliert, sondern zusammen. Wachsamkeit schärft den Blick dafür, wann eine Prüfung nötig ist. Verifikation liefert die Werkzeuge, um sie durchzuführen. Wer beides einübt, schützt sich besser gegen irreführende Informationen, unabhängig davon, ob sie aus einem gefälschten Zitat, einem manipulierten Bild, einer halluzinierten Quelle oder einem Raubverlag stammen. Praktisch heißt das:

- Kritische Auseinandersetzung:
  - Wachsam und sensibilisiert sein
  - Aktiv prüfen und kritisch hinterfragen
- Medienkompetenz
  - Typische Manipulationen kennen
  - Gefährdungsgrad einschätzen
- Werkzeugkoffer
  - Fälschungen sowie Manipulationen prüfen und erkennen

- Methoden als Multiplikator:in verbreiten

Mit diesem Dreischritt ist es möglich, in der Praxis mit überschaubarem Aufwand ein System zu etablieren, das kritisches Einordnen ermöglicht. Unsere Idee mit dieser Handreichung ist, im Umgang mit irreführenden Informationen verbreitete Unsicherheiten zu verringern und Handlungsoptionen aufzuzeigen. Wir möchten aber auch betonen, dass insbesondere eine gesunde Portion Pragmatismus hierfür notwendig ist. Nicht nur wegen der Frage des Aufwandes, sondern auch für den Umgang miteinander, gerade mit Blick auf Desinformation. Denn, wie es *Hanlon´s Razor* ausdrückt, soll man nicht etwas mit boshafter Absicht erklären, wenn man es auch mit Ignoranz oder Fahrlässigkeit erklären kann.[4] Man hat es also häufiger mit Fehl- als mit Desinformation zu tun.

Thematisch orientieren wir uns an den Problemen, die im Hochschulkontext am häufigsten anzutreffen oder sogar hochschulspezifisch sind:

- Fake News
- Deepfakes und andere KI-Fälschungen
- Halluzinierte Quellen
- Predatory Journals
- Retracted Papers
- Social Engineering

Das vorliegende Papier ist so aufgebaut, dass einer kurzen Einordnung des jeweiligen Themas unmittelbar praktische Werkzeuge zur eigenen Anwendung folgen. Wir hoffen, damit für den Hochschulkontext zum Schließen einer Lücke beizutragen.[5] Wir freuen uns auf Rückmeldungen aus der Praxis – bleiben Sie wachsam und der Wahrheitssuche verbunden!

2 Fake News – (k)ein Thema für die Hochschulen?

2.1 Mit Fake News und Fake Science in die „Vertrauenskrise“

Nach dem Brexit-Votum und Trumps Wahlsieg 2016 rückten Falsch- und Desinformationen im Netz ins Zentrum sozialwissenschaftlicher Debatten – oft diskutiert unter dem Schlagwort des „Postfaktischen", das beschreibt, wie geteilte

[4] Sinngemäß übersetzt aus: https://quoteinvestigator.com/2016/12/30/not-malice/.
[5] Dabei schließen wir insbesondere an die Forderungen der Hochschulrektorenkonferenz (s. „Empfehlung der 40. Mitgliederversammlung“, S. 7) und der Deutschen Forschungsgemeinschaft (DFG) (s. „Positionspapier zur Stärkung der Freiheit und Resilienz der Wissenschaften“, S. 7) an.

Gewissheiten über politische Fakten erodieren, Institutionen an Rückhalt verlieren und Fachleuten weniger geglaubt wird. Die Forschung beobachtet dabei zweierlei: Zum einen greift das durch einseitige Kommunikation genährte Misstrauen längst auf Felder wie Wissenschaft und Gesundheit über, zum anderen können organisierte Akteure oder automatisierte Accounts mit gezielter Desinformation demokratische Willensbildung stören (vgl. Abb. 2).[6] Das Phänomen ist nicht neu. Emotionsgetriebene Meinungsmanipulation vor dem Hintergrund populistischer Polarisierung reiht sich in eine lange Geschichte von Propaganda, politischem Spin und Verschwörungserzählungen.[7]

Positiv ist, dass die Wissenschaft allgemein noch immer hohes Vertrauen genießt. Der gesellschaftliche Nutzen und der zugehörige Rückhalt sind noch nicht erodiert (s. Abb. 3).[8] Die gegenwärtige „Vertrauenskrise" macht das nicht weniger unangenehm und sogar gefährlich für die Hochschulen. Gerade Desinformationen stellen sie nämlich vor eine doppelte Herausforderung: Studierende und Lehrende müssen lernen, manipulative Inhalte in sozialen Medien und Nachrichtenquellen zu erkennen. Gleichzeitig sind Hochschulen selbst Ziel von Desinformationskampagnen, wenn etwa ihre Expertise und ihr guter Ruf genutzt werden, um erfundene Studienergebnisse zu legitimieren. Wenn wissenschaftliche Erkenntnisse (etwa zu Klimawandel, Impfungen oder Pandemien) gezielt verzerrt oder aus dem Kontext gerissen werden, müssen Wissenschaft und Hochschulen sich effektiv wehren können. Zu ihren Aufgaben gehört schließlich, Studierende, Lehrende und Forschende vor irreführenden Informationen bestmöglich zu schützen – besonders auch, wenn es um vergleichsweise junge Menschen geht, die oft besonders im Fokus von Desinformationskampagnen stehen.[9]

Das gilt auch und besonders, wenn die Fake News aus dem eigenen Haus kommen, wie es etwa bei „Fake Science" der Fall ist: Immer wieder sind Wissenschaftler:innen für Fördergelder oder den nächsten Schritt auf der Karriereleiter bereit, Ergebnisse zu publizieren, die wissenschaftlichen Standards nicht genügen und schlimmstenfalls frei erfunden sind.[10] Medienkompetenz und kritischer Umgang mit Quellen sind schon immer Kernkompetenzen akademischer Bildung – und bleiben natürlich auch digital wichtig. Hochschulen müssen allerdings mit Augenmaß entscheiden, wann sich eine Auseinandersetzung lohnt und wann Ignorieren die klügere Antwort ist,[11] um nicht Ressourcen unnötig aufzuwenden. Denn die schiere Menge an Fake News allein verunmöglicht bereits eine detaillierte Antwort für jeden Fall.

[6] Vgl. Yan & Schroeder (2025), S. 1195ff.
[7] Vgl. https://blogs.lse.ac.uk/medialse/2026/05/22/why-we-are-not-in-a-post-truth-era/.
[8] Vgl. DFG, „Positionspapier zur Stärkung der Freiheit und Resilienz der Wissenschaften", S. 6.
[9] S. https://www.stiftungen.org/aktuelles/news-aus-stiftungen/detail/studie-zu-desinformation-von-jungen-menschen-in-der-coronakrise-4986.html.
[10] S. https://www.bib.uni-mannheim.de/services/kurse-und-tutorials/fake-science-in-der-wissenschaft/.
[11] Keefe (2019), S. 5.

Abbildung 2: Zahlen zu Fake News[12]

Abbildung 3: Zahlen zum Vertrauen in die Hochschulen

[12] Die Quelle bezieht sich auf folgende Meldung: https://www.bitkom.org/Presse/Presseinformation/9-von-10-Deutschen-stossen-auf-Fake-News.

2.2 Medienkompetent gegen die Desinformation

Medienkompetenz ist keine institutionelle Qualität: Sie wird vorranging durch Individuen beherrscht, gelebt und weitergetragen. Gut ausgebildete Lehrkräfte sind also für den Umgang mit Fake News an der Hochschule entscheidend. Die entsprechende Sensibilisierung ist ein wichtiger erster Schritt. Nur so können Hochschulen resilient gegenüber irreführenden Informationen sein und ihre Rolle als öffentliche Orte der Wahrheitssuche und Foren der demokratischen Debatte wahrnehmen.[13]

Entscheidend für den souveränen Umgang besonders mit Desinformation ist daneben die Kenntnis der entsprechenden Methoden – und der Werkzeuge, um diese zu enttarnen. Der folgende kompakte Überblick ist eine erste Anlaufstelle, um Fake News im Hochschulkontext und darüber hinaus zu entlarven.

Quelle	Beschreibung	Link
CORRECTIV. Faktencheck	Im deutschsprachigen Raum ist der CORRECTIV.Faktencheck der Goldstandard für die Überprüfung von Medienpublikationen. Hier können auch Beiträge eingeschickt werden.	https://correctiv.org/faktencheck/
Bellingcat	Das Recherche-Kollektiv Bellingcat ist auf Recherche mit modernen Technologien und öffentlich zugänglichen Daten spezialisiert. Auf der Website wird anhand von echten Recherchen detailliert gezeigt, wie sich Fake News entlarven lassen.	https://de.bellingcat.com
BMFTR	Für den Hochschulkontext ist besonders interessant, in welchen Dimensionen und Konstellationen sich Fehlinformationen bewegen. Die seit März 2026 geförderten Projekte des BMFTR zum Thema „Desinformation - erkennen, verstehen, abwehren“ zeigen den aktuellen Stand von Forschung, Technik und gesellschaftlicher Debatte.	https://www.forschung-it-sicherheit-kommunikationssysteme.de/forschung/it-sicherheit/desinformation

Tabelle 1: Wichtige Quellen zum Thema „Fake News“

Die technische Entwicklung macht es notwendig, neben klassischen Fake News auch über neuere Entwicklungen wie Deepfakes, andere KI-Fälschungen und von KI halluzinierte Inhalte zu sprechen. Diese haben die Möglichkeiten der Täuschung massiv

[13] S. https://arbeit-der-zukunft.de/artikel/hochschulen-unter-druck-warum-eine-demokratische-und-gerechte-gesellschaft-resiliente-hochschulen-braucht-labora-2025.

erweitert und haben Fake News und Manipulation von Informationen noch wirkungsvoller gemacht. Die beiden nächsten Kapitel zeigen aber, dass man auch diesen Angriffen auf die Wahrheit nicht schutzlos ausgeliefert ist.

3 Deep Fakes und andere KI-Fälschungen

Während bei den Fake News und der Fake Science vor allem gesellschaftliche Debatten im Fokus stehen, wird es bei Deep Fakes etwas technischer. Durch den rapiden Anstieg der Nutzung von immer leistungsstärker werdender künstlicher Intelligenz sind entsprechende Inhalte heute weit verbreitet.[14] Die von der KI ausgeworfenen Ergebnisse sind zumindest in der ersten Anmutung oft von realen Fotos, Videos und Dokumenten nicht zu unterscheiden.

KI-generierte Text-, Bild-, Audio- und Videofälschungen eröffnen im Hochschulkontext neue Missbrauchsszenarien: von gefälschten Aussagen prominenter Wissenschaftler:innen über manipulierte Prüfungsvideos bis hin zu gezielter Rufschädigung von Lehrenden oder Studierenden. Besonders problematisch ist die zunehmend niedrige technische und soziale Hemmschwelle bei der Erstellung solcher Inhalte. Professionelle Kenntnisse oder schwer zugängliche Software sind kaum noch erforderlich. Hochschulen brauchen sowohl technische Detektionswerkzeuge als auch klare Richtlinien und Sanktionsmechanismen, um dieser Bedrohung institutionell begegnen zu können.

Für die Arbeit in der Hochschule heißt das konkret, vor allem den Kontext zu beachten und im Zweifel zu prüfen. Es gibt keine per se schlechte Technologie, sondern es sind der Rahmen und der Umgang, die Missbrauch ermöglichen oder vermeiden können. Hochschulen sind die Heimat des Experiments, können sich auf der anderen Seite aber keine Spielräume bei der Glaubwürdigkeit und wissenschaftlichen Standards erlauben. In diesem Spannungsfeld gilt es, Ruhe zu bewahren und mit Augenmaß vorzugehen. Während sicherheits- oder strafrechtlich relevante Vorgänge konsequentes Handeln der zuständigen Stellen erfordern, empfehlen wir für den alltäglichen Umgang mit KI-Fälschungen folgendes Vorgehen:

1) Nicht verteufeln – verstehen, Rahmen setzen, Verstöße ahnden.
2) Prüfen, was echt ist: Der von der EU gefördert „Fake News Debunker“ hilft dabei mit KI-Unterstützung.
3) Selbst prüfen: Wer weiß, worauf zu achten ist, kann Fake News und Deep Fakes auch ohne technische Hilfe oft schnell erkennen. Mit dieser Checkliste weiß jede:r, worauf zu achten ist.

[14] S. https://www.insurancetimes.co.uk/news/fraud-attempts-with-deepfakes-surge-over-last-three-years/1454509.article.

4 Halluzinierte Quellen

4.1 Kein blindes Vertrauen in KI-Outputs

Ein Problem im Einsatz von sog. *Large Language Models* (LLMs) ist deren Tendenz, Quellen zu erfinden bzw. zu „halluzinieren".[15] Dabei können einzelne bibliographische Details halluziniert sein – oder auch die gesamte Quelle.[16] Dies betrifft jedoch nicht nur Bibliografie-Einträge, sondern auch Zitate oder sogar Scans (s. Abb. 4). Daher ist es wichtig, zitierte Quellen zu verifizieren – selbst wenn man nicht weiß, ob bei der Erstellung des gegebenen Texts KI zum Einsatz kam.

Zum Glück kann man in vielen Fällen mit wenigen Schritten überprüfen, ob eine Quelle echt ist oder nicht. Das allgemeine Ziel ist es, sie im Netz aufzufinden und dadurch alle ihre Details überprüfen zu können – Autorennamen, Erscheinungsjahr, Titel, Journal beziehungsweise Verlag, Erscheinungsort und gegebenenfalls sogar den konkreten Inhalt. Die folgenden Methoden und Tools sollten dabei helfen, dieses Ziel zu erreichen. Einige von ihnen sind genereller Natur, andere wiederum spezifisch für gewisse Situationen. Sie sind trotzdem so arrangiert, dass man sie mehr oder weniger der Reihe nach anwenden kann.

4.2 „Ohne DOI-Check zitiere ich nichts!"

Viele wissenschaftliche Publikationen erhalten eine sogenannte *DOI* („digital object identifier"), vergeben durch die *DOI Foundation*, eine amerikanische non-profit Organisation. Ziel einer DOI ist es, einer Publikation einen *persistenten* Marker zur Identifikation zu geben, welcher selbst bei Änderungen der Metadaten (beispielsweise Änderung des Namens des Herausgebers oder Autor:in) unverändert bleibt. Die Foundation hat auf ihrer Homepage eine Search Engine, mit der man direkt eine DOI überprüfen kann: https://www.doi.org/.

Alternativ kann man eine DOI auch auf spezifischen Datenbanken oder allgemeinen Suchmaschinen recherchieren (s. 4.3 und 4.4). Häufig sind DOIs aber als Hyperlinks in Artikeln selbst hinterlegt, wodurch mittels Anklicken direkt die entsprechende Website des Verlags beziehungsweise Herausgebers geöffnet wird.

Das Recherchieren der DOI erlaubt häufig eine schnelle Verifikation einer Quelle, inklusive ihrer bibliographischen Details, *sollte* die DOI gültig *und korrekt* sein. Es kann aber vorkommen, dass die DOI entweder gar nicht existiert (*ungültig* ist) oder sie einer Veröffentlichung gehört, die aber nicht die ursprünglich genannte ist (*inkorrekt*). Weder der eine noch der andere Sachverhalt lässt auf eine KI-Halluzination schließen – schließlich kann eine DOI auch falsch abgetippt worden sein!

[15] Vgl. Topaz et al. (2026), S. 1779.
[16] Vgl. Aljamaan et al. (2024), S. 2.

4.3 Suchmaschine: „Google is your friend"

Auch in Zeiten von allgegenwärtigen KI-Anwendungen hat die klassische Webrecherche ihren Wert. Wenn keine DOI vorhanden ist, dann ist die Suche nach Autor:in und/oder dem Titel mittels einer gängigen Suchmaschine wie Google, Bing oder DuckDuckGo eine weitere, schnelle Option. Dadurch gelangt man häufig auf die Website des entsprechenden Herausgebers, wo die Quelle mitsamt ihrer bibliographischen Daten zur Verfügung gestellt wird – allerdings nicht immer gratis oder zugänglich ohne Account. Ebenfalls häufig ist die Weiterleitung auf eine Publikationsdatenbank, wie beispielsweise Google Scholar (s. Abschn. 4.4). Es lohnt sich in der Regel auch, verschiedene Suchmaschinen zu benützen, insbesondere falls der persönliche Favorit nichts Verwertbares findet oder man nach Zitaten sucht (vgl. Abschn. 4.6).

Nur weil eine Publikation echt ist, heißt nicht, dass sie von hoher Qualität ist oder der Herausgeber legitime wissenschaftliche Ziele verfolgt – siehe dazu auch Abschnitt 5 zu Predatory Journals. Wer also gleichzeitig mit der Verifikation einer Quelle auch noch deren Qualitätsprüfung durchführen möchte, sollte sich also nicht mit dem alleinigen Vorhandensein von Google-Links zufriedengeben, sondern sollte diese auch kritisch sichten. Gerade im Hinblick auf die Möglichkeit halluzinierter bibliografischer Details ist dies ohnehin unumgänglich.

4.4 Datenbanken: Alles auf dem Index

Nicht alle echten – geschweige denn qualitativ hochwertigen – Publikationen haben eine DOI. Denn um DOIs vergeben zu dürfen, muss ein Herausgeber der DOI Foundation entsprechende Beträge bezahlen, was lokale Journals, Verlage und Konferenzen häufig nicht tun (können). Das Fehlen einer DOI ist also insbesondere kein Indiz für eine KI-Halluzination!

Wie bereits oben erwähnt, führt die Nutzung von Suchmaschinen schnell zu Online-Datenbanken. In der Regel kann dort gezielt mit Autorennamen, Titel und/oder Erscheinungsjahr gesucht werden, was die Verwendung von Datenbanken gegenüber Suchmaschinen zielführender machen kann. Allerdings helfen Datenbanken je nachdem wenig, wenn es um die Verifikation von Zitaten oder Textausschnitten geht. Denn je nach Quelle sitzt der Volltext auch hier hinter einer Paywall (s. Abschn. 4.6).

Es gibt allgemeine Datenbanken für wissenschaftliche Publikationen sowie fachspezifische. Zur ersten Kategorie gehören unter anderem:

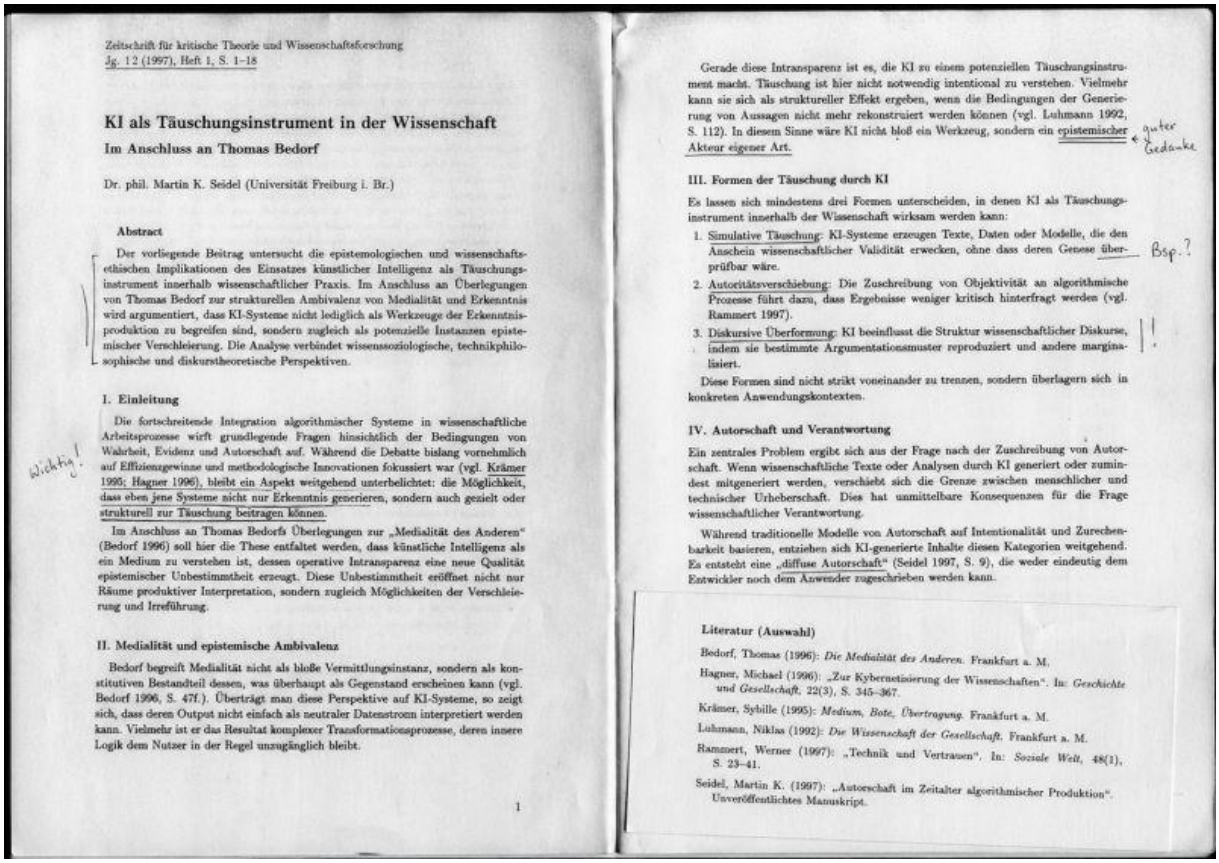
- **Google Scholar**: https://scholar.google.com/
  Eine enorm große Datenbank, die nebst gewöhnlicher akademischer Literatur auch Patente oder Gerichtsentscheide indiziert.
- **Semantic Scholar**: https://www.semanticscholar.org/
  KI-gestützte Suchmaschine, die jedoch auf einem echten Index basiert und somit keine Halluzinationsprobleme aufweist.

- **BASE** (Bielefeld Academic Search Engine): https://www.base-search.net/
  Größte deutsche Online-Datenbank für wissenschaftliche Publikationen.
  Speziell ist die hohe Anzahl an indizierten Open-Access-Publikationen.

Spezifische Datenbanken bestehen praktisch für jeden Fachbereich, beispielsweise **Pubmed** (https://pubmed.ncbi.nlm.nih.gov/, Medizin), **Philpapers** (https://philpapers.org/, Philosophie) oder **Project Euclid** (https://projecteuclid.org/, Mathematik). Viele solcher Datenbanken haben einen hohen Anteil an Open-Access-Veröffentlichungen und indizieren teilweise auch Preprints. Fachspezifische Datenbanken findet man entweder durch die Nutzung von Suchmaschinen (Google, Bing, DuckDuckGo, etc., s. Abschn. 4.3) oder mittels Meta-Datenbanken wie die **DBIS** (https://dbis.uni-regensburg.de/), welche Datenbanken an sich indiziert und nach verschiedenen Kategorien (Fachbereich, Art des Hostings, etc.) sortiert.

4.5 ORCID: Der DOI für Forscher:innen

Eine weitere Option, insbesondere bei Scans oder Screenshots, ist es, die *ORCIDs* („Open Researcher and Contributor ID") der Autor:innen zu überprüfen, falls welche angegeben werden. Genau wie bei der DOI geht es bei ORCID darum, persistente Identifikatoren zu verleihen. Im Gegensatz zur DOI Foundation kann jede:r Forscher:in eine ORCID erhalten, ganz ohne Kosten. Eine ORCID lässt sich darüber hinaus schnell auf der Homepage der ORCID Organisation selbst mittels einer Search Engine überprüfen: https://orcid.org/. Häufig sind dort auch die entsprechenden Publikationen aufgelistet, wodurch die Zuordnung sowie weitere Details ebenfalls schnell überprüft werden können.



Zeitschrift für kritische Theorie und Wissenschaftsforschung
Jg. 12 (1997), Heft 1, S. 1–18

KI als Täuschungsinstrument in der Wissenschaft

Im Anschluss an Thomas Bedorf

Dr. phil. Martin K. Seidel (Universität Freiburg i. Br.)

Abstract

Der vorliegende Beitrag untersucht die epistemologischen und wissenschaftsethischen Implikationen des Einsatzes künstlicher Intelligenz als Täuschungsinstrument innerhalb wissenschaftlicher Praxis. Im Anschluss an Überlegungen von Thomas Bedorf zur strukturellen Ambivalenz von Medialität und Erkenntnis wird argumentiert, dass KI-Systeme nicht lediglich als Werkzeuge der Erkenntnisproduktion zu begreifen sind, sondern zugleich als potenzielle Instanzen epistemischer Verschleierung. Die Analyse verbindet wissenssoziologische, technikphilosophische und diskurstheoretische Perspektiven.

I. Einleitung

Die fortschreitende Integration algorithmischer Systeme in wissenschaftliche Arbeitsprozesse wirft grundlegende Fragen hinsichtlich der Bedingungen von Wahrheit, Evidenz und Autorschaft auf. Während die Debatte bislang vornehmlich auf Effizienzgewinne und methodologische Innovationen fokussiert war (vgl. Krämer 1995; Hagner 1996), bleibt ein Aspekt weitgehend unterbelichtet: die Möglichkeit, dass eben jene Systeme nicht nur Erkenntnis generieren, sondern auch gezielt oder strukturell zur Täuschung beitragen können.

Im Anschluss an Thomas Bedorfs Überlegungen zur „Medialität des Anderen" (Bedorf 1996) soll hier die These entfaltet werden, dass künstliche Intelligenz als ein Medium zu verstehen ist, dessen operative Intransparenz eine neue Qualität epistemischer Unbestimmtheit erzeugt. Diese Unbestimmtheit eröffnet nicht nur Räume produktiver Interpretation, sondern zugleich Möglichkeiten der Verschleierung und Irreführung.

II. Medialität und epistemische Ambivalenz

Bedorf begreift Medialität nicht als bloße Vermittlungsinstanz, sondern als konstitutiven Bestandteil dessen, was überhaupt als Gegenstand erscheinen kann (vgl. Bedorf 1996, S. 47f.). Überträgt man diese Perspektive auf KI-Systeme, so zeigt sich, dass deren Output nicht einfach als neutraler Datenstrom interpretiert werden kann. Vielmehr ist er das Resultat komplexer Transformationsprozesse, deren innere Logik dem Nutzer in der Regel unzugänglich bleibt.

1

Gerade diese Intransparenz ist es, die KI zu einem potenziellen Täuschungsinstrument macht. Täuschung ist hier nicht notwendig intentional zu verstehen. Vielmehr kann sie sich als struktureller Effekt ergeben, wenn die Bedingungen der Generierung von Aussagen nicht mehr rekonstruiert werden können (vgl. Luhmann 1992, S. 112). In diesem Sinne wäre KI nicht bloß ein Werkzeug, sondern ein epistemischer Akteur eigener Art.

III. Formen der Täuschung durch KI

Es lassen sich mindestens drei Formen unterscheiden, in denen KI als Täuschungsinstrument innerhalb der Wissenschaft wirksam werden kann:

1. Simulative Täuschung: KI-Systeme erzeugen Texte, Daten oder Modelle, die den Anschein wissenschaftlicher Validität erwecken, ohne dass deren Genese überprüfbar wäre.
2. Autocitätsverschiebung: Die Zuschreibung von Objektivität an algorithmische Prozesse führt dazu, dass Ergebnisse weniger kritisch hinterfragt werden (vgl. Rammert 1997).
3. Diskursive Überformung: KI beeinflusst die Struktur wissenschaftlicher Diskurse, indem sie bestimmte Argumentationsmuster reproduziert und andere marginalisiert.

Diese Formen sind nicht strikt voneinander zu trennen, sondern überlagern sich in konkreten Anwendungskontexten.

IV. Autorschaft und Verantwortung

Ein zentrales Problem ergibt sich aus der Frage nach der Zuschreibung von Autorschaft. Wenn wissenschaftliche Texte oder Analysen durch KI generiert oder zumindest mitgeneriert werden, verschiebt sich die Grenze zwischen menschlicher und technischer Urheberschaft. Dies hat unmittelbare Konsequenzen für die Frage wissenschaftlicher Verantwortung.

Während traditionelle Modelle von Autorschaft auf Intentionalität und Zurechenbarkeit basieren, entziehen sich KI-generierte Inhalte diesen Kategorien weitgehend. Es entsteht eine „diffuse Autorschaft" (Seidel 1997, S. 9), die weder eindeutig dem Entwickler noch dem Anwender zugeschrieben werden kann.

Literatur (Auswahl)

Bedorf, Thomas (1996): *Die Medialität des Anderen*. Frankfurt a. M.

Hagner, Michael (1996): „Zur Kybernetisierung der Wissenschaften". In: *Geschichte und Gesellschaft*, 22(3), S. 345–367.

Krämer, Sybille (1995): *Medium, Bote, Übertragung*. Frankfurt a. M.

Luhmann, Niklas (1992): *Die Wissenschaft der Gesellschaft*. Frankfurt a. M.

Rammert, Werner (1997): „Technik und Vertrauen". In: *Soziale Welt*, 48(1), S. 23–41.

Seidel, Martin K. (1997): „Autorschaft im Zeitalter algorithmischer Produktion". Unveröffentlichtes Manuskript.

Abbildung 4: Ein Beispiel für einen „Scan-Deepfake"[17]

4.6 Für Hobby-Detektive: Leseproben und Preprints

Während man mit den bisherigen Tools gut Bibliographie-Einträge verifizieren kann, ist damit die Möglichkeit von KI-halluzinierten Zitaten oder Scans noch nicht ausgeschlossen. Falls es sich bei der Quelle um eine Open-Access-Publikation handelt, ist das Verifizieren eines Scans oder Zitats schnell erledigt. Häufig aber ist der Volltext hinter einer Paywall – selbst bei vorhandenem institutionellem Zugang durch eine (Hochschul-)Bibliothek.

Hier schaffen drei Optionen Abhilfe:

1. **Kritische Lektüre**: Häufig erscheint gerade ein Scan nur auf den ersten Blick real. Wer sich aber die Zeit nimmt und die (vermeintliche) Quelle aufmerksam durchliest, kann gegebenenfalls bereits Ungereimtheiten aufdecken. Im gegebenen Bild (s. Abb. 4) beispielsweise sollte schnell klar werden, dass es eher unwahrscheinlich ist, dass im Jahre 1997 bereits etwas zu generativer KI

[17] Das Beispiel wurde entnommen von: https://www.linkedin.com/posts/andreas-giesbert-773968262_chatgpt-ki-taeuuschung-activity-7454449729414979584-ALgX?utm_source=share&utm_medium=member_desktop&rcm=ACoAADZw0WABJ6bLhyOfJqfFqILdAvXIV4erk4I.

mit diesem Inhalt publiziert wurde. Zudem ist die Einleitung für eine echte Publikation zu knapp bemessen – und so weiter und so fort.

2. **Zitatsuche**: Man kann ein Zitat bei den meisten Suchmaschinen auch als ganzen Textabschnitt suchen, indem man es in Anführungs- und Schlusszeichen setzt. Dadurch wird man hin und wieder auf Leseproben verlinkt, die den entsprechenden Abschnitt der Publikation einsehbar zur Verfügung stellen, wie beispielsweise auf Google Scholar oder der sogenannten **Wayback Machine**, einem Archiv für Internet-Websites (http://web.archive.org/).
3. **Preprints**: Viele Autor:innen veröffentlichen ihre Manuskripte vorab auf ihrer Homepage oder einer (fachspezifischen) Datenbank für Preprints. Obschon sich der Inhalt eines Preprints vom Text der publizierten Version unterscheiden kann, lässt sich mittels einem Preprint häufig verifizieren, ob ein Zitat oder ein Seiten-Scan echt ist. Die größten Datenbanken für Preprints sind **arXiv** (https://arxiv.org/) sowie seine fachspezifischen Variationen wie beispielsweise **bioRxiv** (https://www.biorxiv.org/) für Life Sciences.

4.7 Mut zum Kontakt

Zu guter Letzt kann man versuchen, die Autor:innen wissenschaftlicher Arbeiten direkt zu kontaktieren. Nicht nur sind viele Akademiker:innen gerne bereit, von ihrer Arbeit zu berichten, sondern eine Kontaktaufnahme kann gleichzeitig ein Karrierebonus im Sinne des Netzwerkens sein. Gerade bei großer fachlicher Relevanz einer Quelle ist eine Kontaktaufnahme nur zu empfehlen!

4.8 Letzte Bemerkungen

Die in den letzten Abschnitten dargestellten Optionen sollten insbesondere dabei helfen, halluzinierte Zitate oder Scans als solche entlarven zu können. Natürlich handelt es sich hier nicht um die einzigen Optionen, noch sind diese immer einfach durchzuführen. Das Suchen nach Zitaten oder Leseproben kann einige Zeit in Anspruch nehmen, je nachdem, wie obskur die (vermeintliche) Quelle ist. Darüber hinaus gäbe es noch technisch ausgefeiltere Methoden wie inverse Bildsuche sowie die Analyse von Metadaten bei Bildern, die aber an dieser Stelle zu weit führen würden. Unsere Empfehlung ist es, den Aufwand pragmatisch anhand der Relevanz der Quelle zu bestimmen: bei einem zu übernehmenden Zitat sollte der Aufwand sicher größer sein als bei einer Quelle, die als eine von vielen einen weitbekannten Fakt bestätigt und dessen Nennung nur überlesen wird.

5 Predatory Journals

Raubverlage, die gegen Gebühr wissenschaftliche Artikel ohne echte Qualitätssicherung publizieren, untergraben die Integrität des wissenschaftlichen Publikationssystems. Für Nachwuchswissenschaftler:innen, die unter Publikationsdruck stehen, sind sie besonders gefährlich: Wer unwissentlich in einem Predatory Journal veröffentlicht, riskiert nicht nur seinen Ruf, sondern trägt auch zur Verbreitung nicht peer-reviewter und potenziell fehlerhafter Forschungsergebnisse bei. Hochschulen sind gefordert, Forschende frühzeitig über Erkennungsmerkmale unseriöser Verlage aufzuklären und verlässliche Orientierung zu geben.

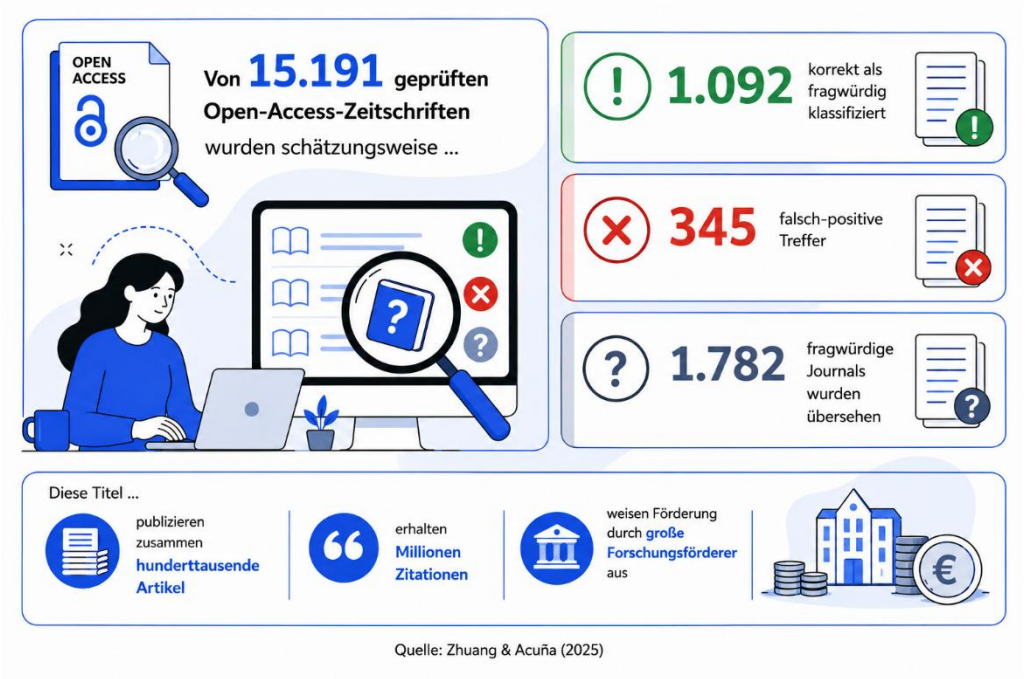


Abbildung 5: Zahlen aus der Studie von Zhuang & Acuña (2025)

Problematisch wird es, wenn Artikel aus Raubverlagen in öffentlichen Debatten zitiert werden. Formal weisen sie DOI, Literaturangaben und den Anschein eines Peer-Review-Verfahrens auf und lassen sich daher kaum von seriösen Publikationen unterscheiden. Das macht sie anfällig für Instrumentalisierung: Eine unbelegte These zu Impfungen, Klimawandel oder Ernährung erhält allein durch die Form der Veröffentlichung einen wissenschaftlichen Anstrich, den der Inhalt nicht rechtfertigt (vgl. Abb. 5).

Raubverlage lassen sich an mehreren Merkmalen erkennen:

- Die Begutachtung dauert oft nur wenige Tage. Ein seriöses Peer-Review-Verfahren braucht Wochen, manchmal Monate. Wer Manuskripte schneller publizieren kann, dem gegenüber sollte man misstrauisch werden.
- Häufig beginnt der Kontakt mit einer unaufgeforderten E-Mail. Der Ton ist auffällig schmeichelhaft. Die eigene Forschung wird als „bahnbrechend" oder

„von großem Interesse" bezeichnet, oft ohne inhaltlichen Bezug. Das ist Marketing, keine fachliche Einschätzung.
- Redaktionsadressen fehlen oder bleiben vage. Manchmal findet sich nur ein Kontaktformular, kein Impressum, kein Standort. Auch Herausgebergremien sind ein Warnsignal. Bekannte Namen werden gelistet, ohne dass die Betroffenen davon wissen. Eine kurze Rückfrage bei den genannten Personen reicht oft, um das aufzudecken.
- Die Kosten werden selten vorab transparent gemacht. Erst nach Annahme des Artikels folgt die Rechnung. Dann steht die eigene Publikation bereits im Raum, der Druck zu zahlen ist entsprechend hoch.

Wer konkrete Zweifel an der Seriosität eines Papers hat, ist bei den folgenden Adressen gut aufgehoben, um kritisch zu prüfen:

1) Das Ombudsgremium für die wissenschaftliche Integrität in Deutschland hat eine detaillierte Handreichung zu Predatory Journals veröffentlicht, die als Orientierung hilfreich ist.
2) Das Directory of Open Access Journals listet qualitätsgeprüfte Open-Access-Zeitschriften mit Peer Review. Um gelistet zu werden, müssen Zeitschriften ein aufwendiges Bewerbungsverfahren durchlaufen.
3) Think, Check, Submit ist ein praktisches Online-Tool im Stil einer Checkliste, mit der Forschende testen können, ob der Verlag, der ihre Beiträge veröffentlichen will, seriös und wissenschaftlich ist.

6 Retracted Papers

Etwas anders als in den Fällen von halluzinierten Quellen, künstlich erstellten Fotos oder Fake News sind zurückgezogene Studien („retracted papers") zwar echt, wurden aber nach ihrer Publikation wieder zurückgezogen. Dies kann aus diversen Gründen geschehen, von unbeabsichtigter Verletzung von Urheberrechten bei einzelnen Grafiken über manipuliertes Peer-Review bis hin zur Verwendung von anderen zurückgezogenen Studien als Quelle.[18] Auch hierzu gibt es paar Tipps für den Umgang, deren Umsetzung sich im Allgemeinen einfach gestaltet:

Tipp #1: (Routine-)Checks zum Status einer Veröffentlichung

An erster Stelle muss natürlich zuerst herausgefunden werden, ob eine Studie überhaupt zurückgezogen wurde oder nicht. Hierzu gibt es einschlägige Datenbanken:

- **Retraction Watch:** https://retractionwatch.com/ (beziehungsweise https://retractiondatabase.org/ für die Datenbank selbst)

[18] S. bspw. https://retractionwatch.com/retraction-watch-database-user-guide/retraction-watch-database-user-guide-appendix-b-reasons/.

Große Datenbank für zurückgezogene Studien, welche immer auch gleich die Gründe für das Zurückziehen angibt.

- **Crossref:** https://www.crossref.org/
  Eine Datenbank, in welcher man nach Autor:in, Titel, DOI, etc. suchen kann und der Status über eine Zurückziehung klar ausgewiesen wird.

Am besten macht man einen solchen Check aber gleich bei der Verifikation beziehungsweise Recherche der Quelle. Häufig wird auf den Websites der Herausgeber angegeben, ob die Studie zurückgezogen wurde oder nicht. Allerdings: dies ist längstens nicht immer gleich transparent dargestellt (vgl. Abb. 6 und 7) – die Website muss also teilweise vorsichtig gelesen werden!

Tipp #2: Gründe recherchieren

Je nachdem, wie wichtig die Quelle für die eigene Arbeit oder Meinungsbildung ist, könnte es wichtig sein, kritisch nachzulesen, *weshalb* eine Studie zurückgezogen wurde. Wie oben erwähnt gibt es zahlreiche Gründe, weshalb eine Studie zurückgezogen werden kann. Dabei ist es für den Umgang mit den Aussagen der Studie entscheidend, ob sie wegen einer Urheberrechtsverletzung zurückgezogen wurde oder ob ihr Design schlecht war. Je nach Grund kann im entsprechenden Journal, bei den Autor:innen oder bei Zitierungen der Quelle durch andere Publikationen weiterverfolgt werden, ob und inwiefern die Aussagen dennoch stichhaltig waren. Nicht selten werden Studien wiederholt oder angepasst – meistens durch dieselben Autor:innen.

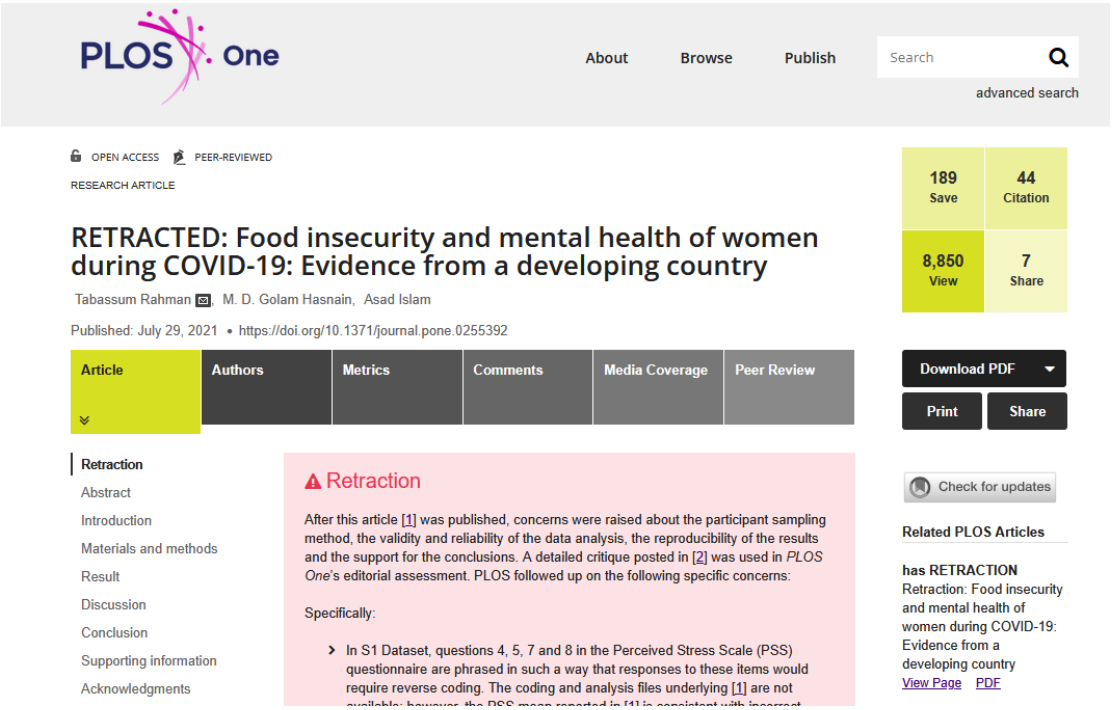


Abbildung 6: Ein Beispiel eines Hinweises auf „Retraction": markant und daher sofort ersichtlich

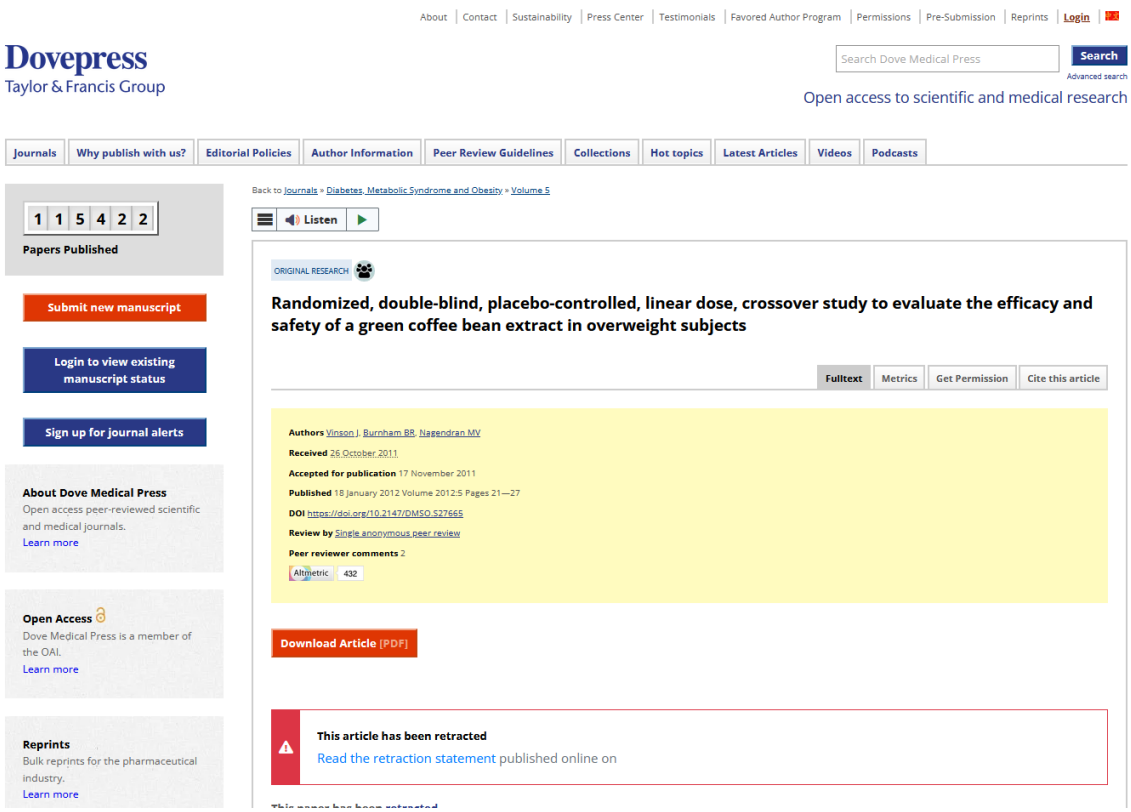


Abbildung 7: Ein weiteres Beispiel eines Hinweises auf „Retraction" – allerdings wesentlich weniger deutlicher als in Abb. 6, wobei die Originalität des Papers immer noch zuerst genannt wird!

Tipp #3: Weiterer Quality-Check

Selbst wenn eine Studie nicht zurückgezogen wurde, kann sie dennoch kontrovers sein. Das muss nicht heißen, dass ihr Inhalt wissenschaftlichen Standards nicht genügt, geschweige denn, dass ihre Aussagen falsch sind. Allerdings kann es für die eigene Meinungsbildung und Sprechfähigkeit – besonders wenn man sich in der eigenen Arbeit auf sie verlassen möchte – wichtig sein, dass man sich ein Bild über die Resonanz von gewissen Studien macht.

Eine klassische Plattform hierzu ist PubPeer (https://pubpeer.com/), welche dem sogenannten *Post-Publication Peer Review* gewidmet ist, also dem Bewerten wissenschaftlicher Literatur durch Experten aus der Community *nach* der Publikation. Dies kann eine bereichernde Quelle von kritischen Kommentaren sein, insbesondere auch jenseits nur aufbauender, kritischer, weiterer Publikationen. Es sei aber darauf hingewiesen, dass die meisten Kommentare dort anonym verfasst sind – mit den entsprechenden Konsequenzen bezüglich Fairness und Inhalt. Dennoch kann es, zusammen mit weiterer Recherche auf anderen Foren, einen Beitrag zur eigenen Meinungsbildung leisten.

Zu guter Letzt soll an dieser Stelle noch auf ein weiteres Phänomen aufmerksam gemacht werden, was je nachdem nicht ganz die Anforderungen für Fehl- oder Desinformation erfüllt, aber in eine ähnliche Richtung geht. Es gibt wissenschaftliche Veröffentlichungen, die nach wie vor zitiert werden, nie zurückgezogen wurden, aber

dennoch in der einen oder anderen Hinsicht als überholt gelten. Die Gründe für die weiteren Zitierungen können vielfältig sein: abgesehen von kritischer Auseinandersetzung mit dem Werk können dabei schlichte Ignoranz oder sogenannte „Zombie-Zitierungen" Ursachen sein. Letzteres bezeichnet das unkritische (Weiter-) Zitieren von Papers nur, weil die eigene Community dies ebenfalls tut.

Beispiel: In einem viel-zitierten Paper berechnen Murtaugh und Schlax (2009) den CO_2-Fußabdruck spezifisch von Eltern *qua* Eltern und kommen beispielsweise auf eine Summe von 9441t pro Mutter, was das 5.7-fache deren lebenslangen Emissionen darstellt.[19] Diese Zahlen werden in der Reproduktionsethik häufig unkritisch entgegengenommen.[20] Wie aber beispielsweise Pinkert und Sticker (2025) argumentieren beziehungsweise bemerken, gingen diese Berechnungen von unveränderten Emissionen der Nachfolgegenerationen bis zum Jahr 2400 aus. Die durchschnittlichen Emissionen haben aber bereits in vielen Teilen abgenommen – ein Umstand, der bei der kontinuierlichen Zitierung von Murtaugh und Schlax (2009) häufig vergessen geht.[21]

Die Wahrscheinlichkeit, solche überholten Ergebnisse weiter zu propagieren, ist (idealerweise) umgekehrt proportional zum eigenen Fachwissen im entsprechenden Gebiet. Es ist daher beim Einbeziehen *fachfremder* Literatur ratsam, entsprechende Kolleg:innen zu Rate zu ziehen oder zumindest auch kritische, später publizierte Gegenliteratur zu lesen – ganz im Sinne des kritischen Denkens, welches für alle hier besprochenen Themen zentral ist!

7 Social Engineering

7.1 Der Mensch als digitaler Risikofaktor

Zur Abrundung der bisher vorgestellten Techniken und Methoden zum Umgang mit irreführenden Informationen sollten wir uns noch dem Thema der digitalen Resilienz im Sinne von *Cybersicherheit* zuwenden. Das Thema ist breit und komplex: Es beinhaltet verschiedene Aspekte und Facetten, die sich nicht selten auf die technische Ebene beziehen. Im Schnittgebiet zum Thema Desinformation rücken wir aber den „Faktor Mensch" in den Fokus. Obschon Cybersicherheit eine zentrale Aufgabe der Hochschulverwaltung und -leitung ist, ist gerade die „Schwachstelle" des Individuums von entscheidender Bedeutung. Zu häufig gelingt es Kriminellen und anderen Angreifern, sich über einzelne Individuen Zugang zu internen Netzwerken zu verschaffen.[22] **Dem**

[19] Murtaugh & Schlax (2009), S.18.
[20] Vgl. bspw. Conly (2016), Earl et al. (2016) und Hedberg (2020).
[21] Pinkert & Sticker (2025), S. 48f.
[22] Vgl. hierzu das Interview des Deutschlandfunks mit Maren Lübcke: https://www.deutschlandfunk.de/warum-cyberkriminelle-oft-hochschulen-angreifen-interview-mit-maren-luebcke-100.html.

entgegenzuwirken ist also zuletzt auch eine Aufgabe jeder Person an einer Hochschule – Studierende und Promovierende inbegriffen.

In der Schnittmenge von Cybersicherheit und Desinformation steht der Begriff des *Social Engineerings*. Dabei geht es um absichtliche Täuschungen und Verbreitung von Desinformation durch Angreifer:innen, um an sensible Informationen zu kommen, Sicherheitsmaßnahmen zu umgehen oder Schadsoftware (auch: *Malware*) in ein System einzuschleusen.[23] Dieses Vorgehen ist im Grunde nichts Neues – mit der Digitalisierung hat es aber neue Facetten bekommen, was durch das Aufkommen von generativer KI nochmals beschleunigt wurde.[24] Die folgenden Ausführungen haben zwei Ziele: über die möglichen Formen von Social Engineering zu informieren und hilfreiche Tipps zum Schutz mitzugeben. Dies sollte in einer entsprechenden Wachsamkeit resultieren, die ihrerseits eine grundlegende und effektive Verteidigungslinie bietet.[25]

7.2 Phishing: Fischen nach Zugangsdaten mittels Desinformation

Die üblichste Form von Social Engineering ist das sogenannte *Phishing*,[26] d.h. das Fischen nach Passwörtern und anderen sensiblen Daten.[27] Dies geschieht, indem ein:e Angreifer:in sich als Bank, IT-Abteilung, Vorgesetzte:r oder ein:e Bekannte:r ausgibt, um mittels diversen Druckmechanismen oder Betrugsmaschen die Preisgabe von sensiblen Informationen zu bewirken. Dies kann auch dadurch geschehen, dass man auf täuschend echt aussehenden Webseiten umgeleitet wird, wo beispielsweise durch das Eingeben von Formulardaten die Informationen gewonnen werden oder durch den Besuch der Seite allein bereits Malware eingeschleust wird. Dabei ist es wichtig zu betonen, dass Phishing vor allem dann zum Zug kommt, wenn technische Schwachstellen nicht (einfach) ausnutzbar sind.

Je nachdem, über welches Medium die Kontaktaufnahme erfolgt, kommen auch anderen Namen zum Einsatz:[28]

- *Phishing* bei E-Mails
- *Vishing* bei Anrufen übers Telefon[29]
- *Smishing* über SMS
- *Quishing* über QR-Codes

[23] Vgl. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Social-Engineering/social-engineering_node.html.
[24] Vgl. Bundesamt für Sicherheit in der Infromationstechnik (BSI), „Künstliche Intelligenz sicher nutzen", S. 16.
[25] S. BSI, „Basisschutz für sichere E-Mail-Nutzung".
[26] Vgl. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Social-Engineering/social-engineering_node.html.
[27] Vgl. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Spam-Phishing-Co/Passwortdiebstahl-durch-Phishing/passwortdiebstahl-durch-phishing_node.html.
[28] S. https://www.bundeskriminalamt.at/212/internet/phishing_smishing.html.
[29] S. https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-bedrohungen/phishing.html.

Gerade Phishing-E-Mails werden häufig massenhaft versendet, mit dem Kalkül, dass mindestens ein paar Personen darauf reinfallen werden. Dank generativer KI lässt sich dies heutzutage noch einfacher gestalten.[30]

Demgegenüber steht das sogenannte *Spear-Phishing*, bei welchem die Täter:innen gezielt über Gruppen oder Einzelpersonen vorab recherchieren. Dies dient dazu, die Täuschung noch effizienter zu gestalten. Entsprechend ist das Spear-Phishing schwieriger zu erkennen und abzuwehren. Wir mögen eher nicht dazu verleitet sein, dem vermeintlichen Anwalt aus dem Ausland mit einer zu vergebenden Summe in Millionenhöhe zu glauben, wohl aber einer vorgesetzten Person, die uns von ihrer tatsächlichen E-Mail-Adresse aus anzuschreiben scheint.

Das Erhalten von sensiblen Informationen dient dabei verschiedenen Zwecken, je nach Identität und Absicht des Angreifers. Häufig stehen bei kriminellen Akteuren der finanzielle Gewinn im Vordergrund. So kann man sich mittels Zugriffs auf ein Online-Shopping- oder Bank-Konto illegitim bereichern, oder aber man erpresst insbesondere Organisationen durch die Androhung eines Leaks der gestohlenen Informationen – ein sogenannter *Ransomware-Angriff*.[31] Manchmal stehen aber auch andere Motive im Hintergrund: vom Anrichten von (politisch motivierten) Reputationsschäden durch Leaks über Spionage durch staatliche Akteure bis hin zu Sabotage-Angriffen auf kritische Infrastrukturen.[32] **Phishing – sowie Social Engineering im Allgemeinen - ist also alles andere als harmlos!**

7.3 Schutz vor Phishing: Basics und Vigilance

Um dieser Gefahr souverän entgegenzutreten, gibt es verschiedene Methoden und Tricks, um auf diversen Ebenen mit solchen Angriffen umzugehen. Diese sind in grob zwei Kategorien unterteilt: *Basic Cybersecurity* und *Vigilance*. Wir beginnen mit Ersterem.

Es gibt etliche Richtlinien und „Good-Practice"-Beispiele zu allgemeiner Cybersicherheit. Diese sollten so oder so befolgt werden – unabhängig davon, ob man sich spezifisch gegen Social Engineering im Cyberraum schützen möchte oder nicht. Hier listen wir vor allem diejenigen auf, die keinen spezifischen Bezug zu Social Engineering haben, sondern deren Nicht-Befolgung das Phishing im Vornherein überflüssig machen würde - einfach weil der Akteur sich bei Nichtbeachtung die Mühe der Täuschung sparen

[30] S. BSI, „Künstliche Intelligenz sicher nutzen", S. 16f.
[31] S. bspw. https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-unternehmen/aktuelle-bedrohungen/ransomware.html. So jüngst geschehen im Saarland, s. https://www.zeit.de/news/2026-07/10/jugendlicher-hacker-soll-uni-des-saarlandes-erpresst-haben.
[32] S. Bundesamt für Verfassungsschutz (BfV), „BfV CYBER INSIGHT, Teil 1", S. 2f., Direktion Staatsschutz und Nachrichtendienst (DSN), „Verfassungsschutzbericht 2024", S. 132f. und Nachrichtendienst des Bundes (NDB), „Jahresbericht 2025", S. 70f.

könnte! Die folgende Liste beinhalten einige dieser allgemeinen Tipps, ohne dabei Anspruch auf Vollständigkeit zu erheben:

1. **Passwort-Management**
   Wer kein starkes Passwort verwendet, bei dem muss auch gar nicht erst „gephisht" werden. Es ist heutzutage einfach, vollständig automatisiert ganze Wörterbücher sowie Standard-Kombinationen durchzuprobieren. Es ist also absolut zentral, ein möglichst starkes Passwort zu verwenden. Insbesondere sollten folgende Punkte beachtet werden:[33]

   a. Keine Standard-Kombinationen, Geburtstage oder einzelne Wörter versehen mit einem Sonderzeichen.
   b. Ein Passwort sollte man sich möglichst gut merken können. Eine Möglichkeit, dies zu bewerkstelligen, wäre die Verwendung eines Satzes oder längeren Worts, welches man rückwärts schreibt und Zahlen sowie Sonderzeichen nach Belieben einführt. Damit lassen sich Eselsbrücken zum Passwort herstellen.
   c. Mindestens 12 Zeichen, wobei möglichst viele Arten von Zeichen benutzt werden (Klein- und Großbuchstaben, Ziffern, Sonderzeichen) - je länger, desto besser!
   d. Man sollte Passwörter nicht für mehrere Zugänge gleichzeitig verwenden – jeder Zugang benötigt ein neues, starkes Passwort! **Im Minimum sollten für die Zugänge an der Hochschule neue, starke Passwörter verwendet werden, die sich nicht mit denen im privaten Leben decken.** Dadurch verhindert man, dass ein Leak auf der einen Seite sich auf die andere auswirkt.
   e. Man soll wachsam bei der Eingabe von Passwörtern sein und sicherstellen, dass niemand zuschaut.

   Natürlich kann sich kein Mensch Dutzende von Passwörtern merken. Es empfiehlt sich daher die Verwendung von *Passwort-Wallets* bzw. sogar *Passkeys*, um die Problematik zu beseitigen.

   Letztlich gilt aber **immer: Passwörter werden nicht an Dritte weitergegeben!**

2. **Multi-Faktor-Authentifizierung**
   Wo möglich sollte mindestens eine Zwei-Faktor-Authentifizierung genutzt

[33] Vgl. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/sichere-passwoerter-erstellen_node.html und https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-themen/schuetzen-sie-ihre-konten.html.

werden. Damit kann selbst beim Diebstahl des Passworts der Zugang verwehrt bleiben. Meistens handelt es sich bei dem zweiten Faktor um einen SMS-Code, eine Ziffernfolge in einer App oder die Eingabe biometrischer Daten.[34] Grundsätzlich besitzen Hochschulen über Zwei-Faktor-Authentifizierungen für ihre IT-Dienste - das Angebot muss aber auch durch die Hochschulangehörigen genutzt werden!

Wichtig: niemals einen Zugang, den man nicht selbst angefordert hat, genehmigen!

3. **Aktualisieren von Software**

   Manchmal nutzen Cyberangrifer:innen auch Schwachstellen in einem Betriebssystem, einer App oder sonstiger Software. Diese werden regelmäßig von den Entwicklern beseitigt, weswegen das regelmäßige – am besten: *automatische* - Aktualisieren von Software einen grundlegenden und wichtigen Schutz bietet. Dies betrifft nicht nur den eigenen Desktop oder Laptop, sondern auch Smartphones, Tablets und weitere elektronische Geräte mit Internetanschluss.[35] **Ganz besonders gilt dies für alle elektronische Geräte, die auf dem Campus zur Verfügung gestellt werden oder auf den Campus gebracht werden!**

4. **Schutz der Geräte vor Fremdzugriffen**

   Selbst die beste Firewall hilft wenig, wenn man sich ein Gerät auch physisch verschaffen und ohne Zugangssperre nutzen kann. Nebst dem Einrichten von Zugangssperren bei Smartphones, PCs und anderen Geräten mittels Passwörtern, PINs, Mustereingaben oder biometrischen Scans (Fingerbadruck, Gesichtserkennung, etc.) gehört hier insbesondere dazu, dass **man seine Geräte nicht unbeaufsichtigt lässt und bei Abwesenheit im Minimum die Zugangssperre aktiviert hat.**[36]

   Dies gilt insbesondere für PCs, die auf dem Campus für die Arbeit benutzt werden!

5. **Gesunde Skepsis in der Nutzung von Apps, Plug-Ins sowie Add-Ons**

   Bezüglich der Nutzung von Apps und Software-Erweiterungen – egal ob auf dem PC, dem Smartphone oder einem anderen Gerät, gelten grundsätzlich folgende Prinzipen:[37]

[34] Vgl. BSI, „Soziale Netzwerke sicher nutzen“, Tipp 2.
[35] S. BSI, „Smartphone, Tablet und Co sicher nutzen“, Tipp 1, und https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-themen/geraeteschutz.html.
[36] S. ebenda.
[37] Vgl. u.a. BSI, „Das Internet der Dinge sicher nutzen“, Tipp 8.

a. Software nur aus vertrauenswürdigen Quellen, wie beispielsweise offizielle Hersteller-Webseiten, beziehen.
b. Bei Unsicherheit: lieber vorher einmal zu viel zu Reviews oder Nutzererfahrungen recherchieren als einmal zu wenig.
c. Die Datenschutzregelungen von Software vorher studieren und abwägen, ob der Nutzen die Weitergabe persönlicher Informationen wert ist.
d. Nicht mehr benötigte Apps, Plug-Ins oder Add-Ons löschen.

Dies gilt natürlich eher nicht für die Apps, die von einer Hochschule zur Verfügung gestellt werden, wohl aber für zusätzliche Software, die auf Endgeräten installiert werden, die im Netz einer Hochschule angeschlossen werden.

Die Überlegung hinter dem letzten Prinzip bezieht sich unter anderem auf Aktualisierungen von AGBs. Denn es kann vorkommen, dass dadurch Datenschutzverordnungen geändert werden und neue Zugriffsrechte auf Daten entstehen.[38] Wer die Anzahl installierter Apps möglichst klein hält, behält den Überblick und verkleinert die Angriffsfläche – für passiven Datenabfluss als auch für Cyberangriffe.[39]

Konkret gegen Phishing hingegen helfen insbesondere die folgenden Vorgehensweisen, die allesamt entweder „Wachsamkeit" oder „Verifizierung" beinhalten:

1. **Gesunde Skepsis bei E-Mails und deren Anhängen**
   Das E-Mail-Postfach ist eines der beliebtesten Einfallstore für Cyberkriminelle - und insbesondere für Phishing-Versuche. Zur Abwehr von letzterem benötigt es vor allem stetige Wachsamkeit bezüglich gewisser typischer Muster von Phishing-Mails. Darüber hinaus gibt es noch technische Aspekte zu berücksichtigen. Die wichtigsten Punkte sind die folgenden:[40]

   a. E-Mails von unbekannten Absendern, die (i) viele Schreib- und Grammatikfehler beinhalten, (ii) zum Anklicken von Links oder Scannen von QR-Codes auffordern sowie (iii) in irgendeiner Form Handlungsdruck erzeugen, sind skeptisch und vorsichtig zu

[38] S. BSI, „Smartphone, Tablet und Co sicher nutzen", Tipp 2.
[39] Hier überschneidet sich das Thema mit dem der *digitalen Souveränität* – das Hochschulforum Digitalisierung informiert hierzu auf folgender Website: https://hochschulforumdigitalisierung.de/dossier/digitale-souveraenitaet/.
[40] S. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Social-Engineering/social-engineering_node.html, https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Spam-Phishing-Co/Passwortdiebstahl-durch-Phishing/passwortdiebstahl-durch-phishing_node.html und https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-themen/social-engineering.html.

betrachten. Gerade letzteres ist eine beliebte Masche, um Menschen zu voreiligem Handeln zu bewegen.

b. Generell gilt es bei verdächtigen E-Mails keine Links zu öffnen. Das Besuchen der Webseite allein kann schon Schadsoftware verbreiten.
c. Dasselbe gilt bei Anhängen, vor allem welche, die mit .exe, .zip oder .scr enden. Hier handelt es sich häufig um ausführbare Programme, die schädlich sein können. Anhänge von verdächtigen E-Mails sollten im Allgemeinen nicht geöffnet werden.
d. QR-Codes von verdächtigen E-Mails sollten nicht gescannt werden. Dadurch kann ein Link zu einer Webseite geöffnet werden, welche wiederum automatisch Schadsoftware einschleust.
e. Teilweise werden Phishing-E-Mails im HTML-Format versendet. Dadurch kann bereits beim Öffnen der E-Mail Schadsoftware auf den Rechner gelangen. Entsprechend gilt es, diese Darstellungen in der Mail-Applikation abzuschalten, wodurch immer nur Text in E-Mails erscheint. Dies lässt sich in der Regel über die Einstellungen der verwendeten E-Mail-App einstellen.

2. **Niemals sensible Daten über E-Mail, Telefon oder sonst wie weitergeben**
Dies mag offensichtlich erscheinen, kann aber je nach Phishing-Versuch schwierig werden. Mit generativer KI beispielsweise kann eine Stimme eines Bekannten am Telefon nachgeahmt werden, oder es wird die Rolle einer Autoritätsperson (Polizei, Feuerwehr, etc.) vorgespielt, die mit drastischen Konsequenzen bei unkooperativem Verhalten droht.

   Grundsätzlich gilt aber: **seriöse Institutionen fragen nicht über E-Mail oder Telefon nach sensiblen (Zugangs-)Daten. Es ist wichtig, sich nicht unter Druck setzen zu lassen oder voreilig zu handeln.** Man soll möglichst einen kühlen Kopf bewahren und kritische Gegenfragen stellen (s. unten).

3. **Unabhängige Verifikation der Echtheit**
Natürlich kann es vorkommen, dass über Telefon, SMS oder E-Mail dringliche Fragen oder Handlungserwartungen gestellt werden. Hier gilt es, im Zweifelsfall die Kontaktaufnahme zu verifizieren. Je nachdem, über welches Medium die Nachricht eintraf, ergeben sich verschiedene Möglichkeiten:

   a. **Bei E-Mails und SMS:** Recherchen über den Absender anstellen.
      i. Überprüfung der E-Mail-Absender-Adresse: Manchmal entlarven sich Cyberkriminelle durch die Verwendung einer Absender-Adresse, die offensichtlich nicht zur vorgegebenen Identität passt. Es ist aber Vorsicht geboten: Absender-Adressen lassen sich leicht faken!
      ii. Bei E-Mails lassen sich je nachdem *unabhängig* weitere Kontaktmöglichkeiten, wie beispielsweise über Telefon,

ermitteln. Über diese kann direkt nachgefragt werden, ob die Nachricht echt war. Wichtig ist, dass man nicht die Kontaktdaten aus dem E-Mail verwendet, da diese ebenfalls falsch sein können. Bei Arbeitskolleg:innen, Kommilliton:innen, Behörden, Bekannten oder Banken dürfte dies besonders einfach sein, da man zu diesen ohnehin andere Kontaktmöglichkeiten besitzt.

iii. Auch Telefonnummern lassen sich häufig gut online recherchieren. Insbesondere die **Verbraucherzentrale** (https://www.verbraucherzentrale.de/, Deutschland), das **Bundesamt für Cybersicherheit** (https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-bedrohungen.html, Schweiz) und die **Internet-Watchlist** (https://www.watchlist-internet.at/, Österreich) informieren regelmäßig zu aktuellen Betrugsmaschen und verdächtigen Telefonnummern.

b. **Bei Telefonaten:** Gerade wegen dem Einsatz von Deepfakes (s. Abschn. 3) kann man je nachdem nur schwer anhand der Stimme erkennen, ob die Identität echt ist. Zudem können auch gewisse Details der zu imitierenden Person recherchiert werden (s. unten), was je nachdem das Imitieren von Sprechmustern oder Inhalten erlaubt. Dennoch: **Gezielte Detailfragen, deren Antworten nur von der entsprechenden Person gewusst werden könnten, bieten ein Schutzschild. Ebenso kann man beispielsweise mit der eigenen Bank Codephrasen abmachen**, um in Zukunft die Identität des Gesprächspartners zu verifizieren.[41]

Durch solche Verifikationen kann man nicht nur Phishing, sondern Social Engineering ganz allgemein begegnen. Es ist wichtig zu betonen, dass selbst wenn solche Recherchen auf den ersten Blick umständlich wirken, sie in Wirklichkeit oft nur einzelne Minuten in Anspruch nehmen. Darüber hinaus werden sie bei wiederholter Verwendung Routine, was entsprechend umso weniger Aufwand generiert. **Um es nochmals zu betonen: der Aufwand lohnt sich, da der Schaden an der eigenen Person und bei anderen enorm sein kann!**

4. **Sensible Nutzung von sozialen Medien**

Im Zusammenhang mit Social Engineering geht es beim sensiblen Umgang mit sozialen Netzwerken vor allem darum, **wie viel man von sich über was und auf**

[41] Vgl. https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Spam-Phishing-Co/Passwortdiebstahl-durch-Phishing/passwortdiebstahl-durch-phishing_node.html.

welche Weise preisgibt. Hierzu gibt es eine zentrale Frage, die man sich vor dem Veröffentlichen von Informationen auf sozialen Medien stellen sollte:[42]

> **Würde ich diese Information im echten Leben einer riesigen Menschenmenge mit unbekannter Zusammensetzung anvertrauen?**

Wahrscheinlich würde man dies für eine Reihe von Informationen eher verneinen: weder der eigene Beziehungsstatus, noch der aktuelle Aufenthaltsort in den Ferien, noch Details über die eigene Arbeit oder den Arbeitgeber würde man auf diese Weise unter den meisten Umständen preisgeben.

Tut man dies trotzdem, so können diese Informationen – da sie öffentlich zur Verfügung stehen – auch von Cyberkriminellen und Organisationen recherchiert und verknüpft werden (Stichwort „gläserner Mensch"). So ist es bereits vorgekommen, dass Einbrecher über soziale Medien das Leerstehen der Zielwohnung recherchiert haben, da die Mieter ihren Aufenthaltsort in Echtzeit auf sozialen Medien publiziert hatten. Gerade aber beim Phishing ist die Fülle der Informationen, die frei über ein potenzielles Opfer aufzufinden sind, entscheidend über die Versiertheit der Täuschung. Entsprechend hilft der sensible Umgang mit sozialen Medien *vor allem gegen Spear-Phishing*. Es gilt:

- → **vorsichtig abzuwägen, was man im Internet von sich preisgibt.**
- → **niemals Details über die Arbeit oder den Arbeitgeber zu veröffentlichen, ohne deren Zustimmung zu haben.**

Obschon diese Tipps und Methoden einen hohen Grad an Schutz bewirken können, gilt auch hier dasselbe wie in der Cybersicherheit allgemein: **einen einhundertprozentigen Schutz gibt es nicht**. Es kann also vorkommen, dass man trotz allen Bemühungen auf einen Phishing-Angriff reinfällt. Dann ist es wichtig, möglichst schnell die richtigen Schritte einzuleiten und nicht etwa aus Scham darauf zu verzichten, den Vorfall zu melden und Hilfe aufzusuchen. Dies führt uns aber über das Thema Desinformation sowie Wachsamkeit und Verifizierung hinaus, weswegen hier nur auf die entsprechende Checkliste des BSI verwiesen werden soll.[43]

Zum Schluss ist es wichtig zu betonen, dass gewisse der obigen Tipps und Methoden zwar dem momentanen Stand der Technik entsprechend formuliert wurden, **aber allgemeinerer Natur sind**. Wer weniger Informationen von sich ins Internet stellt, der erschwert den entsprechenden Akteuren - bei jeglicher technischer Ausstattung - die Möglichkeit, gezielte Recherchen in öffentlichen Medien durchzuführen. Ebenso ist das Stellen von gezielten Fragen zur Identifikationsverifikation unabhängig vom Grad der Deepfake-Technologie effektiv!

[42] Abgewandelt entnommen aus: DSN, „Internet, Social Media und Big Data", S. 6f.
[43] S. BSI, „Phishing: Checkliste für den Ernstfall".

8 Fazit

Die hier behandelten Themen hatten alle mit Fehl- oder Desinformation zu tun, in der einen oder anderen Form. Fake News verzerren wahre Inhalte oder benutzen gezielt Desinformation zu politischen Zwecken, während halluzinierte Quellen zwar ohne Täuschungsabsicht entstehen, aber bei unsorgfältiger Weiterverbreitung dennoch die Integrität des Wissenschaftssystems untergraben. Die beiden wirksamsten Werkzeuge gegen solche irreführenden Informationen hatten wir als „Wachsamkeit" und „Verifikation" bezeichnet. Ersteres bezeichnet die Fähigkeit beziehungsweise Haltung, Inhalten grundsätzlich mit gesunder Skepsis zu begegnen, während Letzteres in Verdachtsfällen die Kompetenz meint, gegebene Inhalte zu überprüfen. Beiden ist dabei der Einsatz des kritischen Denkens gemein: Sowohl die Gewohnheit, nach Kennzeichen von Fehl- und Desinformationen Ausschau zu halten, als auch diese als solche zu erkennen, benötigt logisch-argumentative Kompetenzen. Hier schließt sich der Kreis zur Hochschulbildung erneut. Denn die Vermittlung solcher Kompetenzen ist Kern eines jeden Hochschulstudiums und unterstreicht somit den Beitrag der Hochschulen zu einer gesunden und gelebten Demokratie.

Die Fülle der hier vorgestellten Informationen und Techniken mögen auf den ersten Blick überwältigend wirken. Wie wir aber mehrfach erwähnt haben, ist die Kultivierung eines pragmatischen Umgangs mit irreführenden Informationen zentral. Man kann und muss nicht jedes Bild, jeden Zeitungsartikel, jedes Paper und jede zitierte Quelle auf Echtheit überprüfen. Noch sollte man hinter jeder entlarvten Fehlinformation eine Täuschungsabsicht vermuten. Vielmehr gilt es, den Einsatz eines Werkzeugs je nach Kontext auf sein Nutzen-Kosten-Verhältnis hin zu überprüfen und dann entsprechend gezielt einzusetzen.

Eine Methode, um hierzu einen Erfahrungswert aufzubauen, ist es, sich mit folgenden zwei Aspekten zu beschäftigen:

Täuscherperspektive einnehmen: Zu was lohnt es sich überhaupt, zu täuschen? Wenn man eine Gruppe zum Thema X täuschen wollte, wie würde man es tun?

Selbstreflexion: Zu was könnte man mich besonders gut täuschen, weil ich beispielsweise stark emotional darauf reagiere? Welche persönlichen Schwächen würden dazu ausgenutzt werden?

Anstelle also jedes Mal mit den obigen Tools in die Tiefe zu tauchen, kann man auch pragmatisch die Thematik von der anderen Seite her aufrollen: Wo sollte man im Allgemeinen wachsam bleiben, weil dort ein Interesse an Täuschung am ehesten beziehungsweise überhaupt erst besteht? Und wo ist man persönlich besonders

verwundbar? Zu guter Letzt sei nochmals die Wichtigkeit von Routine erwähnt. Denn mit zunehmender Übung ist ein DOI-, Deepfake- oder Identitätscheck auch schneller und effizienter durchgeführt! In diesem Sinne möchten wir besonders zum Aufbau von Routine und der Entwicklung individueller Strategien und Vorgehensweisen raten.

Um zum Schluss nochmals auf unseren Fokus auf die individuelle Verantwortung – vor allem von Studierenden und Promovierenden – zurückzukommen, möchten wir insbesondere zur Kultivierung einer gesunden *Fehlerkultur* raten. Schließlich sind wir alle fehlbar: Es ist uns allen möglich, auf Fehl- und Desinformationen reinzufallen. In einer konstruktiven Fehlerkultur jedoch ist dies kein Anlass zur Bloßstellung, sondern für konstruktive Kritik, und wird als Lernchance wahrgenommen und gelebt. Dadurch stärken wir unsere Resilienz auf der individuellen Ebene, ohne dabei die Einzelperson „im Stich" zu lassen. So können wir der Gefahr von Fehl- und Desinformation auch in Zukunft umso gestärkter und souveräner entgegentreten.

9 Literaturverzeichnis

Publikationen

Aljamaan et al. (2024): „Reference Hallucination Score for Medical Artificial Intelligence Chatbots: Development and Usability Study“, in: *JMIR Medical Informatics* 12, S. 1-11.

Antos, G. & Ballod, M. (2019): „Web und Wahrheit. Vorbemerkungen zu einer Didaktik informationeller Verlässlichkeit“, in: Beißwenger & Knopp (Hrsg.), *Soziale Medien in Schule und Hochschule: Linguistische, sprach- und mediendidaktische Perspektiven*, Berlin: Peter Lang, S. 23-58.Bundesamt für Sicherheit in der Informationstechnik (2021): „Das Internet der Dinge sicher nutzen“, URL = https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Wegweiser_Checklisten_Flyer/Brosch_A6_Internet_der_Dinge.pdf?__blob=publicationFile&v=9.

Bundesamt für Sicherheit in der Informationstechnik (2021): „Smartphone, Tablet und Co sicher nutzen“, URL = https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Wegweiser_Checklisten_Flyer/Brosch_A6_Mobilkommunikation.pdf?__blob=publicationFile&v=18.

Bundesamt für Sicherheit in der Informationstechnik (2021): „Soziale Netzwerke sicher nutzen“ URL = https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Wegweiser_Checklisten_Flyer/Brosch_A6_Soziale_Netzwerke.pdf?__blob=publicationFile&v=13.

Bundesamt für Sicherheit in der Informationstechnik (2024): „Künstliche Intelligenz sicher nutzen“, URL = https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Wegweiser_Checklisten_Flyer/Brosch_A6_Kuenstliche_Intelligenz.pdf?__blob=publicationFile&v=18.

Bundesamt für Sicherheit in der Informationstechnik (2025): „Basisschutz für sichere E-Mail-Nutzung“, URL = https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Wegweiser_Checklisten_Flyer/Wegweiser_kompakt_E-Mail.pdf?__blob=publicationFile&v=7.

Bundesamt für Sicherheit in der Informationstechnik (2025): „Phishing: Checkliste für den Ernstfall“, URL = https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Broschueren/Wegweiser_Checklisten_Flyer/Checkliste_BSI_ProPK_Phishing.pdf?__blob=publicationFile&v=5.

Bundesamt für Verfassungsschutz (2025): „BfV CYBER INSIGHT: Im Schatten des Cyberspace - Teil 1: Cyberangriffe – Wer? Wie? Was?“, URL = https://www.verfassungsschutz.de/SharedDocs/publikationen/DE/cyberabwehr/2025-09-16-bfv-cyber-insight-cyberspace-teil-1.pdf?__blob=publicationFile&v=6. Conly, S (2016): *One Child: Do We Have a Right to More?* New York: Oxford University Press.

Deutsche Forschungsgemeinschaft (2026): „Positionspapier zur Stärkung der Freiheit und Resilienz der Wissenschaften: Orientierungen und Empfehlungen“, URL = https://zenodo.org/records/21353471.

Direktion Staatsschutz und Nachrichtendienst (2022): „Internet, Social Media und Big Data“, URL = https://www.dsn.gv.at/501/files/cyber_ratgeber/schriftenreihe_cybersicherheit_internetsocial_media_und_big_data_februar_2022_bf_20220216.pdf.

Direktion Staatsschutz und Nachrichtendienst (2025): „Verfassungsschutzbericht 2024“, URL = https://www.dsn.gv.at/501/files/vsb/205_2025_vsb_2024_v20250929_webbf.pdf.

Earl J., Rieder T. N. & Hickey C. (2016): „Population Engineering and the Fight against Climate Change”, in: *Social Theory and Practice* 42 (4), S. 845-870.

Hedberg, T. (2020): *The Environmental Impact of Procreation: The Ethics of Procreation*. London: Routledge.

Hochschulrektorenkonferenz (2025): „Empfehlung der 40. Mitgleiderversammlung: Handlungsdruck für Hochschulen, Länder und Bund – HRK-Empfehlungen zur Cybersicherheit“, URL = https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-01-Beschluesse/2025-05-13_HRK-MV_Empfehlung_Cybersicherheit.pdf.

Keefe, T. E. (2019): „Not All 'Fake News' Is Equal: How Should Higher Education Respond to Fake News and Disinformation in the post-Truth Era“, in: *The Liminal: Interdisciplinary Journal of Technology in Education* 1 (1), S. 1-9.

Murtaugh, P. & Schlax, M. (2009): „Reproduction and the carbon legacies of individuals”, in: *Global Environmental Change* 19 (1), S. 14–20.

Nachrichtendienst des Bundes (2026): „Sicherheit Schweiz – Lagebericht des Nachrichtendienstes des Bundes 2025”, URL = https://cms.news.admin.ch/fileservice/sdweb-docs-prod-nsbcch-files/files/2025/07/01/558fc40a-78ee-4d64-b036-47dd7a42ed14.pdf.

Pinkert, F. & Sticker, M. (2025): „Procreative Prerogatives and Climate Change”, in: *Journal of Applied Philosophy* 42 (1), S. 44-66.

Schroll, A.-L., Rademacher, M. & Metzner, J. (2019): „Zwischen `Fake News´ und `Future Skills´“, Diskussionspapier Nr. 5. Berlin: Hochschulforum Digitalisierung.

Topaz et al. (2026): „Fabricated citations: an audit across 2·5 million biomedical papers", in: *The Lancet* 407, S. 1779-1781.

Weatherall, J. O. (2024): „Fake News!", in: *Philosophy Compass* 19 (6), S. 1-12.

Wissenschaft im Dialog/Verian (2025): „Wissenschaftsbarometer 2025", URL = https://wissenschaft-im-dialog.de/documents/473/Brosch%C3%BCre_Wissenschaftsbarometer2025.pdf.

Yan, P., & Schroeder, R. (2025): „Living in a post-truth era? Online misinformation in everyday life in rural and urban China", in: *Information, Communication & Society* 28 (7), S. 1195–1215.

Zhuang, H., Liang, L. & Acuña, D. E. (2025): „Estimating the predictability of questionable open-access journals", in: *Science Advances* 11 (35), S. 1-12.

Verwendete Internetquellen/-ressourcen

1) https://digital-strategy.ec.europa.eu/en/policies/media-literacy
2) https://quoteinvestigator.com/2016/12/30/not-malice/
3) https://blogs.lse.ac.uk/medialse/2026/05/22/why-we-are-not-in-a-post-truth-era/
4) https://www.stiftungen.org/aktuelles/news-aus-stiftungen/detail/studie-zu-desinformation-von-jungen-menschen-in-der-coronakrise-4986.html
5) https://www.bib.uni-mannheim.de/services/kurse-und-tutorials/fake-science-in-der-wissenschaft/
6) https://www.bitkom.org/Presse/Presseinformation/9-von-10-Deutschen-stossen-auf-Fake-News
7) https://arbeit-der-zukunft.de/artikel/hochschulen-unter-druck-warum-eine-demokratische-und-gerechte-gesellschaft-resiliente-hochschulen-braucht-labora-2025
8) https://correctiv.org/faktencheck/
9) https://de.bellingcat.com/
10) https://www.forschung-it-sicherheit-kommunikationssysteme.de/forschung/it-sicherheit/desinformation
11) https://www.insurancetimes.co.uk/news/fraud-attempts-with-deepfakes-surge-over-last-three-years/1454509.article
12) https://arbeit-der-zukunft.de/artikel/hochschulen-unter-druck-warum-eine-demokratische-und-gerechte-gesellschaft-resiliente-hochschulen-braucht-labora-2025
13) https://hochschulforumdigitalisierung.de/die-hausarbeit-ist-tot-es-lebe-die-hausarbeit-entwicklungsorientierung-wissenschaftliches-arbeiten-und-ki-gemeinsam-denken/
14) https://www.veraai.eu/posts/verification-plugin
15) https://www.iu.de/forschung/studien/medienkompetenz/fake-news-erkennen/
16) https://www.doi.org/
17) https://scholar.google.com/
18) https://www.semanticscholar.org/

19) https://www.base-search.net/
20) https://pubmed.ncbi.nlm.nih.gov/
21) https://philpapers.org/
22) https://projecteuclid.org/
23) https://dbis.uni-regensburg.de/
24) https://orcid.org/
25) https://www.linkedin.com/posts/andreas-giesbert-773968262_chatgpt-ki-taeuuschung-activity-7454449729414979584-ALgX?utm_source=share&utm_medium=member_desktop&rcm=ACoAADZwOWABJ6bLhyOfJqfFqILdAvXIV4erk4I
26) http://web.archive.org/
27) https://arxiv.org/
28) https://www.biorxiv.org/
29) https://ombudsgremium.de/13305/handreichung-zu-predatory-journals/
30) https://doaj.org/
31) https://thinkchecksubmit.org/
32) https://retractionwatch.com/retraction-watch-database-user-guide/retraction-watch-database-user-guide-appendix-b-reasons/
33) https://retractionwatch.com/
34) https://retractiondatabase.org/
35) https://www.crossref.org/
36) https://pubpeer.com/
37) https://www.deutschlandfunk.de/warum-cyberkriminelle-oft-hochschulen-angreifen-interview-mit-maren-luebcke-100.html
38) https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Social-Engineering/social-engineering_node.html
39) https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Cyber-Sicherheitslage/Methoden-der-Cyber-Kriminalitaet/Spam-Phishing-Co/Passwortdiebstahl-durch-Phishing/passwortdiebstahl-durch-phishing_node.html
40) https://www.bundeskriminalamt.at/212/internet/phishing_smishing.html
41) https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-bedrohungen/phishing.html
42) https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-unternehmen/aktuelle-bedrohungen/ransomware.html
43) https://www.zeit.de/news/2026-07/10/jugendlicher-hacker-soll-uni-des-saarlandes-erpresst-haben
44) https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/sichere-passwoerter-erstellen_node.html
45) https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-themen/schuetzen-sie-ihre-konten.html
46) https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-themen/geraeteschutz.html
47) https://hochschulforumdigitalisierung.de/dossier/digitale-souveraenitaet/
48) https://www.ncsc.admin.ch/ncsc/de/home/infos-fuer/infos-private/aktuelle-themen/social-engineering.html

10 Abbildungsverzeichnis

11 Tabellenverzeichnis

12 Impressum

ISSN (Online) 2365-7081; 12. Jahrgang

Zitierhinweis

Vonlanthen, S. D., Koitka-Fieke, C. (2026). Wachsamkeit & Verifizierung. Digitale Resilienz in Hochschulen als Faktor für gelebte Demokratie. Arbeitspapier Nr. 94. Berlin: Hochschulforum Digitalisierung.

Herausgeber

Geschäftsstelle Hochschulforum Digitalisierung beim Stifterverband für die Deutsche Wissenschaft e.V.

Hauptstadtbüro • Pariser Platz 6 • 10117 Berlin • T 030 322982-520

info@hochschulforumdigitalisierung.de

Redaktion

Christoph Koitka-Fieke, Simon D. Vonlanthen, Michael Siegel

Layout

Satz: Christoph Koitka-Fieke, Simon D. Vonlanthen, Michael Siegel

Vorlage: TAU GmbH • Köpenicker Straße 154 A • 10997 Berlin

Das 2014 gegründete Hochschulforum Digitalisierung ist eine gemeinsame Initiative des Stifterverbandes, des CHE Centrums für Hochschulentwicklung und der Hochschulrektorenkonferenz. Förderer ist das Bundesministerium für Forschung, Technologie und Raumfahrt.

www.hochschulforumdigitalisierung.de