

Digitale Gesellschaft(en)

Neue Forschungsansätze zur Digitalen Transformation



Josephine B. Schmitt, Julian Kohne & Johannes Breuer (Hrsg.)

Digitale Gesellschaft(en)

Neue Forschungsansätze zur Digitalen Transformation

Inhalt

Editorial

3

Josephine B. Schmitt, Julian Kohne & Johannes Breuer

Einführung: Digitalisierungsforschung. Wie wir die digitale Transformation wissenschaftlich erfassen können

4

Samuel T. Simon & Josephine B. Schmitt

Agile Science: Co-Creating Research on Digital Transformation

12

Leon Fröhling, Lukas Birkenmaier & Jessica Daikeler

Garbage in – Garbage out? Datenqualität im Umgang mit digitalen Verhaltensdaten

21

Julia Weiß & Sebastian Stier

Die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten

31

Matthias Begenat

**Wissenschaftskommunikation in unsicheren, digitalen Zeiten.
Eine Interviewstudie mit Kommunikator*innen aus der Digitalisierungsforschung**

39



Editorial

Liebe Leser*innen,

die Digitalisierung schreitet unaufhaltsam voran und wirkt auf unsere Gesellschaft in vielfältiger Weise und einem Ausmaß, dass eine umfassende Reflektion über Gesellschaft ohne die Betrachtung sozialer Handlungen im digitalen Raum unvollständig bleiben müsste. Dieser Herausforderung wollen wir in zwei *easy*-Schwerpunktheften begegnen, die neue Forschungsansätze zur Digitalen Transformation präsentieren, und Ihnen verschiedene Perspektiven auf dieses spannende Thema bieten.

Im vorliegenden Heft beleuchten wir verschiedene Facetten der Digitalisierung und zeigen, wie sie unsere Gesellschaften und die Art und Weise, wie wir miteinander interagieren, beeinflusst. Dafür haben wir eine Gruppe von Wissenschaftler*innen, insbesondere von GESIS – Leibniz-Institut für Sozialwissenschaften und vom CAIS – Institut für Digitalisierungsforschung eingeladen, ihre Forschungsperspektiven zu teilen und uns einen Einblick in die konkreten *Arbeitsweisen* der Digitalisierungsforschung zu geben. Im zweiten Heft, dass in der zweiten Jahreshälfte

erscheinen wird, werden wir diese Perspektiven dann an unterschiedlichen Anwendungsbeispielen noch eingehender vorstellen.

Gleichzeitig nähern wir uns in beiden Schwerpunktheften einem Thema, das innerhalb der Sozialwissenschaften immer mehr an Bedeutung gewinnt, den „Digitalen Verhaltensdaten“. Denn das Sammeln, Analysieren und Interpretieren von Daten, die von digitalen Geräten und Plattformen stammen, wird zu einem zunehmend wichtigen Instrument, um Verhaltensmuster und soziale Phänomene zu verstehen. Wir sind überzeugt, dass die Erforschung von digitalen Verhaltensdaten in Zukunft eine bedeutende Rolle spielen wird und uns Schritt für Schritt ermöglicht, digitale Gesellschaften besser zu verstehen.

Lassen Sie sich von den vielen neuen Gedanken und Zugangsweisen dieses dynamischen Forschungsfeldes inspirieren!

Herzliche Grüße von Ihren *easy*-Herausgeber*innen

Johannes Breuer, Philip Janßen, Lydia Repke & Sophie Zervos



Digitalisierungsforschung

Wie wir die digitale Transformation wissenschaftlich erfassen können

Josephine B. Schmitt, Julian Kohne & Johannes Breuer

Die Digitalisierungsforschung ist ein junges und sehr dynamisches Feld. Sie umfasst die Erforschung der Bedingungen und Auswirkungen der sozialen, politischen, wirtschaftlichen, kulturellen und technischen Veränderungen, die mit der Digitalisierung einhergehen. Neben zahlreichen neuen Forschungsfragen, die es zu beantworten gilt, fordert die Digitalisierung die Organisation von Forschung und den Umgang mit Daten heraus. Das ist sowohl Chance als auch Herausforderung für die Forschung. Die vorliegende Ausgabe von „easy_social_sciences“ möchte sich in vier Beiträgen mit verschiedenen Aspekten der Digitalisierungsforschung befassen. Dabei werden unterschiedliche Perspektiven von der Konzeption relevanter Forschungsfragen über die Erhebung und Verarbeitung neuer Datentypen bis zur Kommunikation wissenschaftlicher Ergebnisse beleuchtet.

Research on digital transformation is a young and very dynamic field. It encompasses research regarding the conditions and effects of the social, political, economic, cultural, and technical changes that accompany the digital transformation. In addition to numerous new research questions that need to be answered, digital transformation challenges the organization of research and the handling of data. This is both an opportunity and a challenge for research. This issue of “easy_social_sciences” aims to take a closer look at different aspects of research on digital transformation ranging from the conceptualization of relevant research questions to collecting and processing data and communicating results to the general public. In doing so, different perspectives on the challenges for working with digital data, as well as questions dealing with the dialogue between science and society in the context of digitization research are highlighted.

Keywords: Digitalisierung, digitale Transformation, Interdisziplinarität, digitale Daten

Die wissenschaftliche Forschung befasst sich mit der digitalen Transformation und ist zugleich selbst von ihr betroffen. Neben zahlreichen neuen Forschungsfragen, die es zu beantworten gilt, fordert die Digitalisierung die Organisation von Forschung und den Umgang mit Daten heraus. Das ist sowohl eine Chance als auch eine Herausforderung. So erfordert die Erforschung der digitalen Transformation die Perspektiven und das Zusammenspiel verschiedener wissenschaft-

licher Disziplinen. Wenn es beispielsweise um die Analyse des Einsatzes von künstlicher Intelligenz (KI) im Schulunterricht geht, ermöglicht ein interdisziplinäres Forschungsteam eine ganzheitliche Betrachtung des Gegenstandes: Psycholog*innen bringen etwa wertvolles Wissen über menschliches Wahrnehmen und Lernen, aber auch über die Wirkung von Bildungstechnologien für die Zusammenarbeit von Schüler*innen ein. Erziehungswissenschaftler*innen unterstüt-

zen mit Kenntnissen hinsichtlich der Frage, wie Lehren und Lernen mit KI im Schulunterricht ganz praktisch gestaltet werden kann (Didaktik). Forschende aus den Computerwissenschaften ergänzen dies durch Wissen über die technischen Bedingungen und Gestaltungsmöglichkeiten von KI-gestützten Bildungstechnologien. Jurist*innen klären u.a. Fragen des Datenschutzes und der rechtlichen Möglichkeiten des Einsatzes von KI in der Schule.

» **Verschiedene Begriffe, andere Phänomene: Digitization – Digitalization – Digital Transformation.** «

Dieser Themenbereich ist nur ein Beispiel dafür, warum Forschungsfragen im Kontext der Digitalisierung inzwischen vermehrt in disziplinenübergreifenden Forschungsverbänden oder spezialisierten interdisziplinären Forschungseinrichtungen bearbeitet werden. In den letzten Jahren lässt sich die Entwicklung eines eigenen wissenschaftlichen Feldes beobachten: die *Digitalisierungsforschung*.

Was ist Digitalisierungsforschung?

Der Begriff „Digitalisierungsforschung“ wird bisher vorwiegend im deutschsprachigen Raum verwendet. International hat die englisch geprägte Diskussion um die Forschung zu digitaler Transformation verschiedene Begriffe hervorgebracht, welche auf unterschiedliche Phänomene in diesem Zusammenhang Bezug nehmen. Während sich der englische Begriff „*digitization*“ auf den Prozess der Umwandlung von Informationen in digitale Formate bezieht, beschreibt „*digitalization*“ komplexere Auswirkungen digitaler Kommunikations- und Medienstrukturen auf gesellschaftliche Prozesse und Strukturen

(Brennen & Kreiss, 2016). Der Ausdruck „*digital transformation*“ hingegen kann als Prozess der Umstellung einer Organisation auf neue Arbeits- und Denkweisen durch den Einsatz digitaler, sozialer, mobiler und neuer Technologien definiert werden (Gorenšek & Kohont, 2019). Allgemeingültige fächerübergreifende Definitionen für „*digitalization*“ und „*digital transformation*“ gibt es bisher jedoch nicht (Peng et al., 2013; Reis et al., 2018).

In diesem Editorial sowie in den Beiträgen dieser Ausgabe von *easy* liegt der Fokus auf Prozessen und Veränderungen, die sich am ehesten dem Bereich der „digitalen Transformation“ im Sinne der vorausgegangenen Definition zuordnen lassen. Während der Begriff im Englischen vorwiegend im Unternehmenskontext verwendet wird (z.B. Verhoef et al., 2021), weiten wir ihn hier auf die gesamte Gesellschaft aus. Der Begriff „Digitaler Wandel“ ist für uns synonym dazu zu verstehen.

Das Feld der Digitalisierungsforschung

Die Digitalisierungsforschung ist ein junges und sehr dynamisches Feld. Sie umfasst die Erforschung der Bedingungen und Auswirkungen der sozialen, politischen, wirtschaftlichen, kulturellen und technischen Veränderungen, die mit der digitalen Transformation einhergehen. Aufgrund der Vielfalt und Komplexität ihres Forschungsgegenstandes verknüpft die gegenwärtige Digitalisierungsforschung verschiedene disziplinäre Perspektiven. Dazu gehören u.a. Sozial- und Geisteswissenschaften, Psychologie, Medizin, Rechts-, Wirtschafts-, Computer- und Technikwissenschaften. Bislang ist das Feld jedoch, zumindest in Deutschland, stark sozialwissenschaftlich geprägt.

Die Entstehung neuer wissenschaftlicher Felder ist im Rahmen der Hochschulbildung historisch betrachtet nichts Besonderes. Blickt man in die Wissenschaftsgeschichte, finden

sich zahlreiche Beispiele für neue bzw. sich entwickelnde wissenschaftliche Felder. Ein anderes neueres Beispiel ist die Entwicklung der Geschlechterforschung (Gender Studies). Die Feldentwicklung kann dabei – analog zu (sozialen) Bewegungen, welche gesellschaftlichen Wandel herbeiführen oder aufhalten wollen – als *soziale und intellektuelle* Bewegung verstanden werden, welche durch kritische Ereignisse, wie z.B. technologische Innovationen oder gesellschaftliche Umwälzungen ausgelöst wird (Jacobs & Frickel, 2009).

» **Man geht nicht mehr aktiv und bewusst „ins Internet“, die digitale Technologie ist immer dabei.** «

Je nachdem, welche Stichworte man zugrunde legt, liegt die Geburtsstunde der Digitalisierungsforschung irgendwo zwischen den frühen 1990ern und 2010er Jahren – also parallel zum Beginn der kommerziellen Phase des Internets. Während die Anzahl an wissenschaftlichen Studien zu digitalisierungsbezogenen Themen bis in die 1990er Jahre noch sehr übersichtlich war, ist ab der ersten Dekade des 21. Jahrhunderts eine deutliche Zunahme an Forschungsarbeiten zu verzeichnen. Peng und Kolleg*innen (2013) finden im Rahmen einer Analyse von wissenschaftlichen Publikationen, die zwischen 2000 und 2009 erschienen sind, seit 2005 einen signifikanten Anstieg wissenschaftlicher Studien, welche „Internet“-Bezug aufweisen (z.B. zu *Internet, online, web, cyberspace*). Zwischen 2002 und 2005 wurden mit der Gründung von sozialen Medien wie Myspace, YouTube und Facebook wichtige Grundsteine für das Web 2.0 gelegt. Diese neuen Kommunikations- und Interaktionsmöglichkeiten haben zu einem massiven Zuwachs an Forschungsfragen und wissenschaftlichen Studien geführt. Seitdem differenziert sich das Feld der Digitalisierungsforschung immer mehr aus. Das führte in der internationalen Fachdiskussion in den Folgejahren auch zu einer Verschiebung

in der Verwendung von Begrifflichkeiten. Statt „Internet“ (und seiner Synonyme) wird vermehrt über „*digitization*“, „*digitalization*“ und „*digital transformation*“ gesprochen und geschrieben. Die Menge an Veröffentlichungen, welche Stichworte wie „digital“ oder verwandte Begriffe (z.B. *digitalization, digitization*) enthalten, wächst ab dem Jahr 2013 bzw. 2014 um ein Vielfaches (Holand et al., 2019; Mede, 2021; Parida et al., 2019; Reis et al., 2018). Das ist vermutlich u.a. auf den rasanten Anstieg an mobiler Internetnutzung mit Smartphone, Tablet und Co. zurückzuführen. In Deutschland beispielsweise verdoppelte sich im Jahr 2013 die Zeit, die Menschen unterwegs im Netz verbringen (van Eimeren & Frees, 2013). Damit rückt das Thema eines digitalen, kommunikativen Alltags mit permanenter Vernetzung in den Fokus der wissenschaftlichen Forschung (siehe z.B. Vorderer, 2015); man geht nicht mehr aktiv und bewusst „ins Internet“, die digitale Technologie ist immer dabei.

Auch die Breite des Feldes im Hinblick auf die Themen und die an der Erforschung beteiligten Disziplinen zeigt sich schon in den Anfängen der Digitalisierungsforschung. Peng und Kolleg*innen (2013) identifizieren für die Jahre 2000 bis 2009 beispielsweise „E-Health“, „E-Business“, „E-Society“ und „Mensch-Computer-Interaktion“ als besonders prominente Themen. Der Fokus auf technologische, wirtschafts- und organisationsbezogene Fragestellungen sowie interpersonelle Kommunikation (z.B. über soziale Medien) finden sich auch in anderen (späteren) Überblicksarbeiten zur Digitalisierungsforschung (z.B. Holand et al., 2019; Parida et al., 2019).

Die Erforschung von Themen mit Digitalisierungsbezug ist in den Anfängen des Feldes nicht auf bestimmte Disziplinen limitiert. Sie geschieht teilweise bereits früh über disziplinäre Grenzen hinweg (siehe z.B. Holand et al., 2019; Parida et al., 2019; Peng et al., 2013). Methodisch ist das Feld ebenfalls schon immer divers, wobei bisher eher quantitative Zugänge dominieren (siehe z.B. Holand et al., 2019; Mede, 2021; Peng et al., 2013).

Trotz vielfältiger Bestrebungen, Digita-

lisierungsforschung in dafür spezialisierten interdisziplinären Institutionen zu etablieren und zu verstetigen (in Deutschland z.B. Center for Advanced Internet Studies, Weizenbaum Institut, Bayerisches Forschungsinstitut für Digitale Transformation) ist die weitere Entwicklung des Forschungsfeldes noch recht offen. Entsprechend stellen sich grundlegende Fragen, etwa nach dem Ausmaß und der Ausgestaltung der Zusammenarbeit von Forschenden über Fächergrenzen hinweg oder einer interdisziplinären Perspektive auf die für die Digitalisierungsforschung notwendigen Daten, Theorien und Methoden. Daher lohnt es sich, Digitalisierungsforschung aus einer übergeordneten Perspektive zu betrachten.

Digitalisierungsforschung aus der übergeordneten Perspektive oder „Was passiert eigentlich in dieser Ausgabe?“

Hinsichtlich der Forschungsthemen ist das Feld überaus heterogen und dynamisch. Es in seiner Breite darzustellen, würde weit über den Umfang einer Zeitschriftenausgabe hinausgehen. Um den *easy*-Lesenden einen ersten Einstieg in das Feld zu ermöglichen, stellen wir in dieser Ausgabe vier Querschnittsthemen der Digitalisierungsforschung vor, die für alle Forschungsprojekte im Bereich Digitalisierungsforschung auf die eine oder andere Weise eine Rolle spielen. Entlang üblicher Schritte des Forschungsprozesses – Auswahl der Forschungsfragen, Studiendesign, Datenerhebung und -verarbeitung, und bis hin zur Wissenschaftskommunikation – möchten wir Erkenntnisse und Praktiken der Digitalisierungsforschung aufzeigen und diese für ein interdisziplinäres, aber auch nichtwissenschaftliches Publikum verständlich machen.

» **Man erwartet, dass Digitalisierungsforschung auf die (digitalen) gesellschaftlichen Umbrüche reagiert.** «

Auswahl von Forschungsfragen und Methoden der wissenschaftlichen Zusammenarbeit

Die digitalen Möglichkeiten verändern nicht nur die Forschung selbst, sondern auch den Austausch zwischen Wissenschaft und Öffentlichkeit. So können interessierte Personen digital nicht nur leichter auf wissenschaftliche Erkenntnisse zugreifen, sie können selbst – beispielsweise durch sogenannte Datenspenden, Citizen Science (Bürger*innenwissenschaft) oder digitale Partizipationsformate wie z.B. des im Rahmen des Wissenschaftsjahrs 2022 initiierten „Ideenlauf #MEINEFRAGEFÜRDIEWISSENSCHAFT“ – wissenschaftliche Forschung aktiv mitgestalten. Angesichts der Gefahr einer Vertrauenskrise in die Wissenschaft (Rutjens et al., 2018; Sills & Pittinsky, 2015), wie sie während der Corona-Pandemie bei Teilen der Bevölkerung erkennbar wurde, nimmt die Bedeutung einer gesellschaftlich kommunikationsfähigen Forschung außerhalb des sprichwörtlichen Elfenbeinturms noch weiter zu. Eine Orientierung an den Bedürfnissen gesellschaftlicher Stakeholder wie z.B. Bürger*innen, Politik und Wirtschaft, die für Akzeptanz, Relevanz und Impact der Forschung sorgen kann, bleibt aber in vielen Forschungskontexten noch ungenutzt. Diese Orientierung fängt bereits bei der Entwicklung von Forschungsfragen an. Insbesondere wird von der Digitalisierungsforschung erwartet, dass sie angemessen auf die mit der Digitalisierung einhergehenden gesellschaftlichen Umbrüche reagieren kann. Wie können aber

innovative Forschungsfragen in Reaktion auf aktuelle gesellschaftliche Herausforderungen identifiziert werden? Inwiefern können Bürger*innen und Expert*innen aus der Praxis an der Entwicklung von Fragestellungen für die Digitalisierungsforschung und damit am Forschungsprozess selbst beteiligt werden? Diese Fragen greift der Beitrag „Agile Science: Co-Creating Digital Transformation Research“ von *Samuel Simon* und *Josephine B. Schmitt* in diesem Heft auf. Sie schlagen dazu ein Konzept für eine „agile Wissenschaft“ vor, um die Zukunft einer innovativen und responsiven Digitalisierungsforschung maßgeblich mitzugestalten. Am Beispiel des mehrstufigen Themenfindungsprozesses des Center for Advanced Internet Studies (CAIS) zeichnen sie nach, wie die Identifikation von Forschungsthemen mit den unterschiedlichen gesellschaftlichen Akteur*innen basierend auf partizipativen und co-kreativen Ansätzen entlang des Innovationsframeworks „Double Diamond“ (British Design Council, 2005) gelingen kann. Außerdem verdeutlichen sie, wie mit agilen Elementen interdisziplinäre Zusammenarbeit in der Digitalisierungsforschung (neu-)gestaltet werden kann.

Datenerhebung und -verarbeitung

Neben der hervorgehobenen Bedeutung interdisziplinärer Zusammenarbeit in der Digitalisierungsforschung ist eine bedeutsame Veränderung, dass infolge der Digitalisierung zahlreiche neue Daten (v.a. sogenannte digitale Verhaltensdaten) für die Forschung zur Verfügung stehen. Mit diesen neuen Daten lassen sich menschliches Verhalten und Interaktionen in digitalen Räumen und mit digitalen Geräten (z.B. Diskussionen in sozialen Medien, Nachrichtennutzung, Konsumverhalten und Kommunikation), detaillierter nachvollziehen und analysieren. Allerdings sind diese Daten nicht immer für die Wissenschaft

zugänglich, etwa weil der Zugang zu den Daten von Privatunternehmen kontrolliert wird oder weil fortgeschrittene Programmierkenntnisse und ggf. auch technische Infrastrukturen für deren Erhebung und Verarbeitung notwendig sind. Zudem können derartige Daten aus rechtlichen und ethischen Gründen häufig nur mit Einschränkungen genutzt werden. Weiterhin steht zuweilen die Qualität der digital verfügbaren Daten infrage, da es sich oft um Daten handelt, die nicht explizit für die Forschung geschaffen wurden (wie etwa Umfragedaten), sondern quasi als Nebenprodukt bei der täglichen Nutzung von digitalen Geräten, Programmen oder Apps anfallen.

In ihrem Beitrag „Garbage in – Garbage out? Eine kritische Reflektion zur Qualität digitaler Verhaltensdaten“ nehmen *Leon Fröhling*, *Lukas Birkenmaier* und *Jessica Daikeler* die Qualität von digitalen Verhaltensdaten, also von Spuren menschlichen Verhaltens, die durch die Nutzung digitaler Technologien entstehen, unter die Lupe. Sie zeigen auf, wie sich digitale Verhaltensdaten von anderen Datentypen unterscheiden, und diskutieren, welche Rolle die Datenqualität für die daraus abgeleiteten Forschungsergebnisse spielt. Darauf aufbauend schlagen sie die Nutzung von sogenannten Error Frameworks als Rahmen für die systematische Analyse und Minimierung potenzieller Fehlerquellen vor, die mit digitalen Verhaltensdaten einhergehen können.

Neben den genannten Herausforderungen der neuen Datenquellen profitiert die Digitalisierungsforschung von den erweiterten Möglichkeiten zur Verknüpfung von Daten. Potenziale für die Forschung ergeben sich v.a. durch die Kombination unterschiedlicher Datentypen wie z.B. digitaler Verhaltensdaten und „klassischer“ Umfragedaten. *Julia Weiß* und *Sebastian Stier* beschäftigen sich in ihrem Artikel „Die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten“ eingehender mit diesem Thema. Sie diskutieren darin nicht nur die Vorteile und Wege der Verknüpfung dieser Datenformen, sondern legen auch ausführlich die Herausforderungen dar, die damit

einhergehen. Hier wird schnell klar, dass weder digitale Verhaltensdaten noch Umfragedaten für sich genommen die „besseren Daten“ sind, sondern dass sie unterschiedliche Vor- und Nachteile mitbringen und sich sehr gut wechselseitig ergänzen.

Wissenschaftskommunikation

Ausdruck einer gesellschaftsnahen und reaktionsfähigen Digitalisierungsforschung ist neben der Entwicklung von Forschungsfragen mit gesellschaftlichen Stakeholdern eine frühzeitige, praxisnahe und allgemein verständliche Kommunikation wissenschaftlicher Erkenntnisse. Diese befördert nicht nur die Akzeptanz wissenschaftlicher Arbeiten, sie ermöglicht es auch, mögliche Probleme im Hinblick auf die gesellschaftliche Bedeutung von Forschungsgegenständen rasch zu erkennen und gemeinsam Lösungen zu finden. Vor diesem Hintergrund befasst sich *Matthias Begegnat* in seinem Beitrag „Wissenschaftskommunikation in unsicheren, digitalen Zeiten. Eine Interviewstudie mit Kommunikator*innen aus der Digitalisierungsforschung“ mit Wissenschaftskommunikation in der Digitalisierungsforschung. Der Beitrag gibt Einblick in die Strategien der Wissenschaftskommunikation von Instituten der Digitalisierungsforschung: Der digitale Wandel eröffnet neue Möglichkeiten im Hinblick auf Partizipation und den Transfer von Erkenntnissen in die Gesellschaft – etwa durch die Nutzung digitaler Kanäle und Formate wie Social Media, Blogs oder Podcasts.

Zudem zeigen die Befunde, dass die Professionalisierung der Wissenschaftskommunikation zwar immer weiter voranschreitet, diese bisher aber im Wesentlichen als ‚Learning by Doing‘, also beim Ausführen der Tätigkeit selbst geschieht. Der Beitrag verdeutlicht, dass Kommunikator*innen den wechselseitigen Dialog von Wissenschaft und Gesellschaft zu Digitalisierungsthemen als Gelingensbe-

dingung für Wissenschaftskommunikation betrachten, dabei aber auf Vorbehalte in ihren Organisationen stoßen: Bisher gebe es nur wenige Anreize für Forschende in diesen Dialog umfassend zu investieren.

Ausblick

Wir hoffen mit diesem Special Issue sowohl Wissenschaftler*innen aus anderen Fachgebieten als auch interessierten Laien einen guten Überblick über die Arbeitsweisen der Digitalisierungsforschung in verschiedenen Abschnitten des Forschungsprozesses zu geben und vielleicht sogar weiterführendes Interesse für Themen in diesem Kontext zu wecken. Während wir in dieser Ausgabe den Fokus auf Metathemen der Digitalisierungsforschung gelegt haben, wenden wir uns in der nächsten Ausgabe von *easy* einer stärker inhaltlichen Ebene zu. In der kommenden Ausgabe wird die Diversität an verwendeten Daten und Methoden in der Forschung zur digitalen Transformation anhand konkreter Beispiele sichtbar. So werfen wir einen Blick auf das sogenannte Dark Web und die Frage, wie Forschung dazu aussehen kann. Außerdem geben uns Forschende einen Einblick in die Vielfalt und Komplexität an Daten, die im Rahmen neuer Forschungsprojekte zu und mit digitalen Verhaltensdaten von SmartCities und ihrer Erforschung zur Anwendung kommen.

Die Digitalisierung wird uns alle auch weiterhin in allen Lebensbereichen begleiten und unsere Lebenswelten stetig verändern. Nur wenn die Forschung sich den sich daraus ergebenden Chancen und Herausforderungen umfassend widmet und ihrem Auftrag zum Gewinn und zur Vermittlung von gesellschaftlich relevanten Erkenntnissen gerecht wird, hat sie eine Chance, diesen Prozess aktiv mitzugestalten.

Wir wünschen viel Spaß beim Lesen dieser Ausgabe!

Referenzen

- Begenat, M. (2023). Wissenschaftskommunikation in unsicheren, digitalen Zeiten. Eine Interviewstudie mit Kommunikator*innen aus der Digitalisierungsforschung. *easy_social_sciences* 68, 39-49. <https://doi.org/10.15464/easy.2023.05>
- Brennen, J. S. & Kreiss, D. (2016). Digitalization. In R. T. Craig, K. B. Jensen, E. W. Rothenbuhler & J. D. Pooley (Hg.), *The International Encyclopedia of Communication Theory and Philosophy*. John Wiley & Sons. <https://doi.org/10.1002/9781118766804.wbiect111>
- British Design Council. (2005). *Framework for Innovation: Design Council's evolved Double Diamond*. <https://www.designcouncil.org.uk/our-work/skills-learning/tools-frameworks/framework-for-innovation-design-councils-evolved-double-diamond>
- Fröhling, L., Birkenmaier, L. & Daikeler, J. (2023). Garbage in – Garbage out? Datenqualität im Umgang mit digitalen Verhaltensdaten. *easy_social_sciences* 68, 21-30. <https://doi.org/10.15464/easy.2023.03>
- Gorenšek, T. & Kohont, A. (2019). Conceptualization of Digitalization: Opportunities and Challenges for Organizations in the Euro-Mediterranean Area. *International Journal of Euro-Mediterranean Studies*, 12(2), 93–115.
- Holand, A., Svadberg, S. & Breunig, K. (2019). Beyond the Hype: A Bibliometric Analysis Deconstructing Research on Digitalization. *Technology Innovation Management Review*, 9(10), 38–50. <https://doi.org/10.22215/timreview/1274>
- Jacobs, J. A. & Frickel, S. (2009). Interdisciplinarity: A Critical Assessment. *Annual Review of Sociology*, 35, 43–65. <https://doi.org/10.1146/annurev-soc-070308-115954>
- Mede, N. G. (2021). Charakteristika der Forschung zu Wirkungen digitaler Wissenschaftskommunikation: Ein Systematic Review der Fachliteratur. In J.-H. Passoth, M. Tatari & N. G. Mede (Hg.), *Wissenschaftskommunikation in der digitalen Welt: Kartierung der Forschungslandschaft in zwei Themenfeldern* (S. 37–82). Berlin-Brandenburgische Akademie der Wissenschaften.
- Parida, V., Sjödin, D. & Reim, W. (2019). Reviewing Literature on Digitalization, Business Model Innovation, and Sustainable Industry: Past Achievements and Future Promises. *Sustainability*, 11(2), 391. <https://doi.org/10.3390/su11020391>
- Peng, T.-Q., Zhang, L., Zhong, Z.-J. & Zhu, J. J. (2013). Mapping the landscape of Internet Studies: Text mining of social science journal articles 2000–2009. *New Media & Society*, 15(5), 644–664. <https://doi.org/10.1177/1461444812462846>
- Reis, J., Amorim, M., Melão, N. & Matos, P. (2018). Digital Transformation: A Literature Review and Guidelines for Future Research. In Á. Rocha, H. Adeli, L. P. Reis & S. Costanzo (Hg.), *Trends and Advances in Information Systems and Technologies* (S. 411–421). Springer International Publishing. https://doi.org/10.1007/978-3-319-77703-0_41
- Rutjens, B. T., Heine, S. J., Sutton, R. M. & van Harreveld, F. (2018). Attitudes Towards Science. In J. M. Olson (Hg.), *Advances in Experimental Social Psychology* (Bd. 57, S. 125–165). Academic Press. <https://doi.org/10.1016/bs.aesp.2017.08.001>
- Sills, J. & Pittinsky, T. L. (2015). America's crisis of faith in science. *Science*, 348(6234), 511–512. <https://doi.org/10.1126/science.348.6234.511-a>
- Simon, S. T. & Schmitt, J. B. (2023). Agile Science: Co-Creating Research on Digital Transformation. *easy_social_sciences* 68, 12-20. <https://doi.org/10.15464/easy.2023.02>
- van Eimeren, B. & Frees, B. (2013). Rasanter Anstieg des Internetkonsums – Onliner fast drei Stunden täglich im Netz. *Media Perspektiven*, 7–8, 358–372.
- Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Qi Dong, J., Fabian, N. & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research*, 122, 889–901. <https://doi.org/10.1016/j.jbusres.2019.09.022>
- Vorderer, P. (2015). Der mediatisierte Lebenswandel. *Publizistik*, 60(3), 259–276. <https://doi.org/10.1007/s11616-015-0239-3>
- Weiß, J. & Stier, S. (2023). Die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten. *easy_social_sciences* 68, 31-38. <https://doi.org/10.15464/easy.2023.04>

Josephine B. Schmitt

Center for Advanced Internet Studies (CAIS)

E-Mail Josephine.Schmitt@cais-research.de

Josephine B. Schmitt ist wissenschaftliche Koordinatorin am Center for Advanced Internet Studies (CAIS). Dort befasst sie sich u.a. mit der Erforschung und Entwicklung von innovativen Konzepten für die interdisziplinäre Zusammenarbeit in der Digitalisierungsforschung. Sie forscht weiterhin zu Inhalt, Verbreitung und Wirkung von Hate Speech, extremistischer Propaganda und (politischen) Informations- und Bildungsangeboten im Internet. <https://orcid.org/0000-0002-4689-3049>

Julian Kohne

GESIS – Leibniz-Institut für Sozialwissenschaften

E-Mail julian.kohne@gesis.org

Julian Kohne ist wissenschaftlicher Mitarbeiter bei GESIS, Köln und Doktorand der Psychologie an der Universität Ulm. Bei GESIS arbeitet er an der Entwicklung zugänglicher Smartphone-Datenerfassungsinfrastrukturen für Sozialwissenschaftler*innen. In seinem Promotionsprojekt entwickelt er transparente Methoden zur Datenspende für WhatsApp-Chatprotokolle und untersucht, wie unterschiedliche soziale Beziehungen in Online-Kommunikationsmustern ausgedrückt werden. Sein übergreifendes Forschungsinteresse gilt der Nutzung digitaler Verhaltensdaten, insbesondere der Text- und Netzwerkdaten, um sozialwissenschaftlicher Forschungsfragen zu bearbeiten. <https://www.juliankohne.com>

Johannes Breuer

GESIS Leibniz-Institut für Sozialwissenschaften & Center for Advanced Internet Studies (CAIS)

E-Mail johannes.breuer@gesis.org

Johannes Breuer ist Senior Researcher im Team Survey Data Augmentation bei GESIS, Köln und Ko-Leiter des Teams Research Data & Methods am Center for Advanced Internet Studies (CAIS). Seine Forschungsinteressen sind die Nutzung und Wirkung digitaler Medien, Computational Methods, digitale Spurdaten, Datenmanagement und Open Science. <https://www.johannesbreuer.com>



Agile Science: Co-Creating Research on Digital Transformation

Samuel T. Simon & Josephine B. Schmitt

The dynamics of the digital transformation generate important and complex research questions: disruptive technological upheavals are entangled with serious social consequences and their effects and mechanisms need to be researched to be better understood. But the complex societal changes brought along by digital innovations also challenge science and research. So far, research on digital transformation often does not adequately meet the challenges created by the intersection of social and technological aspects. Borrowing from participatory and co-creative innovation approaches, we suggest the concept of “Agile Science”, i.e., a balanced structure for disciplined work and interdisciplinary collaboration, which allows for adaptability and participation. With this, we want to shape the future of innovative and responsive research on digital transformation. We aim to support a shift toward an understanding of, and accountability for, increasing complexities while staying in touch with affected citizens and generating relevant findings and solutions for them. The present paper introduces the main ideas of this concept and illustrates this exemplarily by describing the Research Innovation Hub (RIH) at the Center for Advanced Internet Studies (CAIS).

Die komplexen gesellschaftlichen Veränderungen, die mit der digitalen Transformation einhergehen, fordern auch Wissenschaft und Forschung heraus. Die bisherige Forschung zu digitaler Transformation wird den Herausforderungen, die sich aus der Überschneidung von sozialen und technologischen Aspekten ergeben, jedoch oft nicht gerecht. In Anlehnung an partizipative und co-kreative Innovationsansätze schlagen wir vor diesem Hintergrund das Konzept der „Agilen Wissenschaft“ vor. Darunter verstehen wir eine balancierte Struktur von interdisziplinärer Kollaboration und disziplinärem Arbeiten. Diese soll Partizipation und Anpassungsfähigkeit gewährleisten. Wir wollen die Zukunft einer innovativen und reaktionsfähigen Forschung zur digitalen Transformation so gestalten, dass die zunehmende Komplexität verstanden und verantwortet wird, während gleichzeitig der Kontakt zu den betroffenen Bürgerinnen und Bürgern aufrechterhalten wird und relevante Erkenntnisse und Lösungen für sie geschaffen werden. Der vorliegende Text stellt die Grundgedanken dieses Konzepts vor und veranschaulicht es exemplarisch anhand des Forschungsin kubators des Center for Advanced Internet Studies (CAIS).

Keywords: Agile Science, innovation, co-creation, participatory science, digital transformation

Digital Transformation as a Societal Challenge and Subject of Research

The digital transformation of society has reached an extent where technological changes

influence not only technical things but all areas of our lives as well. For example, smartphones change our daily interactions and communication styles. The way our social interactions work, in turn, impacts how digital innovations come to life. Therefore, it's not exaggerated to say that the digital transformation affects all aspects of life from work to communica-

tion, shopping, business, or administrative tasks. Consequently, there is also a great amount of scientific research investigating the implications and effects of social, political, economic, cultural, and technical changes brought about by the digital transformation. Of course, scientific research itself also does not remain unaffected by social developments. The dynamics of the digital transformation with disruptive technological upheavals and serious social consequences also challenge science and research. However, the established realities of the science system (e.g., the long time it takes from a technological disruption to generating a research idea and publishing it) often fail to meet these challenges. Against these backdrops, in this article, we want to introduce readers to an innovative approach for research on digital transformation, which takes perspectives of various societal stakeholders strongly into account: *Agile Science*.

But first, who are we? We, Josephine and Samuel, are working at the Center for Advanced Internet Studies (CAIS), where we developed and tested the idea of *Agile Science* together with different colleagues¹. The CAIS is a research institute for digital transformation research in North Rhine-Westphalia funded by the regional ministry for culture and science. Developing potentials of digital transformation for human interests, needs and capabilities is one of the central challenges of democratic societies. By proposing evidence-based solutions, at CAIS we contribute to shaping the digital transformation in a way that puts people at the center. As we want to contribute to the people-friendly and trustworthy design of a digital society, our research also includes a strong focus on the dialog with the public. With this paper we would like to introduce our idea of *Agile Science* in more detail to a broader public. With this we do not only want to provide mere information about what we were - and are - doing. We also want to open the space for a discussion about aims, tasks, and responsibilities of scientific

research in general and research on digital transformation in particular. Before we elaborate on Agile Science and show how scientific research can learn from other societal fields such as business innovation, we want to dive deeper into the challenges for digital transformation research.

Challenges for Digital Transformation Research

We identified three main challenges for digital transformation research.

Challenge 1: The General Public as important Stakeholders

In digital transformation research, the interests and needs of the general public should be considered from the very beginning as they are the context in which the digital transformation happens and shape its conditions. This is important for several reasons: The mission of science is to gain knowledge for the benefit of the public. It is therefore necessary to produce scientifically sound insights into the manifestations and consequences of the digital transformation that are useful for society at large. In many countries, scientific research is, furthermore, mainly funded through public funds. Thirdly, the importance of research that can interact with society outside the proverbial ivory tower seems to be important to ensure relevance.

In many other contexts (e.g., economics), responding to the needs of customers or stakeholders in the development of products and services is critical to success (Stickdorn et al., 2017). While, of course, science is intentionally different from producing industries, let's still play with the analogy for a moment. Focusing on the customer directly helps the success of a service. This is not to say that science should stick to economic principles of selling things quickly and in huge amounts. Here, the success would be societal relevance. In this case,

¹ We thank Matthias Begenat, Christoph Bieber, Maximilian Brenker, Anne Goldmann und Daniel Zimpel.

the product or service would be the solutions created by scientists. This does not necessarily call for immediate solutions, keeping in mind that, for example, Arts and Humanities aim for rather long-term impacts on society. If scientific research is supposed to successfully create impact, and real-world applicability, scientists should orient their work towards the needs of societal stakeholders. Ideally, while doing scientific work, the people – as the target audience for the scientific service (i.e., relevance-oriented research) – could even be included at specific points in the process (e.g., in the context of citizen science). Additionally, it is possible to enable exchange with relevant stakeholders through reappearing feedback loops at multiple moments in the research process.

Unfortunately, these ideas often remain unused. Regularly, research topics are defined top-down, meaning that institutions that fund research specify topics which they consider as societally important. Or, in other cases, scientists define them, for example, based on personal interests or a certain research project they are part of. Society usually interacts with research either as a passive subject to be explored or only at the end of a project when results are presented to the public, for example, in the media. The danger here is that scientific research – even when done with the utmost professional excellence – might miss the points that are most relevant to society.

Challenge 2: The Complexity of the Field

A second challenge lies in the constantly evolving novelty of the field of digital transformation. The complexity and rapid change of research questions increase the need for productive synergies through innovative inter- and transdisciplinary collaborations, for instance between social scientists and computer scientists. For example, when investigating how new technologies like social networks and dating platforms are changing our social relationships, both sets of expertise are crucial: We need to know how social

relationships work *and* we need to know how the digital platforms operate. This can be difficult to achieve because interdisciplinary collaboration does not just mean putting researchers of different backgrounds together. It also requires time, structure, and effort to make sure researchers can reach and maintain disciplinary expertise on top of having to work in an interdisciplinary manner. Moreover, it is usually necessary to reach common definitions of theories, methods, and data, to arrive at a shared understanding. Yet, this is often hindered by rigid hierarchical, and often disciplinary organizational structures at universities and other research institutions, or by the uncritical pursuit of disciplinary research paths. The problem with this approach of limited disciplinary perspectives is that this does not account for the complexities of real life, and, therefore, does not create the desired impact for other societal stakeholders. In short: Complex research questions about the digital transformation require interdisciplinary teams with diverse backgrounds to work on them. However, such diverse teams also need suitable supporting structures to render them effectively.

Challenge 3: The Dynamics of the Field

Third, research should be adaptable and react flexibly to changing conditions to achieve socially significant results. Science usually has a long reaction time for processes, such as topic identification, data collection, or publication. This can be problematic, especially for digital transformation topics due to rapid technological change. Take, for example, the platform Twitter, which has recently undergone intensive and quick changes, which, in turn, also affect the work of the scientific community. In contrast to this, traditional project management approaches with long-term and rigid planning have been the norm in many parts of the science system up to now. This leaves little to no room for the adaptation of research questions and processes. It only allows for a few cumbersome feedback phases and may

lead to creating a perfectionist error culture, which inhibits flexibility and innovation.

» ***The necessity of a strictly structured process that creates fertile ground for focused disciplinary expertise and productive, participatory collaboration.*** «

Against this background, we at the Center for Advanced Internet Studies (CAIS) are developing a concept for *Agile Science* to shape the future of responsive and responsible research on digital transformation. At first, the task of fostering productive research activities may seem paradoxical or at least never fully achievable if flexibility and openness to uncertainty as well as relevant application-oriented results are demanded. Nevertheless, as we will show later, there are ways of structuring uncertain processes without limiting adaptability to invite a constructive approach toward uncertainty and facing potential failure. This can be called structured heterogeneity, meaning intentional diversity within a process that still provides an organized and beneficial way of dealing with challenges, diverse stakeholders, and unpredictability.

How can Science become “Agile”?

As Agile Science, we understand science in accordance with the analogy from the beginning of this paper: Science as a service to society. From this perspective, Agile Science refers to scientific research that can continuously integrate new and socially significant questions and challenges, as well as the affected stakeholders, into the research processes. Ideally, it combines interdisciplinary research with responsiveness and adaptability in the face of

a rapidly changing subject matter. To be clear, this does not mean that scientific research should be done rapidly and superficially. On the contrary, it points to the necessity of a strictly structured process that creates fertile ground for focused disciplinary expertise and productive, participatory collaboration.

For the concept of Agile Science, we adapt approaches from agile project management. The concept of agile project management has come a long way since first being introduced in software development in the early 1990s. Since then, it has been experimented with and adapted for various areas of application. One of the current conceptual versions – called „Modern Agile“ (modernagile.org) – remembers the initial core principles and generalizes them in a way so that they can be applied to any project or situation. The focus of agile project management is on individuals and interactions, the functionality of systems, customer orientation, and the permanent willingness to adapt project goals and processes in order to be able to react to changing conditions in the best possible way (agilemanifesto.org, 2001). Moreover, for our understanding of Agile Science, practical experience and methods from the field of strategic organizational development (e.g., design thinking, service design) are examined regarding their transferability to digital transformation research.

We suggest applying Agile Science at least to two areas in the scientific process: 1) identification of research topics and 2) scientific collaboration, as shown in the following.

Identifying Relevant Research Topics in a Complex and Dynamic Environment

As discussed above, we understand digital transformation research as an area of dynamic and sometimes unpredictable changes. New platforms emerge, new technologies are invented or significantly improved, the legislation keeps changing rapidly, and so are the social norms that guide how technologies are used. Because of this, achieving participation, interdisciplinarity, and relevance is particularly

challenging in this field. Consequently, a structured approach is needed to successfully integrate the needs of relevant stakeholders (see Challenge 1) and an iterative procedure amidst the uncertainties of the context (see Challenge 3). The innovation framework Double Diamond (British Design Council, 2005, see Figure 1) provides such a structure.

The main elements of this framework are the iterative changes between diverging and converging phases that allow both the focus on exploring as well as specifying a problem. Thus, the framework is especially helpful for identifying potentially relevant research topics in digital transformation research that are yet unknown. The structure, moreover, calls for integrating diverse perspectives of various societal stakeholders to meet the scope of the problem.

The first diamond holds the phases *Discover* and *Define*, which serve to “understand [...] what the problem is (e.g., which topics are relevant in digital transformation research). It involves speaking to and spending time with people who are affected by the issues. [...] The insight gathered [...] can help you to define the challenge in a different way” (ibid.). The second diamond contains the phases *Develop* and *Deliver*. It “encourages people to give different answers to the clearly defined problem [...] and co-designing with a range of different people” (ibid.). Meaning that: Potential solutions are tested at early stages of the process with relevant stakeholders (e.g., citizens, politicians), rejected, or improved. The separation between the two “diamonds” helps overcome one crucial challenge that – as already pointed out above – we often see unaddressed: Making sure to identify the *right* problem, before attempting to solve it.

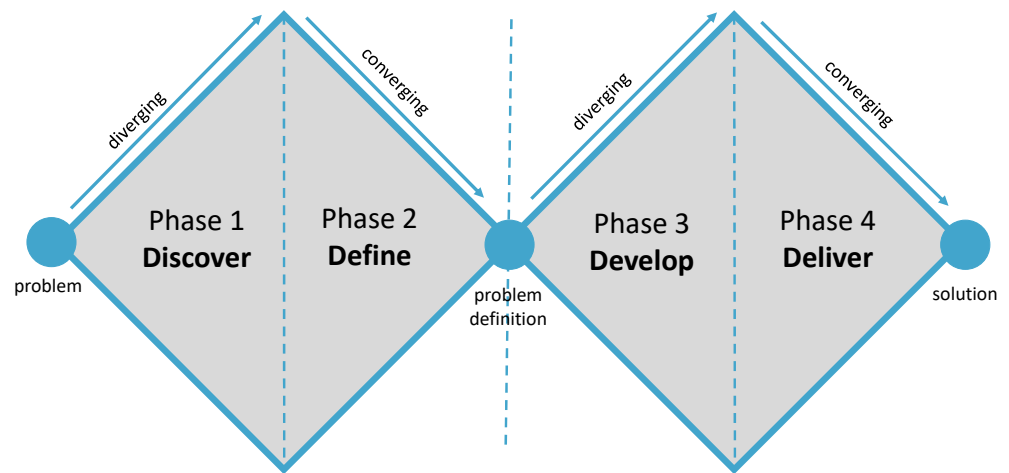


Figure 1 The four phases of the Double Diamond

» **Principles of agile project management are very helpful to foster interdisciplinary exchange and collaboration between researchers.** «

Following the path laid out by the Double Diamond, researchers can define research topics and ensure participation, interdisciplinarity, and societal relevance. Building on to that structure with elements of agile project management the challenges of reacting flexibly to changing framework conditions can be met productively.

Agile Elements in Scientific Collaboration

To address the second challenge of digital transformation research, principles of agile project management are very helpful to foster interdisciplinary exchange and collaboration between researchers. These principles are usually applied through three features working together – which are worth to be transferred to research on digital transformation: stakeholder-centricity, the encouragement of taking a constructive approach to failure through frequent feedback cycles, and an iterative procedure. The first feature can be implemented, for example, as a process as described

in the previous paragraph. The other could mean feeding back interim research results to key societal stakeholders at regular intervals in order to review relevance and impact and, if necessary, strengthen them through readjustments in the research process (e.g., adaptations of research questions, methods, etc.). This approach is complemented by a clear distribution of roles within the team. According to agile project management, these roles are task-based rather than referring to a strict hierarchy in the traditional sense. One of these roles is a so-called facilitator, aiming to help team members to be courageous and solution-oriented with their expertise instead of shying away from a potentially critical opinion.² The facilitator takes up a role similar to that of a workshop host, a project manager, or what has recently been described as a *neo-generalist* (Kern, 2021): a person especially skilled to bring together experts from specific backgrounds – here researchers from different disciplines – and helping them to collaborate effectively. This serves to shift the responsibility of managing the overall structure of the process of a research project to the facilitator. This, in turn, frees the other team members to focus on the content-related challenges at hand. In short: The facilitator plans, oversees, and manages the structure of the process, and, thereby, creates enough room and security for the other team members to maneuver through any potential change in circumstances and the uncertainties brought about by them. Both aspects – the concise structure of the process through iterative phases and the use of a facilitator – serve as the main adaptations to make scientific research processes more agile, flexible, and participatory while still being focused on the societal relevance of findings and solutions. In the following, we will illustrate the implementation of Agile Science exemplarily by giving an overview of the Research Innovation Hub (RIH) at CAIS.

² Cf. previous similar concepts such as the ‘boundary spanner’ (e.g., Bednarek et al., 2018) or the ‘knowledge broker’ (e.g., Bergenholtz, 2011) carry great resemblance to the ‘facilitator’.

The CAIS Research Innovation Hub as a Practical Example

In 2017, the CAIS started as an institution for funding international guest researchers, working groups, and smaller conferences. However, the main aim of the CAIS always has been to develop into a proper research institute with permanent researchers investigating societally important topics in the field of digital transformation. Nevertheless, it had to wait for the appropriate funding from the state of North Rhine-Westphalia. In the starting phase between Summer 2019 and Spring 2021, when the CAIS prepared for becoming a full research institute, the RIH was established to develop and test a concept of agile science. This had two main areas of responsibility: 1) identifying topics for the future CAIS research programs in a participatory process ensuring actuality and societal relevance, and 2) developing and testing methods for interdisciplinary collaboration in digital transformation research.

Identifying Topics for CAIS Research Programs

In a structured process that combined surveys, workshops, and group discussions with researchers as well as representatives of the public and various fields of practice, topics were identified for the first four CAIS research programs. This aimed at the two-fold goal of identifying socially relevant research topics in the field of digital transformation research while, at the same time, including different stakeholders as a participatory human-centered process. The project was structured along the four phases of the Double Diamond (see above). Each phase included at least one qualitative or quantitative study conducted between September 2019 and February 2021. For an overview of the process and the different studies included in the process see Figure 2.³

³ You can find an overview of *all studies, methods, and results here*.

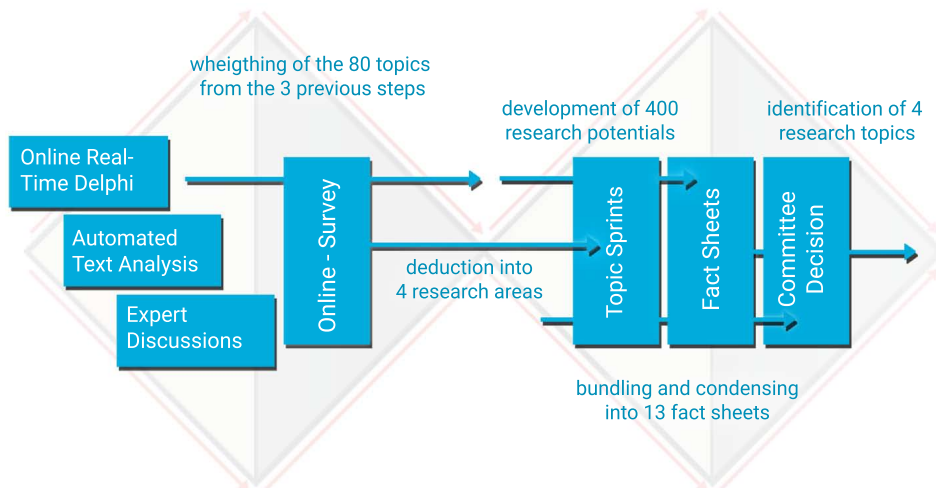


Figure 2 Overview of the process of identifying research topics.

We involved the whole spectrum of stakeholders (e.g., researchers, practitioners, laypeople) to make sure that the future research output will match actual societal needs. Key stakeholders were frequently engaged and repeatedly included into the different steps of the process to validate and improve on intermediate results. Based on this process, the four topics for the CAIS research programs were selected: a) Politics, Inclusion, and Participation, b) Education and Digitalization, c) Ethics and AI, and d) Sustainability.

With the topic-finding process, we aimed to account for the dynamic relations between society and science, exploring current topics around digital transformation and their requirements for participatory research. Structuring the phases along the Double Diamond supported a participatory, co-creative, and solution-oriented process. The systematic collection of topics is the starting point for a long-term monitoring, which intends to map how opinions on, and relevance of digital transformation topics change over time. The further institutionalization of this process is an important steppingstone for understanding the underlying dynamics of digital transformation, being able to effectively research them, and create sustainable outcomes with real societal impact.

Adding a specific role that carries the responsibility of structuring and managing the process – the facilitator – helped us to be less

bound by hierarchy and power structures that might discourage openness. Together with the clearly structured but adaptive setting, the use of a facilitator fostered psychological safety for all involved and thereby enabled a constructive approach to potential setbacks. This, in turn, was a fertile ground for real co-creative work: everyone in the process

could contribute honestly and freely without hesitation, drawing inspiration from the diversity of everyone involved.

Practicing Interdisciplinary Collaboration

In the RIH, we further examined and improved cross-domain networking and innovative interdisciplinary collaboration between researchers in digital transformation research in six pilot research projects. We used agile methods to improve the exchange between the research teams. Among other things, we employed methods from the field of design thinking, which are methodologically designed primarily for activation, feedback, creativity, and reflexivity.

Additionally, in this early stage of building the institute, we conducted group discussions with researchers of different disciplinary backgrounds.⁴ Through that, we identified researchers' needs and concerns in interdisciplinary settings as well as methods (e.g., researchers' call for a facilitator) to be applied and tested within the above-mentioned pilot

⁴ For more detailed information regarding the results of this study, see <https://www.cais-beta.de/forschung/inkubator/forschungsinkubator-expertinnen-diskussion/>.

» ***Responsible integration of various stakeholders while working within an iterative structure that allows for flexibility.*** «

projects. This served as a live testing phase with short feedback loops to generate insights on how aspects of Agile Science can be implemented in future research programs at CAIS.

Conclusion and Outlook

By implementing agile elements in different phases of the research process, we aim to address the above-mentioned challenges for digital transformation research. Focusing on agile principles in the RIH enabled the responsible integration of various stakeholders into different parts of the research process while working within an iterative structure that allows for flexibility. Keeping focus despite changing circumstances and ensuring the relevance of the identified research topics was achieved. At the same time, in the interdisciplinary working settings, this allowed for many points of contact and opportunities for participation and iterative collaboration with our broad set of stakeholders. Consequently, our agile, participatory approach not only provided us with a comprehensive and detailed number of insights but also safeguarded the relevance of output from the very outset. Of course, achieving such high levels of contextualization and human-centricity while managing the process requires additional effort. But investing this effort and, thereby, accounting for increased complexities with the extra effort of an agile approach only seems logical.

The use of the facilitator in digital transformation research is strongly recommended also

in other application contexts, for example, onboarding new members of research teams, enabling a common understanding of different disciplinary perspectives, creating space for early and safe feedback, or managing the balance between disciplinary expertise and interdisciplinary collaboration. Nevertheless, the potentially added value of those functions of the facilitator also needs to be investigated and evaluated systematically in the future.

Unfortunately, in recent years, the term “agile” has degenerated into a rather empty buzzword. What has been lost is in many contexts the knowledge, that agile refers to a certain mindset on one hand, but also to concrete methods for structuring solution-oriented work processes. Re-establishing this awareness is an important and necessary task for implementing innovative and reflexive research in a world characterized by volatility, uncertainty, complexity, and ambiguity. A facilitator could support this – not only in research on digital transformation.

To conclude, the agile, participatory, and co-creative approaches within the RIH at CAIS are crucial elements for successfully conducting need- and relevance-based research. Of course, all the methods described above require additional effort. To account for the complexities of the digital transformation and to guarantee societal relevance and applicability of responsive and responsible research on digital transformation this is not a sacrifice but both an essential requirement as well as an extremely promising long-term investment. This would seem a fitting approach to go along with the dynamic nature of digital transformation research. However, on the side of the university policy and the research institutions, openness to innovative and agile working methods is necessary so that they can be implemented sustainably.

References

- agilemanifesto.org. (2001). *Manifesto for Agile Software Development*. <https://agilemanifesto.org/>
- Bednarek, A. T., Wyborn, C., Cvitanovic, C., Meyer, R., Colvin, R. M., Addison, P. F. E., Close, S. L., Curran, K., Farooque, M., Goldman, E., Hart, D., Mannix, H., McGreavy, B., Parris, A., Posner, S., Robinson, C., Ryan, M., & Leith, P. (2018). Boundary spanning at the science–policy interface: The practitioners’ perspectives. *Sustainability Science*, 13(4), 1175–1183. <https://doi.org/10.1007/s11625-018-0550-9>
- Bergenholtz, C. (2011). Knowledge brokering: Spanning technological and network boundaries. *European Journal of Innovation Management*, 14(1), 74–92. <https://doi.org/10.1108/146010611111104706>
- British Design Council. (2005). *Framework for Innovation: Design Council’s evolved Double Diamond*. <https://www.designcouncil.org.uk/our-work/skills-learning/tools-frameworks/framework-for-innovation-design-councils-evolved-double-diamond/>
- Kern, S. (2021). Scanner-Persönlichkeiten, Multipassionates, Future-Skill-Held:innen & Co. In S. Kern (Ed.), *Future Skill Vielseitigkeit* (pp. 21–27). Springer VS. https://doi.org/10.1007/978-3-658-35519-7_4
- Modernagile.org. (n. d.). Modern Agile. <https://modernagile.org/>
- Sills, J., & Pittinsky, T. L. (2015). America’s crisis of faith in science. *Science*, 348(6234), 511–512. <https://doi.org/10.1126/science.348.6234.511-a>
- Stickdorn, M., Hormess, M., Lawrence, A., & Schneider, J. (2017). *This is service design doing: Applying service design thinking in the real world: A practitioner’s handbook*. O’Reilly Media.

Josephine B. Schmitt

Center for Advanced Internet Studies (CAIS)

E-Mail Josephine.Schmitt@cais-research.de

Josephine B. Schmitt is research coordinator at the Center for Advanced Internet Studies (CAIS). Her work includes researching and developing innovative concepts for interdisciplinary collaboration in digitization research. She also conducts research on the content, dissemination and impact of hate speech, extremist propaganda and (political) information and education offerings on the Internet. <https://orcid.org/0000-0002-4689-3049>

Samuel T. Simon

Center for Advanced Internet Studies (CAIS)

E-Mail Samuel.Simon@cais-research.de

Samuel T. Simon is the project lead for “Agile Academia” at the Center for Advanced Internet Studies (CAIS). Besides his practical work of conducting workshops and other co-creative work formats for an agile, interdisciplinary research on digital transformation among the researchers at CAIS, he is currently researching and specifying the role of the facilitator in academic contexts in his dissertation project. <https://orcid.org/0000-0002-0127-4498>



Garbage in – Garbage out?

Datenqualität im Umgang mit digitalen Verhaltensdaten

Leon Fröhling, Lukas Birkenmaier & Jessica Daikeler

Während in den quantitativen Sozialwissenschaften Umfragedaten seit jeher das Herzstück der Informationsgewinnung bilden, spielten Beobachtungsdaten und andere Datenquellen eine eher untergeordnete Rolle. Soziale Medien und mobile Endgeräte lassen nun digitale Verhaltensdaten immer mehr in den Mittelpunkt sozialwissenschaftlicher Forschung rücken. Doch selbst die innovativsten und umfangreichsten Datenmengen sind unzureichend, wenn sie nicht von hoher Qualität sind. Dieser Artikel diskutiert anhand eingängiger Beispiele die grundlegenden Herausforderungen bei der Analyse digitaler Verhaltensdaten und präsentiert einen zentralen Ansatz zur Evaluation ihrer Qualität.

While survey data has always been at the heart of information gathering in the quantitative social sciences, observational data and other data sources have played a rather subordinate role. Social media and mobile devices are now making new forms of digital behavioral data increasingly central to social science research. However, even the most innovative and comprehensive data sets are insufficient if they are not of high quality. This article discusses the basic challenges of analyzing digital behavioral data using several use cases. Ultimately, it presents one central framework to evaluate the applicability of digital behavioral data for social science research.

Keywords: Beobachtungsdaten, digitale Verhaltensdaten, Repräsentativität, Validität, Datenqualität

Digitale Verhaltensdaten für die Sozialwissenschaften

Eine Vielzahl wirtschaftlicher, sozialer und gesellschaftlicher Vorgänge baut mehr und mehr auf digitalen Technologien und Onlineplattformen auf. Von der Kommunikation mit Freund*innen und Verwandten, über die Nutzung von digitalen Medieninhalten, bis hin zum Kauf und Vertrieb von Produkten und Dienstleistungen, prägen digitale Formate und Werkzeuge zunehmend unseren Alltag und führen zu großen Mengen digitaler Verhaltensdaten. Digitale Verhaltensdaten sind

digitale Spuren menschlichen Verhaltens, wie sie beispielsweise von Online-Plattformen, smarten Geräten und speziellen Forschungssensoren erfasst werden. In Verbindung mit neuen, computergestützten Auswertungsmethoden bieten diese digitalen Verhaltensdaten neuartige Potentiale zur Beschreibung sozialer und politischer Prozesse – von globaler Vernetzung über politische Polarisierung hin zur Beschreibung von Interaktionsmustern im digitalen Raum (King, 2011).

Der Mehrwert einer systematischen Auswertung dieser Daten ist enorm. In der Wirtschaft etwa passen Internetkonzerne wie Google, Facebook oder Amazon ihre Produktempfehlungen und Werbeanzeigen mit

digitalen Daten individuell an die einzelnen Kund*innen an, und entwickeln sich so mehr und mehr zu global dominierenden Konzernen.

In der sozialwissenschaftlichen Forschung finden digitale Verhaltensdaten ebenso regelmäßig Anwendung. An der Schnittstelle zwischen den Sozialwissenschaften (*Social Science*) und der Informatik (*Computer Science*) entsteht das neue Forschungsfeld der *Computational Social Science*, welches einerseits klassisch sozialwissenschaftliche Phänomene mit neuen Daten und Methoden untersucht, aber auch neue, durch die zunehmende Digitalisierung erst auftretende Phänomene erforscht.

Datenqualität – eine zentrale Herausforderung

Obwohl digitale Verhaltensdaten für Forschende eine neue und vielversprechende Datenquelle darstellen, ist die Erhebung und Auswertung dieser Daten in der Praxis oft mit großen Herausforderungen verbunden. Bereits kleine Entscheidungen in der Sammlung, Verarbeitung und Auswertung digitaler Verhaltensdaten können einen gewichtigen Einfluss auf die Qualität von Daten und damit auch die Ergebnisse ganzer Studien haben. Ein populäres Beispiel hierfür ist der „Google Flu Trends“-Algorithmus (GFT), welcher erstmals 2008 von Google veröffentlicht wurde (Ginsberg et al., 2009). Die Idee hinter der simplen Anwendung ist, auf Basis von Google-Suchanfragen zu Grippe-symptomen in Echtzeit zu erkennen, wie sich die Grippe an verschiedenen

Orten der Welt ausbreitet. Obwohl Grippewellen so anfangs noch mit beeindruckender Präzision vorhergesagt werden konnten, verschlechterte sich in den Jahren darauf die Vorhersagequalität teils erheblich, was schließlich zur Einstellung des Projektes im Jahr 2015 führte. Abbildung 1 zeigt den Verlauf der Grippe zusammen mit den GFT-Prognosen für die Jahre 2009 bis 2013. Während GFT in den Anfangsjahren noch eine hohe Übereinstimmung mit den Daten der nationalen Gesundheitsbehörde in den USA (CDC) hatte, nahm die Vorhersagegenauigkeit über die Jahre zunehmend ab. Das ging so weit, dass für 2013 der Anteil der vorhergesagten Grippeerkrankungen mehr als doppelt so hoch wie der Anteil der tatsächlich gemeldeten Erkrankungen war. Grund für die fehlerhaften Vorhersagen waren zum einen kleine Änderungen des Google-Suchalgorithmus sowie des Verhaltens der Nutzenden, welche einen direkten Effekt auf die Vorhersagequalität hatten (Lazer et al., 2014). Zum anderen konnten Lazer et al. (2014) zeigen, dass sich die Symptome jahreszeitlicher Erkrankungen wie Erkältungen und die der Grippe, und damit auch die jeweiligen Suchanfragen, zu stark ähneln. Da das GFT-Modell diese aber

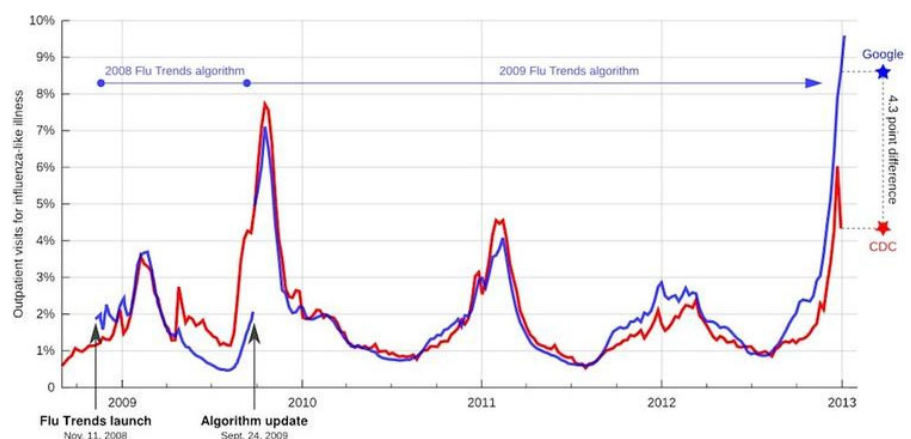


Abb 1 Vorhersage von Grippewellen mithilfe digitaler Verhaltensdaten. In Blau der Anteil der von Google Flu Trend vorhergesagten Grippeerkrankungen und in Rot der Anteil der tatsächlich registrierten Grippeerkrankungen in der Gesamtbevölkerung, über den Zeitverlauf von 2009 bis 2013. Spätestens ab Mitte 2012 ist eine deutliche Abweichung der Vorhersage von den tatsächlich gemeldeten Daten zu erkennen.

Quelle: <https://www.wbur.org/news/2013/01/13/google-flu-trends-cdc> (Zuletzt abgerufen am 16.12.2022).

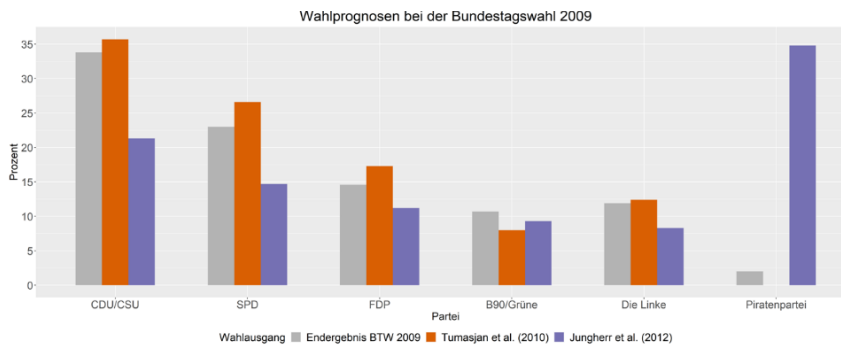


Abb. 2 Vergleich der tatsächlichen Stimmanteile ausgewählter Parteien bei der Bundestagswahl 2009 mit den Vorhersagen von Tumasjan et al. (2010) und Jungherr et al. (2012).

nicht unterscheiden konnte, funktionierte es als Vorhersagemodell jahreszeittypischer Erkrankungen (“winter detector”), aber nicht spezifisch für die Grippe.

Der Versuch, Wahlergebnisse durch Social Media Daten vorherzusagen, dient als weiteres Beispiel für die Herausforderungen in der Arbeit mit digitalen Verhaltensdaten. Ein solches Modell, das einzig anhand der Häufigkeit der Erwähnungen von politischen Parteien auf der Plattform Twitter Stimmanteile bei politischen Wahlen prognostizieren soll, stellen Tumasjan et al. (2010) vor. Ihr Modell, so die Autor*innen, hätte den Ausgang der Bundestagswahl 2009 mit nur minimaler Abweichung vorhergesagt. In einer Replik zeigen Jungherr et al. (2012) aber, dass die Plausibilität der Ergebnisse erheblich von den ohne weitere Begründungen getroffenen Entscheidungen der Forschenden abhängt. Eine dieser Entscheidungen stellt der Verzicht auf Parteien, die sich 2009 nicht im Bundestag befanden, dar. Insbesondere die Piratenpartei, die 2009 als aufstrebende politische Kraft stark im digitalen Raum präsent war, wurde ohne Begründung von Tumasjan et al. (2010) nicht berücksichtigt. Ebenfalls unbegründet und in ihren Auswirkungen unreflektiert blieb die Entscheidung, die Datenerhebung acht Tage vor der Bundestagswahl zu beenden. Jungherr et al. (2012) zeigen, dass die Einbeziehung der acht nicht-berücksichtigten Tage bis zur Bundestagswahl zu einer erheblich schlechteren Vorhersagequalität geführt hätte.

Abbildung 2 vergleicht die Ergebnisse der Bundestagswahl 2009 mit den Vorhersagen von Tumasjan et al. (2010) und Jungherr et al. (2012). Obwohl die nur von Jungherr et al. (2012) berücksichtigte Piratenpartei am Ende nur 2 % der Wähler*innenstimmen bekam, hätte sie nach dem Prognosemodell mit knapp 35 % aller Stimmen den Wahlsieg holen müssen. Diese hohe Diskrepanz erregt

Zweifel, ob die zugrundeliegende Methodik in der Lage ist, verlässliche und belastbare Wahlprognosen zu generieren.

Neben den unten noch näher erläuterten Problemen in der Fallstudie existieren eine Vielzahl weiterer Herausforderungen, die das Arbeiten mit digitalen Verhaltensdaten erschweren. So ist bekannt, dass die Nutzenden von Social-Media-Plattformen wie Twitter oftmals jünger, eher männlich und gebildeter als der Durchschnitt der deutschen Bevölkerung sind (Blank, 2017; Sloan, 2017). Außerdem ist unklar, welcher Anteil der ausgewerteten Tweets von automatisierten Accounts (*Bots*) oder von anderweitig nicht-wahlberechtigten Accounts (z.B. Unternehmen oder Medienanstalten) verfasst wurde. Dies verdeutlicht, dass sich auf Basis von Twitter-Daten nur schwer Aussagen über das allgemeine Wahlverhalten machen lassen.

» **Es fehlt, im Gegensatz zur Umfrageforschung, an allgemeingültigen Verfahren zur Einschätzung der Datenqualität.** «

Ebenfalls zeigen die Beispiele, dass Forschende in der Arbeit mit digitalen Verhaltensdaten eine Vielzahl an Entscheidungen treffen müssen, welche die Ergebnisse ihrer Forschung maßgeblich beeinflussen. Diese reichen von der Sammlung und Aufbereitung der Daten, über die Auswahl und Spezifizierung

der Auswertungsmethode, bis hin zur kritischen Auseinandersetzung und der Interpretation der eigenen Ergebnisse. Insbesondere weil digitale Verhaltensdaten in erster Linie nicht für Forschungszwecke erstellt, sondern lediglich für diese nachgenutzt werden, fehlen in vielen Fällen oft wichtige Hintergrundinformationen zu ihrer Entstehung und ihrem Kontext. Zudem macht die Größe vieler Datensätze mit häufig Millionen von Datenpunkten eine manuelle Überprüfung und Verarbeitung der Daten meist praktisch unmöglich. Während Forschende bei einer kleinen Stichprobe die Möglichkeit haben, beispielsweise die gesammelten Tweets selbst zu überprüfen, kann dies bei großen Datensätzen nur automatisiert durchgeführt werden. Schließlich fehlt es, im Gegensatz zur Umfrageforschung, an allgemein gültigen Verfahren zur Einschätzung der Datenqualität. Es haben sich noch keine Standards etabliert, um alle Faktoren, die für die Qualität der Daten, die Interpretation der Ergebnisse und deren Reproduzierbarkeit durch andere Forschende, einheitlich zu dokumentieren und zu bewerten. Dies ist besonders problematisch, da wissenschaftliche Forschung solche Standards benötigt, damit Forschende die Zuverlässigkeit und Transparenz ihrer Forschung sicherstellen können.

Error Frameworks zur Qualitätsprüfung

Zur Lösung der umrissenen Probleme können sogenannte *Error Frameworks* hilfreich sein. Die Idee der *Error Frameworks* stammt aus den Sozialwissenschaften, insbesondere der Umfrageforschung, wo diese seit den 2000er Jahren dazu verwendet werden, den Forschungsprozess systematisch auf potenzielle Fehlerquellen zu durchleuchten (Groves & Lyberg, 2010). Ziel eines *Error Frameworks* in der Umfrageforschung ist es, für jede Phase im Forschungsprozess, von der Stichprobenziehung bis zur Datenerhebung und -auswertung,

mögliche Fehler (*Errors*) und deren Quellen zu identifizieren. Der Begriff Fehler wird hier nicht in den Dimensionen „richtig“ und „falsch“ verwendet, sondern ist als *Verzerrung* zu verstehen. Eine Verzerrung ist eine systematische Beeinflussung der Daten, die von den Forschenden weder beabsichtigt noch kontrolliert ist, und sich damit potenziell in einem verzerrten Ergebnis der Analyse niederschlägt. *Error Frameworks* ermöglichen, die Fehler über den gesamten Forschungsprozess zu identifizieren und zu aggregieren. Dadurch können Aussagen über die Qualität der Forschung sowie über die Aussagekraft ihrer Ergebnisse gemacht werden.

In den letzten Jahren wurden *Error Frameworks* speziell für die Anforderungen der Forschung mit digitalen Verhaltensdaten angepasst (Amaya et al., 2020; Hsieh and Murphy, 2017; Sen et al., 2021). Abbildung 3 gibt einen Überblick über das *Total Error Framework for Digital Traces of Human Behavior on Online Platforms (TED-On)* von Sen et al. (2021). Dieses ist explizit für die Arbeit mit digitalen Verhaltensdaten von Online-Plattformen entwickelt worden. Das TED-On Framework orientiert sich an fünf Phasen des Forschungsprozesses mit digitalen Verhaltensdaten (Definition des Konstrukts, Auswahl der Plattform, Sammlung, Aufbereitung und Analyse der Daten). Es unterscheidet zwischen zwei verschiedenen Kategorien von Fehlern: Mess- und Repräsentationsfehlern. Messfehler beziehen sich darauf, wie das Konstrukt (z.B. die Präferenz für eine politische Partei) aus den Daten gemessen wird (z.B. durch die Anzahl ihrer Erwähnungen auf Twitter), während Repräsentationsfehler sich auf Fehler in der Erfassung der für die Studie relevanten Personengruppe beziehen (ob etwa die Zusammensetzung der analysierten Twitter-Nutzenden die gesamte Wahlbevölkerung gut repräsentiert). In den folgenden Absätzen wird nun anhand eines Fallbeispiels das TED-On mitsamt seiner verschiedenen Fehlerkategorien angewandt. Es wird aufgezeigt, wie die Verwendung des *Error Frameworks* Forschende bei der kritischen Reflektion ihrer Arbeit mit

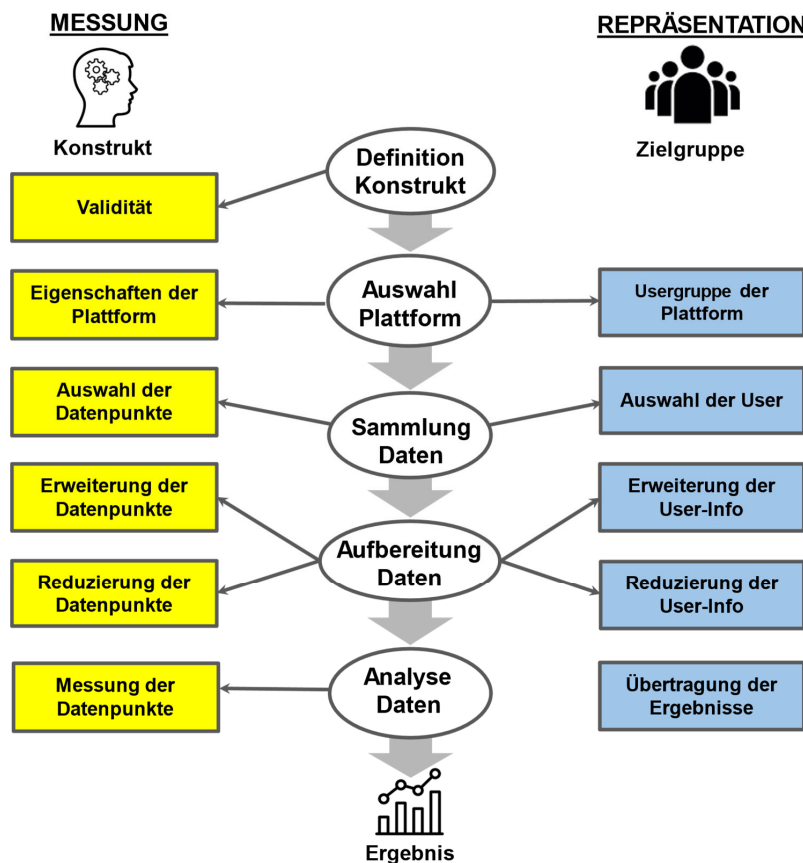


Abb.3 TED-On Framework (übersetzte Abbildung aus Sen et al., 2021).

digitalen Verhaltensdaten unterstützen kann. Als Fallbeispiel wird die in Abschnitt 2 bereits eingeführte Studie von Tumasjan et al. (2010) zur Vorhersage des Bundestagswahlergebnisses anhand von Twitter-Daten betrachtet.

Anwendung des TED-On Frameworks auf ein Fallbeispiel

Definition und Operationalisierung des Konstrukts

Als ersten Schritt im Forschungsprozess mit digitalen Verhaltensdaten benennt das TED-On Framework die Definition des zu untersuchenden Konstrukts. Ein Konstrukt ist das Phänomen, welches in der Studie gemessen und untersucht werden soll. Da das zu untersuchende Konstrukt zumeist nicht direkt beobachtbar ist (beispielhaft etwa die Konstrukte politische Beeinflussung oder

Ideologie), greifen Forschende auf sogenannte Operationalisierungen zurück. Eine Operationalisierung stellt dabei die Verbindung zwischen dem Konstrukt und den Daten her. Sie beschreibt also die Methode, wie das Konstrukt aus den Daten zu messen ist. Dieser Zusammenhang zwischen Konstrukt und Daten ist im besten Fall theoretisch herleitbar und empirisch (z.B. durch andere Studien) belegbar. Erfolgt die Definition des Konstrukts gar nicht oder nicht ausreichend präzise, oder wird durch die gewählte Operationalisierung das Konstrukt nicht genau gemessen,

entstehen Probleme mit der Validität des Forschungsdesigns. Validität kann hier verstanden werden als „Angemessenheit“, ob und wie weit also das Forschungsdesign geeignet ist, die Forschungsfrage zu beantworten.

Im Fallbeispiel ist das Konstrukt schnell gefunden: Es sollen die Stimmanteile der Parteien in der Bundestagswahl 2009 anhand von Twitter-Daten vorhergesagt werden. Die verwendete Operationalisierung ist unmittelbar nachvollziehbar, so wird der Anteil der Erwähnungen der betrachteten Parteien in den Tweets, verfasst in einem bestimmten Zeitraum, verwendet, um den Stimmenanteil bei der Wahl zu prognostizieren. Der in der Operationalisierung unterstellte Zusammenhang und damit die Validität der Studie kann zumindest angezweifelt werden. Wenn etwa Tweets, gerichtet an den Account einer bestimmten Partei, ausschließlich Ausdruck negativer Gefühle gegenüber der Partei sind, etwa im Kontext eines *Shitstorms*, dann ist nicht unmittelbar ersichtlich, wieso sich dies ausgerechnet in einem besseren Wahlergebnis der

betroffenen Partei widerspiegeln sollte. Die in der Studie gewählte Form der Operationalisierung würde diesen Wirkzusammenhang aber direkt implizieren.

Auswahl einer Plattform zur Sammlung der Daten

Nachdem das Konstrukt definiert und die Messung dieses Konstrukts aus den noch zu sammelnden Daten festgelegt ist, sieht der zweite Schritt im Forschungsprozess die Wahl einer passenden Plattform als Datenquelle vor. Der Begriff Plattform meint an dieser Stelle hauptsächlich Soziale Medien als Orte sozialer Online-Interaktionen. Bei der Auswahl ist zu beachten, dass Eigenschaften der Plattform häufig einen direkten Einfluss auf die verfügbaren Daten haben. Es ist daher wichtig zu unterscheiden, welches in den Daten zu beobachtende Verhalten unabhängig von der Plattform Rückschlüsse auf die Person hinter dem Verhalten zulässt, und welches Verhalten von den Eigenschaften der Plattform bestimmt wird. Ferner haben verschiedene Plattformen unterschiedliche User*innengruppen, wodurch bereits die Wahl der Plattform vorgibt, welche demographischen Gruppen in der Studie überhaupt betrachtet werden können.

In dem Fallbeispiel wird der Kurznachrichtendienst Twitter als Plattform gewählt. Neben der bislang guten Zugänglichkeit der Daten über die von Twitter bereitgestellte [API for Academic Research](#) spricht für Twitter auch der Ruf, eine besonders *politische* Kommunikationsplattform zu sein. Dort tauschen sich vornehmlich Akteur*innen aus dem Journalismus, den Medien, der Wissenschaft und eben der Politik aus, häufig auch über (tages-)politische Entwicklungen. Während zumindest der inhaltliche Diskurs auf der Plattform thematisch passend erscheint, bleibt die Frage, ob die User*innengruppe repräsentativ für den wahlberechtigten Teil der deutschen Gesellschaft ist, der für das gewählte Konstrukt (Stimmanteil bei der Bundestagswahl 2009) maßgeblich sein sollte, was für Twitter-Nutzende nicht zutrifft. Darüber hinaus reflektieren auch schon die

Autor*innen des Fallbeispiels, inwiefern sich das auf Twitter geltende Zeichenlimit (zur Zeit der Studie durfte ein Tweet nicht mehr als 140 Zeichen umfassen, aktuell liegt das zulässige Zeichenlimit bei 280) auf den Informationsgehalt der Tweets auswirkt – ein klassisches Beispiel für beobachtetes Verhalten, das direkt durch die Eigenschaften der Plattform beeinflusst wird.

Sammlung der Daten

Als nächstes ist im Forschungsprozess die Festlegung einer geeigneten Methode zum Sammeln der benötigten Daten vorgesehen. Mögliche Probleme und Verzerrungen sind hier in zwei gegensätzliche Richtungen denkbar. Zum einen kann es sein, dass Daten gesammelt und für die Analyse berücksichtigt werden, die für die Messung des Konstrukts und damit die gesamte Studie nicht relevant sind. Zum anderen kann es passieren, dass durch die gewählte Methodik der Datensammlung nicht alle relevanten Beobachtungen erfasst und in den Datensatz aufgenommen werden. Sowohl im Falle der zu liberal als auch im Falle der zu restriktiv gewählten Kriterien für die Inklusion von Daten führt dies zu verzerrten Ergebnissen.

Eindrucksvoll aufgezeigt wird diese Problematik wie oben bereits eingeführt in der Replikationsstudie von Jungherr et al. (2012) zu unserem Fallbeispiel. Nach Jungherr et al. (2012) verzerren die Auswahl der für die Datensammlung berücksichtigten Parteien und der festgelegte zeitliche Rahmen, für den Daten gesammelt wurden, das Ergebnis der Studie. Durch die Nichtberücksichtigung der bei Twitter-Nutzenden damals populären Piratenpartei bei Tumasjan et al. (2010) für die Prognose der Stimmanteile verschieben diese sich deutlich zu Gunsten der übrigen Parteien und nähern sich somit dem tatsächlichen Wahlausgang an (vgl. Abbildung 2). Jungherr et al. (2012) zeigen also, wie eine andere, ebenfalls plausible Entscheidung in der Auswahl der berücksichtigten Parteien und in der Festlegung des Zeitraums der Datensammlung für die Qualität und Aussa-

gekraft der Daten entscheidend sind. Jungherr et al. (2012) verdeutlichen, wie die erwarteten Stimmanteile einzelner Parteien um bis zu 6,5 Prozentpunkte schwanken, wenn der für die Auswertung betrachtete Zeitraum an Twitter-Aktivität um wenige Tage verschoben wird.

Aufbereitung der Daten

In einer typischen Studie mit digitalen Verhaltensdaten schließt sich an den Schritt der Datensammlung die Aufbereitung der Daten für die abschließende Analyse an. Zweck dieses Schrittes ist es, die gesammelten „rohen“ Daten in das für die Auswertung benötigte Format zu bringen. Die Art der Aufbereitung hängt dabei sowohl von den Eigenschaften der gesammelten Daten als auch von den Anforderungen der Analysemethode ab. Bei erweiternden Methoden der Aufbereitung werden die gesammelten rohen Daten mit zusätzlichen Informationen versehen, und bei reduzierenden Methoden werden einzelne Datenpunkte oder Informationen entfernt. Wenn es bei der Erweiterung oder der Reduzierung zu systematischen Fehlern kommt, werden dadurch auch die Daten und damit potenziell auch die Ergebnisse systematisch verzerrt.

Im Fallbeispiel kommen weder erweiternde noch reduzierende Aufbereitungsschritte zum Einsatz, da die Messung des Konstrukts direkt auf den gesammelten rohen Daten aufbaut. Häufig werden bei Twitter-Studien Tweets, die als Spam eingeordnet werden, aus den gesammelten Daten entfernt. Dies ist in vielen Fällen sinnvoll, da diese Tweets keine Auseinandersetzung mit dem jeweiligen Thema beinhalten, sondern ausschließlich Aufmerksamkeit erregen sollen. Bei der Entscheidung, ob es sich bei einem Tweet um Spam handelt oder nicht, besteht die Gefahr, dass Tweets fälschlicherweise als Spam klassifiziert („Erweiterung“ der Daten um die Information, ob es sich um Spam handelt oder nicht) und aus den Daten entfernt werden („Reduzierung“ der Daten durch das Herausfiltern bestimmter Beobachtungen). Wenn dies systematisch geschieht, etwa weil die Tweets bestimmter

User*innengruppen ähnliche Eigenschaften wie Spam-Tweets aufweisen, führt dies zu einer systematischen Nichtberücksichtigung dieser User*innengruppen für die Auswertung. Im Fallbeispiel wäre das gleichzusetzen damit, dass die Stimmen bestimmter Nutzender für die Berechnung der Wahlergebnisse nicht gezählt werden.

Analyse der Daten

Abschließender Schritt des idealtypischen Forschungsprozesses ist die tatsächliche Messung des Konstrukts aus den gesammelten und aufbereiteten Daten. Zu treffende Entscheidungen für die statistische Analyse sind etwa, auf welcher Ebene die Daten aggregiert werden, und wie aus den verschiedenen Aggregationsebenen das finale Resultat berechnet wird.

Im Fallbeispiel ist das zu messende Konstrukt die Stimmanteile der Parteien in der Bundestagswahl 2009. Tumasjan et al. (2010) berechnen die Stimmanteile für die berücksichtigten Parteien, in dem sie die Häufigkeit ihrer Erwähnungen in Tweets durch die Gesamtzahl der gesammelten Tweets teilen. In einer alternativ denkbaren Form der Aggregation könnten Tweets auf der Ebene der einzelnen Tweet-Autor*innen aggregiert und in Stimmen für eine Partei ausgewertet werden. So würden etwa Nutzende, die mit mehreren Tweets im Datensatz vertreten sind, nicht mehrfach für die Berechnung der Stimmanteile berücksichtigt.

Während viele der Entscheidungen, die Tumasjan et al. (2010) in ihrem Forschungsdesign getroffen haben, von Jungherr et al. (2012) kritisch hinterfragt und hinsichtlich ihrer Auswirkungen auf die Resultate der Studie untersucht wurden (vgl. Übersicht der hier erläuterten Fehlerpotentiale in Tabelle 1) ist dies vermutlich nicht die größte Schwäche dieser Studie. Kritischer einzuordnen ist vielmehr die fehlende Dokumentation und Erläuterung der Prozesse der Entscheidungsfindung, da so Entscheidungen beliebig erscheinen und die Datenqualität nicht unmittelbar ersichtlich

| Phase des Forschungsprozesses | Fehler-Kategorie | Identifiziertes Fehlerpotential |
|-------------------------------|--------------------------------|---|
| Definition Konstrukt | Validität | Passt die Operationalisierung (Nennung der Partei) zuverlässig zum Konstrukt (Stimmanteil)? |
| Auswahl Plattform | Eigenschaften der Plattform | Können kurze Tweets den nötigen politischen Informationsgehalt haben? |
| Auswahl Plattform | User*innengruppe der Plattform | Sind Twitter-User*innen repräsentativ für die stimmberechtigte Bevölkerung? |
| Sammlung Daten | Auswahl der Datenpunkte | Sind alle relevanten Parteien in den Daten berücksichtigt? Welcher Zeitraum wird für die Auswertung berücksichtigt? |
| Analyse Daten | Messung der Datenpunkte | Wie werden Tweets mit verschiedenen genannten Parteien gezählt? Wie wird aggregiert? |

Tabelle 1 Übersicht über die mittels des TED-On identifizierten Fehlerpotentiale im Fallbeispiel der Bundestagswahlstudie von Tumasjan et al. (2010).

wird. Für die Arbeit mit digitalen Verhaltensdaten ist dies im Sinne der Nachvollziehbarkeit und Reproduzierbarkeit der Ergebnisse jedoch unbedingt wünschenswert.

» Menge und Verfügbarkeit digitaler Verhaltensdaten – eine Zeitenwende. «

Wie das Fallbeispiel aufzeigt, kann das TED-On dazu verwendet werden, den Prozess der kritischen Auseinandersetzung mit den eigenen Forschungsdesigns von Anfang an anzuleiten, und Forschende bei der Findung und Begründung ihrer Entscheidungen zu unterstützen.

To do: Qualitätsstandards und interdisziplinärer Austausch zu Datenqualität weiterentwickeln

Die Sozialwissenschaften erleben durch die immense Menge und Verfügbarkeit digitaler Verhaltensdaten gerade eine Zeitenwende. Dieser Beitrag versucht jedoch aufzuzeigen, dass selbst die innovativsten und umfang-

reichsten Daten unzureichend sind, wenn sie nicht von hoher Qualität sind. Diese Qualitätsdefizite können einerseits in den datengenerierenden Prozessen ihren Ursprung haben, aber auch im Umgang mit den Daten durch die Forschenden selbst. Digitale Verhaltensdaten werden immer häufiger genutzt. Gerade diese vermehrte Nutzung untermauert die Notwendigkeit zur Etablierung von Qualitätsstandards, um bestmögliche wissenschaftliche Rückschlüsse zu generieren, welche wiederum die Grundlage politischer Entscheidungen darstellen können. Die bisherigen Qualitätskonzepte (siehe für eine Übersicht zu Qualitätskonzepten Daikeler et al., 2022) und Dokumentationsstandards zu *TED-On* (z.B. TES-D von Fröhling et al., 2022), können nach den ersten Jahren des Ausprobierens insbesondere als erste Schritte gewertet werden, um Qualitätsstandards zu etablieren. Diese Standards sind, wie in unserem Beispiel erläutert, oftmals von dem Format und der Nutzung der Daten abhängig.

Die Ergänzung von Umfragedaten mit digitalen Verhaltensdaten hat neue Datenqualitätspotentiale, jedoch auch neue Herausforderungen, mit sich gebracht (Weiss & Stier, 2023, in diesem easy-Band). Datenqualitätsstandards hängen neu etablierten Datenformaten oftmals hinterher. Denken wir diese Entwicklung neuer Datenformate nun noch einen Schritt weiter, beispielsweise in die



Abb. 4 Datenerfassung von Bild-, Sprach-, Video und Sensordaten in einer virtuellen Realität, Photo by Eugene Capon, CC0 Public Domain

Richtung von *virtual* und *augmented reality*, oder auch „nur“ in die automatisierte Analyse von Bild- und Videodaten, stehen wir mittel- und langfristig vor weiteren, nie dagewesenen Forschungsmöglichkeiten (z.B. zur Untersuchung von Vertrauen wie in Miller et al., 2019) – aber auch vor neuen Herausforderungen für die Datenqualität. Insbesondere durch die Verschmelzung von Text-, Video-, Sprach-, Bild- und Sensordaten (siehe Abbildung 4 für ein Beispiel) stehen die Sozialwissenschaftler einer riesigen Datenquelle mit all ihrer potenziellen Datenverzerrungsproblematik gegenüber. Gerade vor diesem Hintergrund ist der Blick in die Qualitätsstandards von lang etablierten Disziplinen der Computer Science, der Umfrageforschung und der Sensortechnik essenziell und eine intensive Zusammenarbeit unerlässlich.

Zudem bleibt zu bedenken, dass Forschende oft noch keinen Zugang zu den neu entstehenden Datenmengen haben. Das liegt insbesondere an Datensicherheitsbedenken und kommerziellen Interessen vieler Plattformen. Aktuell ermöglichen insbesondere YouTube, Reddit und Mastodon für Forschende kostenlosen und weitreichenden Zugang zu veröffentlichten Daten, während bei vielen anderen Plattformen (z.B. Facebook, Insta-

gram, TikTok und neuerdings Twitter) der Zugang zu den Daten stark eingeschränkt wird (Bruns, 2019; Freelon, 2018). Der richtige Umgang mit dieser Problematik ist eine weitere kurz- und mittelfristige Herausforderung für die sozialwissenschaftliche Forschung.

Zuletzt wird auch das hier diskutierte Problem der Repräsentation der Stichprobe auf die Gesamtbevölkerung erhalten bleiben. So wird es trotz hoher Nutzendenzahlen in der nächsten Dekade vermutlich nicht möglich sein, Rückschlüsse für die gesamte deutsche Bevölkerung über soziale Medien zu erheben.

Literatur

- Amaya, A., Biemer, P. P. & Kinyon, D. (2020). Total error in a big data world: adapting the TSE framework to big data. *Journal of Survey Statistics and Methodology*, 8(1), 89–119. <https://doi.org/10.1093/jssam/smz056>
- Bruns, A. (2019). After the ‘APIcalypse’: Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665–668. <https://doi.org/10.1080/10584609.2018.1477506>
- Groves, R. M. & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public opinion quarterly*, 74(5), 849–879. <https://doi.org/10.1093/poq/nfq065>
- Hsieh, Y. P. & Murphy, J. (2017). Total twitter error. In P. P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, C. Tucker & B. T. West (Hg.), *Total survey error in practice* (S. 23–46). <https://doi.org/10.1002/9781119041702.ch2>
- Jungherr, A., Jürgens, P. & Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. “Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment”. *Social science computer review*, 30(2), 229–234. <https://doi.org/10.1177/0894439311404119>
- Sen, I., Flöck, F., Weller, K., Weiß, B. & Wagner, C. (2021). A total error framework for digital traces

- of human behavior on online platforms. *Public Opinion Quarterly*, 85(S1), 399–422.
<https://doi.org/10.1093/poq/nfab018>
- Daikeler, J., Sen, I., Birkenmaier, L., Froehling, L., Gummer, T., Silber, H., Lechner, C. & Weiß, B. (2022). *Assessing Data Quality in the Age of Digital Social Research: A Systematic Review*. Daikeler. ESA RN 21 Quantitative Methods Mid-Term Conference, Salamanca, Spain.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457(7232), 1012–1014.
<https://doi.org/10.1038/nature07634>
- Kim, M. & Lu, Y. (2020). Testing partisan selective exposure in a multidimensional choice context: Evidence from a conjoint experiment. *Mass Communication and Society*, 23(1), 107–127.
<https://doi.org/10.1080/15205436.2019.1636283>
- King, G. (2011). Ensuring the Data-Rich Future of the Social Sciences. *Science*, 331(6018), 719–721.
<https://doi.org/10.1126/science.1197872>
- Lazer, D., Kennedy, R., King, G. & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), 1203–1205.
<https://doi.org/10.1126/science.1248506>
- Moeller, J. & Helberger, N. (2018). *Beyond the filter bubble: Concepts, myths, evidence and issues for future debates*. <https://hdl.handle.net/11245.1/478edb9e-8296-4a84-9631-c7360d593610>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.
- Ross Arguedas, A., Robertson, C., Fletcher, R. & Nielsen, R. (2022). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism.
- Scharkow, M., Mangold, F., Stier, S. & Breuer, J. (2020). How social network sites and other online intermediaries increase exposure to news. *Proceedings of the National Academy of Sciences*, 117(6), 2761–2763.
<https://doi.org/10.1073/pnas.1918279117>
- Tumasjan, A., Sprenger, T., Sandner, P. & Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *Proceedings of the International AAAI Conference on Web and Social Media*, 4(1), 178–185.
<https://doi.org/10.1609/icwsm.v4i1.14009>
- Weiß, J. & Stier, S. (2023). Die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten. *easy_social_sciences* 68, 31–38.
<https://doi.org/10.15464/easy.2023.04>

Leon Fröhling

GESIS – Leibniz-Institut für Sozialwissenschaften

E-Mail leon.froehling@gesis.org

Leon Fröhling ist wissenschaftlicher Mitarbeiter und Doktorand in der GESIS-Abteilung Computational Social Sciences. Er forscht zu digitalen Verhaltensdaten, Datenqualität und dem Zusammenspiel von Technologie und Gesellschaft.

Lukas Birkenmaier

GESIS – Leibniz-Institut für Sozialwissenschaften

E-Mail lukas.birkenmaier@gesis.org

Lukas Birkenmaier ist wissenschaftlicher Mitarbeiter und Doktorand in der GESIS-Abteilung Survey Design and Methodology. Seine Forschungsschwerpunkte liegen im Bereich digitale Verhaltensdaten, Polarisierung und Validität.

Jessica Daikeler

GESIS Leibniz-Institut für Sozialwissenschaften

E-Mail jessica.daikeler@gesis.org

Dr. Jessica Daikeler arbeitet als Post-Doc in der GESIS-Abteilung Survey Design and Methodology. Sie forscht zu Repräsentativität von digitalen Verhaltensdaten und Umfragedaten, Bereitschaft zur Teilung von Daten, Modus- und Deviceeffekte in Befragungen.



Die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten

Julia Weiß & Sebastian Stier

Während die Sozialwissenschaften seit Jahrzehnten Daten aus Umfragen verwenden, stehen in den letzten Jahren zunehmend auch digitale Verhaltensdaten, z.B. Textdaten aus sozialen Medien oder Informationen über das Nutzungsverhalten von Smartphones, zur Verfügung. Beide Datenformen bringen dabei spezifische Potentiale und Limitationen mit sich, welche durch deren Verknüpfung adressiert werden können. Dennoch findet diese Verknüpfung bisher nur in verhältnismäßig wenigen Studien statt. Der vorliegende Beitrag erörtert, warum es sinnvoll sein kann Befragungen und digitale Verhaltensdaten miteinander zu verknüpfen, wie dadurch Forschung zu neuen Themen ermöglicht wird, welche Wege es für die Verknüpfung gibt und welche Herausforderungen damit einhergehen.

While the social sciences have been using survey data for decades, digital behavioral data, e.g., text data from social media or information about behavior on smartphones, have become increasingly available in recent years. Both data forms have specific potentials and limitations, which can be addressed by linking both with each other. Although the combination of these two forms of data opens up new potentials for social science research, this has so far only been done in comparatively few studies. This article discusses why it can make sense to link surveys and digital behavioral data, how this enables research on new topics, what ways there are for linking them, and what challenges arise.

Keywords: Data linking, Umfragedaten, digitale Verhaltensdaten

Die Digitalisierung bringt im Privaten, wie auch auf gesellschaftlicher Ebene einen Strukturwandel mit sich. Wir alle agieren tagtäglich mit soziotechnischen Systemen, beispielsweise lesen oder schreiben wir etwas in sozialen Medien, tragen eine smarte Uhr, die unseren Standort aufzeichnet, oder kaufen online etwas ein. Die dadurch entstehenden Interaktionen zwischen Menschen und digitaler Technologie können dabei soziale Strukturen und Prozesse wie z.B. die Verstärkung von Meinungsunterschieden (Polarisierung)

beeinflussen oder sogar neue Phänomene wie die kollaborative Wissensproduktion und -verbreitung über Wikipedia entstehen lassen. Ob die Digitalisierung unter dem Strich insgesamt wünschenswert oder problematisch ist, soll an anderer Stelle beurteilt werden, klar ist jedoch, dass man für eine solche Beurteilung diese Interaktionen und Einflüsse zwischen Menschen und digitaler Technologie näher betrachten muss. Aus Sicht der Sozialwissenschaften eröffnet sich hier ein großes Forschungsfeld und damit einhergehend die

Frage, welche Daten man zur Beforschung dieser Phänomene zur Verfügung hat oder neu erhalten kann.

Seit Jahrzehnten basiert ein erheblicher Teil der empirischen Sozialforschung auf Umfragedaten. Täglich finden Umfragen zu allerlei Themen statt: Sei es im Bereich der kommerziellen Markt- und Meinungsforschung, zur Wahrnehmung von Themen und Beurteilung von Produkten oder in der akademischen Forschung, etwa im Kontext von größeren Ereignissen wie etwa Wahlen und den Fragen rund um das Wahlverhalten. Die mit den resultierenden Daten stattfindende Forschung ist dabei vielfältig. Sie stößt aber insbesondere im Bereich der Beforschung des Zusammenspiels von Menschen und digitaler Welt an ihre Grenzen. Mit der immer größer werdenden Verfügbarkeit von digitalen Verhaltensdaten eröffnen sich neue Möglichkeiten zur Lösung dieser Herausforderung in der Forschung, und gleichzeitig führt der neue Datentyp zur Etablierung eines eigenständigen Forschungsfeldes. Digitale Verhaltensdaten kann man als Aufzeichnungen von Aktivitäten, die über ein digitales Informationssystem durchgeführt wurden, definieren (Howison et al., 2011). Dabei handelt es sich um eine Vielzahl von digitalen Informationssystemen, wie etwa Webseiten, Social-Media-Plattformen, Smartphone-Apps oder Sensoren (z.B. GPS oder Bluetooth). Hierbei können beispielsweise Meinungsäußerungen in Form von Textdaten auf Plattformen, das Interaktionsverhalten mit Freund*innen oder auch die Klickzahlen für Nachrichtenseiten von Interesse sein. Wie das Zusammenspiel aus Umfragedaten und digitalen Verhaltensdaten erfolgen kann, warum es überhaupt sinnvoll sein kann die beiden Datenformen miteinander zu verbinden und welche zukünftigen Wege denkbar und wünschenswert sind, werden wir im Folgenden darlegen.

Warum kann es sinnvoll sein, digitale Verhaltensdaten mit Umfragedaten zu verknüpfen?

Sowohl digitale Verhaltensdaten als auch Umfragedaten bringen jeweils Vor- und Nachteile mit sich. Die Vorteile digitaler Verhaltensdaten liegen dabei insbesondere in der Möglichkeit der aktiven und passiven Messung von „natürlichem“ menschlichen Verhalten in einem zuvor nicht dagewesenen Maßstab und Detailgrad. Digitale Verhaltensdaten können dabei in vielen Formen und aus vielen Quellen für wissenschaftliche Zwecke relevant sein. Als grobe Gruppierungsmöglichkeit hat sich die Unterscheidung zwischen „found data“, also z.B. öffentliche Kommunikation in sozialen Netzwerken, und spezifisch für wissenschaftliche Zwecke erhobenen Daten („designed data“), z.B. das Aufzeichnen des Internetnutzungsverhalten von Personen, die im Rahmen eines Projektes dem zugestimmt haben, bewährt. Gefundene digitale Verhaltensdaten („found data“) sind vielfältig, sie umfassen Texte, Bilder oder Videos, die allesamt für sozialwissenschaftliche Fragestellungen relevant sein können. Sie werden in großer Menge erzeugt und weisen oft eine genauere räumliche und zeitliche Auflösung als Umfragedaten auf, zum Beispiel bis hin zu einem sekundengenauen Zeitstempel, wann ein Inhalt erzeugt wurde. Zudem können Wissenschaftler*innen gefundene digitale Verhaltensdaten in Echtzeit erheben, da diese ständig und unabhängig von einzelnen Forschungsprojekten durch die Nutzung digitaler Angebote entstehen. Nachteilig ist jedoch, dass gefundene digitale Verhaltensdaten, z.B. die wissenschaftlich viel genutzten Twitter-Daten, meist nur begrenzt Informationen über die Produzierenden der Daten umfassen, die selten über den pseudonymisierten Twitter-Accountnamen und äußerst lückenhaft berichtete Informationen (wie z.B. dem Wohnort, Beruf oder Geschlecht in der freiwilligen Profilbeschreibung) hinaus gehen.

Eigens für wissenschaftliche Zwecke generierte Daten („designed data“) können beispielsweise Erhebungen durch eine Smartphone-App zur mobilen Befragung und Erhebung von Sensordaten sein oder aber die Erfassung der Webseiten-Besuche durch ein Browser-Plugin. Diese Arten von Datenerhebungen geben Wissenschaftler*innen mehr Transparenz und Kontrolle zur Sicherung wissenschaftlicher Standards, als wenn man beispielsweise Datensätze von Firmen erhält (z.B. Kaufverhalten auf digitalen Marktplätzen oder Suchanfragen in digitalen Suchmaschinen), die diese kommerziell erheben und kaum bis keine Informationen über das Entstehen der Datensätze preisgeben. Gleichzeitig erfordern Erhebungen von „designed data“ einen erheblich höheren Ressourceneinsatz der Wissenschaftler*innen, da sowohl mit der Erhebung solcher Daten selbst Kosten einhergehen als auch spezifische Kenntnisse zur technisch korrekten Erhebung häufig erst einmal erlernt werden müssen.

» **Verbindet man digitale Verhaltensdaten mit Umfragedaten, so können sich die jeweiligen Nachteile ausgleichen und produktive Synergien entstehen.**

Beiden Typen von digitalen Verhaltensdaten ist jedoch gemein, dass sie für sich stehend beschränkte oder gar keine Informationen über die datengenerierenden Personen enthalten. Basierend auf digitalen Verhaltensdaten ist so zum Beispiel generell keine direkte Messung von Einstellungen möglich, wir kennen also z.B. die Meinungen und Überzeugungen der Personen hinter einem Beitrag in den sozialen Medien nicht. Darüber hinaus enthalten digitale Verhaltensdaten keine Informationen zu den Offlineverhaltensweisen der Individuen. All dies mündet in der Frage, inwiefern mit der Analyse dieser Form der

Daten überhaupt Aussagen über die gesamte Bevölkerung oder zumindest bestimmte Gruppen getroffen werden können. Im Falle der für wissenschaftliche Zwecke generierten („designed“) Daten ist es möglich die Stichprobe, also eine Auswahl an Personen, die stellvertretend für eine Grundgesamtheit (z.B. die deutsche Bevölkerung oder alle Twitter-Nutzenden) ausgewählt werden, zu bestimmen, um so später Aussagen über die Gesamtbevölkerung treffen zu können. Bei den gefundenen Daten ist es dagegen nicht möglich festzulegen, wie sich eine Stichprobe zusammensetzt. So sind später auch keine Aussagen über die Grundgesamtheit möglich.

Der Vorteil von Umfragedaten liegt darin, dass sie die Erhebung verschiedener Konzepte ermöglichen, die Teilnehmende am besten selbst einschätzen können, wie z.B. persönliche Werte, Einstellungen und Überzeugungen. Es kann aber auch Offline- und Onlineverhalten erfragt werden. Darüber hinaus ist es hier möglich, eine für die Grundgesamtheit stellvertretende Auswahl an Befragten festzulegen. Nachteilig ist jedoch, dass Selbstauskünfte der Befragten zu ihrem Onlineverhalten nur bedingt valide sind (Parry et al., 2021). So kann es sein, dass Befragte sich schlicht nicht erinnern oder nur schwer einschätzen können, wie viele Stunden sie beispielsweise täglich oder auf welchen Websites sie Nachrichten konsumieren. Darüber hinaus kann soziale Erwünschtheit die Ergebnisse einer Befragung beeinflussen, wenn sich also Befragte bei ihrer Antwort an sozialen Normen und nicht an ihrer tatsächlichen Einstellung oder ihrem tatsächlichen Verhalten orientieren. Schließlich haben Umfragen mit sinkender Teilnahmebereitschaft, also mit einem zunehmend geringeren Anteil an tatsächlich realisierten Interviews an der ursprünglich ausgewählten Stichprobe, zu kämpfen.

Verbindet man digitale Verhaltensdaten mit Umfragedaten so können sich die jeweiligen Nachteile ausgleichen und produktive Synergien entstehen (Stier et al., 2020). Beispielsweise liegen durch die Umfragedaten

Informationen über die Produzierenden der Daten vor, wir wissen also etwa wie alt die befragte Person ist und welche politischen Einstellungen sie hat. Mit Hilfe der digitalen Verhaltensdaten werden detaillierte Informationen beispielsweise zum Onlinenutzungsverhalten erhoben, welche nun nicht mehr durch Schätzungen der Befragten erhoben werden müssen. Das Verbinden beider Datenformen bringt neben der Lösung zuvor vorhandener Probleme, wie etwa dass Befragte sich nun nicht mehr zu ihrem Onlineverhalten erinnern müssen, auch spezifische eigene Vorteile mit sich. So eröffnen sich für die Wissenschaft neue Forschungsfelder und für die Befragten kann sich der Aufwand reduzieren, da die Belastung durch die Teilnahme an einer Umfrage reduziert wird, wenn Informationen stattdessen als digitale Verhaltensdaten erhoben werden.

Wie lassen sich digitale Verhaltensdaten und Umfragedaten verknüpfen?

Grundsätzlich sind für die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten mindestens zwei Wege denkbar. Zum einen kann man ausgehend von einer Umfrage die Teilnehmenden einladen, ihre digitalen Verhaltensdaten zu teilen. Dieses Teilen der digitalen Verhaltensdaten kann dabei auf unterschiedliche Arten erfolgen: Die Abfrage der Profilnamen der Nutzenden in sozialen Medien ermöglicht Forschenden, im Anschluss die Daten der entsprechenden Profile in den sozialen Medien zu erheben (Al Baghal et al., 2020). Durch Studienteilnehmende installierte Apps (Keusch et al., 2022) oder Erweiterungen (Plugins) für den Webbrowser (Christner et al., 2022) erheben über einen gewissen Zeitraum hinweg in Echtzeit deren Verhalten. Schließlich besteht die Option der sogenannten Datenspenden, bei denen die Teilnehmenden einer Umfrage ihre digitalen

Verhaltensdaten selbst herunterladen und an die Forschenden übermitteln (Boeschoten et al., 2022). Zum anderen können, ausgehend von digitalen Verhaltensdaten, Teilnehmende für eine Umfrage gewonnen werden. So ist es möglich, über Werbung auf sozialen Medien oder die direkte Ansprache von Nutzenden digitaler Anwendungen Menschen für die Beantwortung einer Umfrage zu rekrutieren.

Insgesamt kann das Verknüpfen der Daten dabei in unterschiedlichen Dimensionen variieren. In Bezug auf die zeitliche Abfolge der Verknüpfung ist es sowohl möglich Umfragedaten und digitale Verhaltensdaten gemeinsam, etwa innerhalb eines Projekts, zu erheben (ex-ante Linking), als auch einen bereits bestehenden Datensatz mit einem nachträglich erhobenen Datensatz zu verbinden (ex-post Linking) (Stier et al., 2020). Darüber hinaus kann die Verknüpfung von Umfragedaten und digitalen Verhaltensdaten auf verschiedenen Ebenen stattfinden. Also sowohl auf individueller Ebene, wenn etwa eine Person an einer Umfrage teilnimmt und von genau dieser Person auch digitale Verhaltensdaten erhoben werden, als auch auf Aggregatsebene, beispielsweise durch die Verknüpfung von Durchschnittsergebnissen aus Umfragen und digitalen Verhaltensdaten über verschiedene Zeitpunkte oder geografische Regionen.

Welche Herausforderungen gehen mit der Verknüpfung einher?

Die Herausforderungen, welche mit der Verknüpfung von digitalen Verhaltensdaten und Umfragedaten einhergehen, umfassen praktische Hürden sowie rechtliche und ethische Aspekte.

Die praktischen Hürden bei der Verknüpfung können vielfältig sein (Breuer et al., 2022). Geht es um die Erhebung einer bestimmten Nutzungsart, z.B. die Smartphone-Nutzung, so können diejenigen, die kein Smartphone

» **Verknüpfung von digitalen Verhaltensdaten und Umfragedaten auf individueller Ebene erst nach Aufklärung und Einwilligung der Teilnehmenden.** «

besitzen oder von den Forschenden zur Verfügung gestellt bekommen, nicht an der Studie teilnehmen. Auf Seiten der Teilnehmenden einer Umfrage, welche gebeten werden ihre digitalen Verhaltensdaten zu teilen, können fehlende technische Kenntnisse zu Problemen führen. Vermeintlich einfache Dinge, wie das Wiederaufrufen von Zugangsdaten oder Nutzendennamen können die Erhebung von digitalen Verhaltensdaten erschweren. Dies kann dazu führen, dass bestimmte Gruppen mit dieser Erhebungsform nicht erreicht werden können. So werden beispielsweise bei Onlineumfragen Menschen ohne Internetzugang oder bei Erhebungen mittels einer Erweiterung (Plugin) für den Webbrowser, Personen mit mangelnden technischen Kenntnissen von der Teilnahme ausgeschlossen. Gleichzeitig ist es Aufgabe der Forschenden, Erhebungen möglichst so zu gestalten, dass praktische Hürden auf Seiten der Befragten minimiert werden. Dafür sind auch auf Seiten der Forschenden entsprechende technische Kenntnisse erforderlich.

Die konkrete Verknüpfung von digitalen Verhaltensdaten und Umfragedaten birgt selbst einige Herausforderungen. Die Frage nach der Teilnahme an einer Erhebung von digitalen Verhaltensdaten kann von den Befragten in einer Umfrage unbeantwortet bleiben oder sogar zum Abbruch der Befragung führen. Für die Forschenden bringt das nicht nur höhere Kosten für die Erzielung ihrer benötigten Fallzahlen mit sich, sondern auch durch potenziell verzerrte Daten und wissenschaftliche Fehlschlüsse, wenn beispielsweise online-affine Gruppen mit schwach aus-

geprägten Einstellungen zum Datenschutz in Studiensamples überrepräsentiert sind. Bisherige Forschung zeigt, dass in Panelbefragungen, wenn man also mit zeitlichem Abstand immer wieder dieselben Personen befragt, mehr Befragte bei der Frage, ob sie bereit wären ihre digitalen Verhaltensdaten zur Verfügung zu stellen, abbrechen bzw. diese Frage unbeantwortet übergehen als bei anderen Anfragen. Die Wirkung lässt jedoch in den folgenden Wellen nach. Das erscheint hinnehmbar angesichts des Mehrwerts, der durch die Verbindung von Umfrage- und Verhaltensdaten erreicht werden kann (Trappmann et al., 2022).

Aus rechtlicher und ethischer Perspektive kommen weitere Herausforderungen für die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten hinzu. So sind beispielsweise im Falle der Erhebung von Daten aus sozialen Medien stets die Nutzungsbedingungen der Plattform zu beachten. Diese unterlagen zum Teil die automatisierte Erhebung von öffentlich einsehbaren Informationen oder setzen sogar technische Hürden ein, um dies zu verhindern. Darüber hinaus muss der Datenschutz gewahrt und die seit 2018 anwendbare Datenschutzgrundverordnung (DSGVO) eingehalten werden. Dies bedeutet für die Forschenden nun umfangreicher die Einhaltung des Datenschutzes dokumentieren zu müssen, eröffnet aber gleichzeitig eine neue Form des Zugangs zu digitalen Verhaltensdaten. Konkret wahrt die DSGVO für die Nutzenden das Recht auf Auskunft, also das Recht darauf, von den Betreibenden sozialer Medien die über sie gespeicherten Daten zu erhalten (Boeschoten et al., 2022). Diese können dann im Rahmen von Datenspenden für die Forschung nutzbar gemacht werden. Generell kann die Erhebung sowie die Verknüpfung von digitalen Verhaltensdaten und Umfragedaten auf individueller Ebene erst nach Aufklärung und Einwilligung der Teilnehmenden stattfinden. Die Aufgabe der Forschenden ist es dabei, den Befragten alle nötigen Informationen zu geben, damit diese informiert entscheiden können, ob

sie einer Datenerhebung zustimmen oder nicht. Dies muss in derart strukturierter und überschaubarer Form stattfinden, dass die Befragten nicht mit unnötigen Informationen überfordert werden und somit wohlmöglich ihre Zustimmung geben, ohne die Folgen ihrer Zustimmung verstanden zu haben (Beuthner et al., 2022; Walzenbach et al., 2022).

» **Öffentliche Infrastrukturen, die Forschenden den Zugang zu digitalen Verhaltensdaten erleichtern, sind deshalb unverzichtbar.** «

Sind die Daten einmal erhoben und miteinander verbunden, ergeben sich neue Herausforderungen im Bereich der Analyse, Archivierung und Wiederverwendbarkeit des entstandenen Datensatzes. In den Sozialwissenschaften gängige statistische Analyseverfahren sind nicht ohne Weiteres auf derart vielschichtige Datensätze anwendbar (Stier et al., 2020). Stattdessen bedarf es der Verwendung und Kombination unterschiedlicher Methoden und für individuelle Forschende womöglich auch die Erlernung neuer Verfahren zur Auswertung der Daten.

Handelt es sich bei dem entstandenen Datensatz um Daten, die nicht öffentlich verfügbar sind, stellt sich die Frage wie der Datensatz an andere Forschende weitergegeben werden kann. Es ist ein wichtiges Grundprinzip einer „offenen Wissenschaft“ (*Open Science*), die Überprüfung von Forschungsergebnissen und Sekundärnutzung von Daten zu ermöglichen. Zudem geben Drittmittelgeber oder auch Verlage und Fachzeitschriften regelmäßig vor, dass Daten für die Öffentlichkeit verfügbar gemacht werden müssen (Breuer et al., 2020). Dabei entstehen Zielkonflikte mit den Nutzungsrichtlinien von Online-Plattformen und der Notwendigkeit, dem Datenschutz genüge zu tragen.

Was kann die Zukunft bringen?

Eine eigene Forschungsinfrastruktur aufzubauen, ist für Individuen und viele Forschungsgruppen technisch herausfordernd und kostenintensiv. Kommerzielle Anbieter zur Erhebung von Umfrage- und digitalen Verhaltensdaten sind teuer, arbeiten zum Teil intransparent und sind meist grundsätzlich auf die Marktforschung ausgerichtet. Für Forschende verschließt all dies häufig den Zugang zu einem möglichen Forschungsfeld. Institutionalisierte öffentliche Infrastrukturen, die Forschenden den Zugang zur Erhebung und Auswertung von digitalen Verhaltensdaten erleichtern, sind deshalb unverzichtbar.

Darüber hinaus bedarf es vielfältiger Forschung zu diesem immer noch relativ „neuen“ Datentyp. Dies beinhaltet theoretische Arbeiten, welche die außerordentliche Informationsvielfalt von verknüpften Datensätzen aufgreifen und methodische Zugriffe, mit denen sich diese Datenkombinationen bestmöglich analysieren lassen. Während die Auswertung von strukturierten Befragungsdaten wissenschaftlich standardisiert ist, stellt sich die Frage, wie man aus digitalen Daten sinnvolle Messinstrumente ableiten kann. Es ist eine erhebliche Herausforderung, Konstrukte wie Populismus oder *Hate Speech* (dt.: Hassrede) mit computergestützten Verfahren in großen Datensätzen valide zu messen. Zudem bedarf es weitergehender Forschung dazu, wie individuelle Kosten-Nutzen-Strukturen, also etwa finanzielle Abhängigkeiten der Befragten von einer monetären Entlohnung für die Teilnahme, sowie allgemeine Einstellungen von Befragten die Bereitschaft zum Teilen von digitalen Verhaltensdaten beeinflussen und wie sich dies zwischen unterschiedlichen Formen digitaler Verhaltensdaten unterscheidet (Silber et al., 2022). Daran schließen sich Fragen soziodemografischer Verzerrungen an sowie weitere experimentelle Ansätze, die

Erhebung von digitalen Verhaltensdaten für Befragte möglichst einfach zu gestalten. Teilen beispielsweise Befragte von Online-Befragungen häufiger ihre digitalen Verhaltensdaten als Befragte von Telefoninterviews und sollte dies der Fall sein, liegt das an der höheren Affinität zu digitalen Inhalten oder an der Art der Anfrage digitale Daten zu teilen?

Fazit

Die Digitalisierung verändert die individuelle und gesellschaftliche Realität. Die Interaktion von Mensch und digitaler Technologie führt dabei zu ganz real erfahrbaren Phänomenen, wie beispielsweise Hassrede in sozialen Medien. Für die Sozialwissenschaften ergeben sich hierdurch zahlreiche neue Forschungsthemen und -fragen. Gleichzeitig eröffnen digitale Verhaltensdatensätze die Beforschung etablierter Forschungsfragen und theoretischer Fragen, beispielsweise zu menschlichen Interaktionen und ihren individuellen Folgen. Die Verbindung von Befragungen und digitalen Verhaltensdaten und ihrer jeweiligen Vorteile kann die sozialwissenschaftliche Forschung erheblich voranbringen. Gleichzeitig ist zu berücksichtigen, dass eine solche Datenverbindung selbst mit verschiedenen Herausforderungen einhergeht.

Durch die Verbindung von Umfragen und digitalen Verhaltensdaten sind an der Schnittstelle zwischen Sozialwissenschaften und Informatik neue Möglichkeiten für die Wissenschaft entstanden. Zukünftige Forschung wird nicht nur die Transparenz und die praktische Umsetzbarkeit des *Linking* beider Datentypen verbessern, sondern auch verschiedene Forschungsfelder mit innovativer inhaltlicher Forschung bereichern.

Literatur

- Al Baghal, T., Sloan, L., Curtis, J., Williams, M. & Burnap, P. (2020). Linking Twitter and Survey Data: The Impact of Survey Mode and Demographics on Consent Rates Across Three UK Studies. *Social Science Computer Review*, 38(5), 517–532. <https://doi.org/10.1177/08944393198280>
- Beuthner, C., Keusch, F., Silber, H., Weiß, B. & Schröder, J. (2022). Consent to Data Linkage for Different Data Domains - The Role of Question Order, Question Wording, and Incentives. *SocArXiv*, 1–39. <https://doi.org/10.31235/osf.io/qh93g>
- Boeschoten, L., Ausloos, J., Moeller, J., Araujo, T. & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, 4(2), 388–423. <https://doi.org/10.5117/CCR2022.2.002.BOES>
- Breuer, J., Bishop, L. & Kinder-Kurlanda, K. (2020). The practical and ethical challenges in acquiring and sharing digital trace data: Negotiating public-private partnerships. *New Media & Society*, 22(11), 2058–2080. <https://doi.org/10.1177/1461444820924622>
- Breuer, J., Kmetty, Z., Haim, M. & Stier, S. (2022). User-centric approaches for collecting Facebook data in the ‘post-API age’: Experiences from two studies and recommendations for future research. *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2022.2097015>
- Christner, C., Urman, A., Adam, S. & Maier, M. (2022). Automated Tracking Approaches for Studying Online Media Use: A Critical Review and Recommendations. *Communication Methods and Measures*, 16(2), 79–95. <https://doi.org/10.1080/19312458.2021.1907841>
- Howison, J., Wiggins, A. & Crowston, K. (2011). Validity Issues in the Use of Social Network Analysis with Digital Trace Data. *Journal of the Association for Information Systems*, 12(12), 767–797. <https://doi.org/10.17705/1jais.00282>
- Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., Trappmann, M. & Eckman, S. (2022). Non-Participation in Smartphone Data Collection Using Research Apps. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(S2), S225–S245. <https://doi.org/10.1111/rssa.12827>
- Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H. & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, 5(11), 1535–1547. <https://doi.org/10.1038/s41562-021-01117-5>
- Silber, H., Breuer, J., Beuthner, C., Gummer, T., Keusch, F., Siegers, P. & Stier, S. (2022). Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behavior. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(S2), S387–S407. <https://doi.org/10.1111/rssa.12954>

Stier, S., Breuer, J., Siegers, P. & Thorson, K. (2020). Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field. *Social Science Computer Review*, 38(5), 503–516. <https://doi.org/10.1177/0894439319843669>

Trappmann, M., Haas, G.-C., Malich, S., Keusch, F., Bähr, S., Kreuter, F. & Schwarz, S. (2022). Augmenting survey data with digital trace data: Is there a threat to panel retention? *Journal of Survey Statistics and Methodology*. <https://doi.org/10.1093/jssam/smac023>

Walzenbach, S., Burton, J., Couper, M. P., Crossley, T. F. & Jäckle, A. (2022). Experiments on multiple requests for consent to data linkage in surveys. *Journal of Survey Statistics and Methodology*. <https://doi.org/10.1093/jssam/smab053>

Julia Weiß

GESIS – Leibniz-Institut für Sozialwissenschaften

E-Mail julia.weiss@gesis.org

Dr. Julia Weiß (GESIS, Mannheim) ist wissenschaftliche Mitarbeiterin bei der deutschen nationalen Wahlstudie (GLES). Ihre Forschungsinteressen liegen in den Bereichen Computational Social Science, politischer Partizipation und Arbeitsmarktforschung.

Sebastian Stier

GESIS – Leibniz-Institut für Sozialwissenschaften

E-Mail sebastian.stier@gesis.org

Dr. Sebastian Stier (GESIS, Köln) ist wissenschaftlicher Mitarbeiter und kommissarischer Abteilungsleiter der Abteilung Computational Social Science. Seine Forschungsinteressen liegen in den Bereichen Computational Social Science, digitale Medien, politische Kommunikation und vergleichende Politikwissenschaft.



Wissenschaftskommunikation in unsicheren, digitalen Zeiten

Eine Interviewstudie mit Kommunikator*innen aus der Digitalisierungsforschung

Matthias Begeat

*Der digitale Wandel verändert die Art und Weise, wie wir zusammenleben. Der Umbruch verunsichert viele Menschen und erhöht den Bedarf an wissenschaftlicher Forschung in diesem Bereich. Wie sachlich eine Gesellschaft die neuen Herausforderungen diskutiert, hängt maßgeblich davon ab, wie sich Organisationen, die zum digitalen Wandel forschen, in die Debatte einbringen. Vor diesem Hintergrund untersuchen wir in unserer Studie die Wissenschaftskommunikation dieser Organisationen. Dazu haben wir Experten*innen-Interviews mit Kommunikationsverantwortlichen geführt und ausgewertet. Die Institute kommunizieren auf vielfältige Weise und setzen stark auf digitale Kanäle. Die Kommunikator*innen sehen den Austausch mit den Bürger*innen vor allem als wechselseitigen Dialog, stoßen dabei aber teilweise auf Widerstände innerhalb ihrer Organisationen. Obwohl sich die Wissenschaftskommunikation in der Digitalisierungsforschung professionalisiert, qualifizieren sich Quereinsteigende oft erst über ihre Tätigkeit.*

Digital transformation is changing the way we live together. The radical change brings about uncertainty for many people and increases the need for scientific research in this area. How objectively a society discusses these new challenges depends on how organizations researching digital transformation engage in the debate. Against this background, we examine the science communication of these organizations in our study. For this purpose, we conducted and evaluated interviews with communication executives from these institutions. The institutes communicate in diverse ways and use digital channels to a great extent. The communicators see the interaction with citizens primarily as a two-way dialog, but, in some cases, they encounter resistance within their organizations. Although science communication in digitization research is becoming more professional, career changers often qualify themselves through their work.

Keywords: Wissenschaftskommunikation, Expert*innen-Interviews, Digitalisierungsforschung

Der digitale Wandel und der Bedarf an Wissenschaftskommunikation

Fake News, Hate Speech, Filterblasen – dies sind nur einige Begriffe, die markieren, wie sehr sich öffentliche Diskussionen und die Art,

wie sich Personen informieren und sich ihre Meinung bilden durch digitale Kommunikation verändert hat. Für viele andere Bereiche unseres Lebens gilt ähnliches. Von der Schule über das Berufs- und Familienleben bis hin zur Freizeitgestaltung merken wir überall, dass die Digitalisierung unser Zusammenleben prägt und alte Gewohnheiten herausfordert. Gerade in Phasen des Umbruchs und der Unsicherheit, wie eben im digitalen gesellschaftlichen

Wandel, schauen viele Menschen in Richtung Wissenschaft. Dass in den letzten Jahren zahlreiche Forschungsinstitute und -verbände wie zum Beispiel das Weizenbaum-Institut für die vernetzte Gesellschaft, das Bayerische Forschungsinstitut für Digitale Transformation (bidt), das Verbundforschungsprojekt digilog@bw oder das Center for Advanced Internet Studies (CAIS) gefördert wurden, die sich mit den Auswirkungen der Digitalisierung befassen, ist ein starkes Anzeichen für den gewachsenen Bedarf an Daten und Wissen.¹ Im besten Fall können die Organisationen mit ihren Forschungsergebnissen den Bürger*innen und der Politik Orientierung bieten.

Nun setzen sich aber die wenigsten Menschen hin und lesen wissenschaftliche Artikel zum Einsatz von Künstlicher Intelligenz in der Schule oder in den Personalabteilungen von Unternehmen. Vielmehr erfahren sie von Forschungsbefunden (wenn sie davon erfahren), indem sie Interviews mit Wissenschaftler*innen in (Online-)Zeitungen lesen, TV-Talkshows mit Expert*innen schauen, Podcasts zu wissenschaftlichen Themen streamen, in sozialen Medien Beiträge von wissenschaftlichen Organisationen in die Timeline gespült bekommen oder Veranstaltungen besuchen, auf denen Forschende über ihre Arbeit berichten. Dies alles lässt sich unter dem Begriff der Wissenschaftskommunikation zusammenfassen. Wissenschaftsorganisationen kommunizieren ihre Arbeit und Ergebnisse nach außen bzw. in die Gesellschaft. Dies tun sie, indem sie die oben genannten Formate oder andere Wege nutzen.

Wie *sachlich* und *evidenzbasiert* eine demokratische Gesellschaft mit drängenden Herausforderungen der Digitalisierung umgeht, hängt dabei auch davon ab, wie genau Wissenschaftsorganisationen kommunizieren, die in diesem Feld forschen. Wie bringen sich diese in die gesellschaftliche Auseinandersetzung ein? Liefern sie Daten und Befunde, die verständlich und nachvollziehbar sind? Erreichen

sie mit ihren Angeboten die Bürger*innen?

Um sich diesen Fragen zu nähern, haben wir in einer qualitativen Studie sieben Kommunikationsverantwortliche unterschiedlicher Organisationen aus dem Bereich der Digitalisierungsforschung interviewt.

Und was genau ist jetzt Wissenschaftskommunikation?

Wir kehren gleich zur Digitalisierungsforschung zurück. Zunächst werfen wir kurz einen Blick auf die Wissenschaftskommunikation im Allgemeinen. Zur Wissenschaftskommunikation zählen nach Schäfer et al. (2015) „alle Formen von auf wissenschaftliches Wissen oder wissenschaftliche Arbeit fokussierter Kommunikation, sowohl innerhalb als auch außerhalb der institutionalisierten Wissenschaft, inklusive ihrer Produktion, Inhalte, Nutzung und Wirkung“ (S. 13). Wenngleich auch Unternehmen (im Bereich der Digitalisierung z.B. Meta oder Google), Verbände (z.B. Bitkom oder eco – Verband der Internetwirtschaft), politische Parteien, Thinktanks und Nichtregierungsorganisation (z.B. Algorithm-Watch, Initiative D21, iRights oder die Stiftung Neue Verantwortung) zunehmend über wissenschaftsbezogene Themen kommunizieren (Schäfer & Fähnrich, 2020), haben wir uns in unserer Studie auf wissenschaftliche Organisationen, die sich an Zielgruppen *außerhalb* des Wissenschaftssystems richten und die von den wissenschaftlichen Akteuren und Organisationen selbst betrieben wird (Beck & Dogruel, 2016) konzentriert.

Wissenschaftsorganisationen sind wichtig, wenn es darum geht, Forschung zu kommunizieren, da sie am ehesten die personellen und finanziellen Mittel haben, um sich dauerhaft in die öffentliche Debatte einzubringen (Schäfer & Fähnrich, 2020). Nicht geschmälert werden sollen dabei der Einsatz und die durchaus wirkungsvollen Aktivitäten einzelner Forschender in der Öffentlichkeit, die losgelöst

1 Beispiele. Die Institutionen der interviewten Expert*innen werden hier anonymisiert.

von Forschungsorganisationen agieren. Eine gelingende, professionalisierte Forschungskommunikation kann die öffentliche Auseinandersetzung versachlichen. Auf die Spitze getrieben könnte man sagen, dass sie für eine Demokratie notwendig ist, die auf rationale Entscheidungen setzt (im Sinne von sachlich richtigen Handlungsoptionen) (Renn, 2017). Für die Wissenschaftskommunikation gelten im Vergleich zu anderen Formen der Kommunikation besondere Regeln und Ziele. Wissenschaftskommunikation muss im Gegensatz zur kommerziellen Kommunikation keine Produkte verkaufen, sie muss keine Wählerstimmen gewinnen wie die politische Kommunikation. Wissenschaftskommunikation zielt vielmehr „auf die öffentliche Vermittlung von Wissenschaft oder auf die gesellschaftliche Verständigung über Wissenschaft“ ab (Raupp, 2017, S. 147). Allerdings kann auch die Wissenschaftskommunikation interessengeleitet und strategisch ausgerichtet sein, zum Beispiel um Fördermittel von staatlicher oder privater Seite zu erhalten (Raupp, 2017).

Ziel der Studie war es zum einen, die Wissenschaftskommunikation im spezifischen Forschungsgebiet der sich noch entwickelnden und daher hoch dynamischen, interdisziplinären und unmittelbar gesellschaftlich relevanten Digitalisierungsforschung zu erkunden. Innerhalb der Forschung über Wissenschaftskommunikation möchte die Studie mit dem empirischen Zugriff auf die Meso-Ebene der Organisationen und auf die strategische Planung in eine Lücke stoßen, die zwischen Beiträgen klafft, die auf der Makro-Ebene die Bedeutung der Wissenschaftskommunikation verhandeln (z.B. in Zeiten der Pandemie oder des Populismus) oder die Evaluation einzelner Fallbeispiele oder Instrumente vornimmt (Raupp 2017; Schäfer & Fähnrich, 2020). Zum anderen wollten wir einen praktischen, anwendungsbezogenen Mehrwert für Wissenschaftskommunikator*innen im Sinne von best practices oder auch worst practices bieten.

Was wollten wir herausfinden?

In unserer Studie haben wir uns angesehen, wie Institute nach außen kommunizieren, die sich mit den Auswirkungen der Digitalisierung auf die Gesellschaft befassen. Uns interessierten dabei folgende Forschungsfragen:

- Welche *Kommunikationsinstrumente* verwenden die Institute, um sich in die *öffentliche Debatte* einzubringen?
- Welches *Selbstverständnis* von Wissenschaftskommunikation haben die Kommunikator*innen und Institute?
- Wie *professionell und strategisch* gehen die Institute in Bezug auf ihre Wissenschaftskommunikation vor?

Welche Instrumente nutzen die Institute?

Auf ihren Webseiten formulieren Institute, die zur digitalen Transformation forschen, ihre Ansprüche und Kommunikationsziele. Zusammengefasst lauten sie: Die Digitalisierung soll verantwortungsvoll gestaltet, dem digitalen Wandel sollen innovative Impulse verliehen und es soll problemrelevantes Wissen für Politik, Wirtschaft und Zivilgesellschaft geschaffen werden.

In den Interviews haben wir die Kommunikationsverantwortlichen gefragt, wie die Institute aus dem Bereich der Digitalisierungsforschung diese Ansprüche praktisch umsetzen: Welche Pläne verfolgen die Institute? Welche Kanäle werden für welche Inhalte genutzt? Wer sind die Zielgruppen, die erreicht werden sollen? Werden durch die thematische Nähe gerade digitale Kanäle genutzt (etwa Social Media)? Oder werden – z.B. auch durch die Problematisierung von Kommunikation in sozialen Netzwerken – diese Kanäle eher gemieden?

Welches Selbstverständnis dominiert?

Was ist mit der Frage nach dem Selbstverständnis gemeint? Wie könnten unterschiedliche Auffassungen aussehen? Forschende, die analysieren, wie wissenschaftliche Organisationen und Wissenschaftler*innen öffentlich kommunizieren, sprechen häufig von drei Phasen oder Modellen der Wissenschaftskommunikation (Broer & Hasebrink 2022, S. 236 ff.; Schäfer et al., 2015, S. 14 ff.).

(1) Abschottung

In einer ersten, historisch frühen Phase kommunizierte die Wissenschaft in erster Linie mit und an sich selbst. In Folge der Aufklärung im 18. Jahrhundert bildeten sich einzelne wissenschaftliche Disziplinen heraus, die ihre eigenen Methoden und Fachbegriffe entwickelten. Abgeschottet und distanziert zogen sich die Wissenschaftler*innen zurück, um mit Muße ihrer Forschung nachzugehen (Broer & Hasebrink, 2022; Schäfer et al., 2015, S. 14 ff.). Hier hat die häufig bemühte Metapher des Elfenbeinturms ihren Ursprung. Nur in Ausnahmefällen, wenn wissenschaftliche Expertise von Entscheidungstragenden eingeholt wurde, kommunizierten Forschende außerhalb ihrer Fachgemeinschaften (Broer & Hasebrink, 2022). Dieses Modell ist heute als Selbstverständnis in der Forschung mit Sicherheit nicht mehr dominant. Wir werden allerdings in unseren Befunden noch sehen, dass einige Denkmuster dieser Art immer noch existieren.

(2) Transfer

Nach der beschriebenen frühen Phase hat sich die Wissenschaftskommunikation in den letzten Jahren und Jahrzehnten stark gewandelt. Es etablierte sich ein Verständnis dafür, dass Forschungsergebnisse an Entscheidungstragende und die Öffentlichkeit weitergegeben werden sollten (Marcinkowski, 2014). Auch weil wissenschaftliche Forschung häufig aus Steuermitteln finanziert wird, geriet die Wissenschaft unter den Druck, ihre Leistungen auch öffentlich abzubilden. In dieser Phase galt der (Wissenschafts-)Journalismus als die

zentrale Instanz, die Forschung einem großen Publikum vermittelte. Wissenschaftsorganisationen reagierten darauf, in dem sie ihre Abteilungen für die Pressearbeit ausbauten. Eine gute Pressearbeit führte – in dieser Logik – zu Berichterstattung und so zu gesellschaftlicher Akzeptanz von Forschung. Da diese gesamte Phase der Wissenschaftskommunikation auf der Vorstellung basiert, dass Bürger*innen zu wenig über Wissenschaft wissen und daher die Akzeptanz schwindet, wird dieses Modell auch Defizitmodell genannt.

(3) Dialog

Mittlerweile hat sich eine weitere Vorstellung etabliert, die weniger auf eine einseitige Vermittlung von wissenschaftlichen Ergebnissen in Richtung der Bürger*innen setzt, sondern vielmehr einen vielfältigen Dialog mit den Menschen sucht. Als Gründe dafür werden verschiedene, sich überlappende Entwicklungen gesehen. Die Wissenschaft ist für andere Teilbereiche der Gesellschaft wie die Politik oder die Wirtschaft und auch für die Zivilgesellschaft immer wichtiger geworden, es wird sogar von einer „Verwissenschaftlichung der Gesellschaft“ gesprochen (Schäfer et al., 2015, S. 20; Weingart, 2001; Weingart, 2005). Darüber hinaus führten technologische Innovationen zu einem Medienwandel, der es Forschenden und den Wissenschaftsorganisationen ermöglicht, selbst Medieninhalte zu produzieren und sich an ein größeres Publikum zu wenden – denken Sie hier an die eingangs erwähnten Beispiele von Podcasts bis hin zu Social Media (Neuberger et al., 2021, S. 24; Schäfer, 2017).

Zwar werden die Modelle der Wissenschaftskommunikation häufig als zeitliche Abfolge beschrieben. Nichtsdestotrotz können sie als Selbstverständnis weiterhin nebeneinander existieren.

In den Interviews haben uns folgende Fragen interessiert: Welche Vorstellungen gibt es in den Instituten für Digitalisierungsforschung? Sind in diesem recht jungen Forschungsfeld ausschließlich Vorstellungen zu

finden, die auf Dialog mit den Bürger*innen setzen? Wie wichtig ist der einseitige Transfer noch? Herrscht Einigkeit im Institut oder gibt es sich widersprechende Positionen?

Wie professionalisiert und strategisch wird kommuniziert?

Wir wollten außerdem wissen, wie *professionell* die gesamte Wissenschaftskommunikation in der Digitalisierungsforschung bereits aufgestellt ist und wie *strategisch und geplant* die Institute vorgehen.

Vor diesem Hintergrund sind wir in den Interviews folgenden Merkmalen bzw. Fragen nachgegangen (Blöbaum, 2012; Mieg, 2018): Gibt es eigenständige Fachabteilungen für die Wissenschaftskommunikation? Wird den Kommunikator*innen innerhalb der Institute eine fachliche Autorität zugestanden? Arbeiten dort Personen, die spezifische Studiengänge absolviert haben?

Um zu verstehen, wie planvoll die Institute in ihrer Kommunikation agieren, wollten wir von den Befragten auch wissen, ob Konzeptpapiere existieren, die beschreiben, an welcher Stelle in der Organisation über die langfristige Ausrichtung der Wissenschaftskommunikation entschieden wird und welcher Stellenwert der Kommunikation insgesamt eingeräumt wird. Wie beschrieben interessierte uns außerdem, ob die Kommunikation auch strategisch ausgerichtet ist – also, ob z.B. bestimmte Maßnahmen gezielt eingesetzt werden, um Fördermittel einzuwerben und weniger in den Dialog mit den Bürger*innen zu treten.

Wie sind wir methodisch vorgegangen?

Insgesamt haben wir 15 Institute in Deutschland identifiziert, die sich (auch) aus sozialwissenschaftlicher Perspektive mit den gesellschaftlichen Auswirkungen der Digitalisierung befassen. Unter den Einrichtungen befanden sich drei universitäre Einrich-

tungen und 12 außeruniversitäre Institute. Wir haben aus allen Instituten versucht, Gesprächspartner*innen für unsere Studie zu gewinnen. Insgesamt haben wir zehn Zusagen für die Interviews erhalten, aus verschiedenen Gründen haben drei Personen ihre Teilnahme kurzfristig zurückgezogen. Insgesamt konnten wir im Zeitraum zwischen Oktober und Dezember 2020 sieben Interviews realisieren.

In unserer Studie haben wir mit insgesamt sieben Kommunikationsverantwortlichen von Einrichtungen aus dem Digitalisierungsforschungsbereich ausführlich gesprochen und dabei einen sogenannten qualitativen Ansatz verfolgt. Im Gegensatz zu einer quantifizierenden Vorgehensweise haben wir keine geschlossenen Fragen gestellt und Antworten in Zahlen überführt, so wie es z.B. in Fragebogenstudien üblich ist. In unseren langen Expert*innen-Interviews haben wir uns vielmehr für die Sichtweise, die Einschätzungen und die Begründungen der interviewten Personen interessiert.

Für die Gespräche haben wir einen Leitfaden erarbeitet, der alle für uns interessanten Fragen umfasst. Der Leitfaden dient dazu, die Gespräche vergleichen zu können. In allen Interviews sollen die gleichen Aspekte strukturiert thematisiert werden. Gleichzeitig sollen sich offene Gespräche entwickeln, in denen sich die Personen detailliert erklären können.

Die Gespräche dauerten bis zu 60 Minuten. Durch die Pandemie konnten wir die Gesprächspartner*innen nicht persönlich treffen. Die Interviews wurden daher über die Videokonferenzplattform Zoom geführt. Für die Analyse interessierte uns nur das gesprochene Wort. Wir haben daher (mit Einverständnis der Befragten) die Tonspur der Interviews aufgezeichnet und die Gespräche anschließend verschriftlicht. Die Mitschriften haben wir in einer computergestützten qualitativen Inhaltsanalyse mit Hilfe der Software MAXQDA ausgewertet. Dabei werden Textstellen aus den Interviews codiert, d.h. dass einzelne inhaltliche Abschnitte Kategorien zugeordnet werden. Wenn etwa die Befragten

über ihre Ausbildung gesprochen haben, wurden diese Angaben der Kategorie „Ausbildung“ zugewiesen. Mittels der Software lassen sich später alle Aussagen der Interviewten miteinander vergleichen.

Was wir herausgefunden haben: Wissenschaftskommunikation zwischen den Stühlen

Zwischen vielfältigen Kommunikationswegen und Zurückhaltung

Die Interviews zeigen, dass die Wissenschaftskommunikation – je nachdem, wie viele Mittel zur Verfügung stehen – in einem recht unterschiedlichen Umfang umgesetzt wird. Die Anzahl und auch die Vielfältigkeit der genutzten Instrumente variiert stark. Inhalte und Zielgruppen sind ebenfalls mehr oder weniger klar umrissen.

In fast allen Instituten sind Veranstaltungen mit lokaler Reichweite ein bewährtes Mittel, um mit ihren Zielgruppen in Kontakt und ins Gespräch zu kommen. Dies können bekannte Formate wie Podiumsdiskussionen, aber vielfach auch innovative Veranstaltungen sein, die die Teilnehmenden aktivieren und digital geprägt sind (genannt werden z.B. „Hackathons“ oder „Crypto-Partys“, auf denen Teilnehmende gemeinsam und fokussiert (Soft- oder Hardware)-Lösungen für Problemstellungen entwickeln oder sich über Verschlüsselungstechniken austauschen). Kleinere Institute führen dabei scheinbar weniger, kleinere und traditionellere Veranstaltungen durch. In der Pandemie haben viele Kommunikator*innen versucht, die Veranstaltungen in den digitalen Raum zu überführen. Äußerungen der Kommunikator*innen zeigen, dass es sowohl in Präsenz als auch im digitalen Raum oftmals schwierig ist, Zielgruppen zu erreichen.

„Im Grunde ist es schon von offline zu online gleichgeblieben. Offline hatten wir das so gedacht, dass es nett wäre, wenn es eine private Runde ist. Wir haben das in einer Unibibliothek stattfinden lassen, mit viel indirektem Licht und gemütlich mit Büchern usw., jeder sitzt da, wo er gerne möchte. Das war sehr nett. Die Leute fühlten sich sehr abgeholt. Zum Offline-Termin waren dann vor allem ältere Leute dabei, die sich dafür interessiert haben. Online ist das dann allerdings nicht herübergeschwappt. Die haben wir also ein Stück weit verloren. Online waren es immer so 20, die teilgenommen haben, wobei der Wissenschaftsanteil größer gewesen ist. Es war nachher also nicht mehr die Kommunikation mit Menschen aus der Stadtgesellschaft, sondern tendenziell eher interner Austausch.“

Die klassische Medien- und Pressearbeit wird von der Mehrheit der interviewten Personen weniger stark in den Fokus gerückt. Häufig beschränken sich die Institute auf die Bearbeitung von Medienanfragen oder veröffentlichten Pressemitteilungen. Nur zwei Befragte, die mit größeren Teams arbeiten, berichten von strategischer und umfangreicherer Zusammenarbeit mit Journalist*innen.

„Wir haben tatsächlich auch Kooperationen mit großen Zeitschriften [...], wo wir gemeinsame Features zu bestimmten Themen machen [...]. Das wird dann ganz konkret mit Wissenschaftsjournalisten erarbeitet. [...] Wir bieten auch selbst kleine Workshops für Journalisten an, für Datenjournalisten ganz speziell im Bereich Data Science.“

Für alle Kommunikator*innen spielen digitale Kanäle eine große Rolle. Die natürliche Nähe zu digitalen Plattformen der Digitalisierungsforschungsinstitute bringt sowohl eine Selbstverständlichkeit als auch eine reflektierte Einordnung der Nutzung mit sich. Die Frage, welche Social-Media-Kanäle wie bespielt werden, hängt stark von den vorhandenen personellen Möglichkeiten ab. Ob sie genutzt werden, ist meist keine Frage. Twitter ist dabei das soziale Netzwerk, das für die Kommunikator*innen die höchste Priorität

besitzt, darauf folgen Instagram und Facebook. Für alle Plattformen gilt, dass die Interaktion mit den Nutzenden eher gering ausfällt. Auftritte in berufsbezogenen Netzwerken wie Xing und LinkedIn gibt es bei Instituten, die über mehr Personal verfügen.

» **Bedarf an Orientierung und die Polarisierung – teilweise Chance, teilweise Risiko.** «

Ebenso verhält es sich mit Videoinhalten auf YouTube und anderen Plattformen. TikTok ist zum Zeitpunkt der Interviews noch kein Thema.

„Wir haben schlicht nicht genug Personal, um einen Facebook-Account oder Instagram-Account gut zu bespielen. Social Media braucht einige Stunden Vorbereitung, wenn man es wirklich gut macht und vor allem, wenn es dann responsiv wird, weil Facebook oder Instagram sehr viel responsiver sind als Twitter. Dann muss man das auch betreuen [...]. Da haben wir im Moment einfach nicht die Ressourcen zu. Also für uns ist Twitter das Tool an sich.“

Als thematische Besonderheiten der Digitalisierungsforschung werden die Aktualität und Dynamik der Themen, der hohe Orientierungsbedarf sowie der hohe Grad an Polarisierung genannt (z.B. bei der Meinungsbildung in sozialen Netzwerken). Häufig seien die Themen und Phänomene, mit denen sich die Wissenschaftskommunikation befasst, relativ neu für die Öffentlichkeit. Der Bedarf an Orientierung und die Polarisierung wird dabei teilweise als Chance, teilweise aber auch als Risiko für eine öffentliche Kommunikation eingestuft.

„Ich hatte mir Sorgen bei dem Thema ‚Feminismus im Netz‘ gemacht, weil es häufig von Rechten o.ä. gekapert wird, die mit solchen Themen nicht zurechtkommen. Da weiß man nie, wie sowas abläuft. Bei sowas [...] habe ich mir schon Gedanken gemacht, ebenso bei dem Thema ‚Demokratie im Netz‘, weil so Medien-Guerilla-Strategien der Neuen Rechten vorhan-

den und bekannt sind. Da sorgt man sich schon, wie Themen vereinnahmt werden könnten.“

Die Aktualität von Digitalisierungsthemen wird übereinstimmend positiv bewertet. Der Einsatz von Künstlicher Intelligenz wird von fast allen Interviewten als übergeordnetes Thema hervorgehoben, das aktuell starkes Interesse hervorruft.

In der Analyse der Instrumente zeigen sich Merkmale, die für die Wissenschaftskommunikation im Feld der Digitalisierungsforschung durchaus charakteristisch sein dürften. Unterschiede und Gemeinsamkeiten zu anderen Forschungsfeldern müssten über einen weiterführenden Vergleich geprüft werden. Zunächst erscheint bemerkenswert, dass Veranstaltungsformate (auch jenseits von Kontaktbeschränkungen in der Pandemie) häufig digitale Elemente und Innovationen aus einer digitalen Netzkultur enthalten. Eine umfassende direkte Kommunikation über eigene digitale Kanäle scheint gegenüber einer traditionellen Medien- und Pressearbeit präferiert zu werden. In Aussagen, die über die kommunizierten Themen gemacht werden, ist die hohe Dynamik und zum Teil auch die Unsicherheit und Polarisierung greifbar.

Zwischen Dialog und wissenschaftskulturellen Widerständen

In den Gesprächen mit den Kommunikationsverantwortlichen zeigt sich – bis auf wenige Ausnahmen – ein Selbstverständnis, das auf einen wahrhaftigen Dialog mit den Bürger*innen und anderen Zielgruppen ausgerichtet ist.

*Mein Team und ich haben ein „[...] breiteres Verständnis von Wissenschaftskommunikation als viele Wissenschaftler*innen. Für mich ist es eben nicht nur das geschriebene Wort, der Artikel oder der Vortrag, sondern ein großer Fokus liegt auf dem Dialogischen, [...] Der Entwicklung bzw. Generierung von Forschungsfragen mit Stakeholdern zusammen. Das ist ganz wichtig. [...] Für mich ist Wissenschaftskommunikation also vor allem der Dialog mit der Gesellschaft und das auf Augenhöhe-Kommunizieren [...]“*

„Das ist auch gut, weil es auch mehr Aufgabe ist, der Wissenschaft, die sich durch Steuergelder finanziert, das Wissen in die Gesellschaft zu tragen und in den Dialog zu treten. Und beim Dialog, glaube ich, ist schon eine Herausforderung, dass das nicht nur ein Transfer ist, also, dass man Wissen in die Gesellschaft schmeißt, sondern dass auch in wirklichen Gesprächen echtes Interesse und Dialog entsteht.“

In ihren Instituten sehen sich die Kommunikator*innen häufig als Dienstleistende und Service-Einrichtungen, die die Forschenden entlasten. Übereinstimmend wird angemerkt, dass die Kommunikation nach außen ein Tätigkeitsfeld ist, das viel Aufwand erfordert, der für Wissenschaftler*innen kaum zu stemmen sei.

„Das ist etwas, was Wissenschaftler kaum leisten können. Das ist gar nicht deren Job und denen das auch noch obendrauf zu setzen, ist vielleicht ein bisschen viel des Guten.“

Neben der Aufgabenteilung erscheint mit Blick auf das Selbstverständnis der Forschenden noch ein anderer Punkt interessant: Wenn die Befragten über die größten Herausforderungen berichten, wird übereinstimmend die Kultur im Wissenschaftssystem benannt. Viele Forschende sind, laut den Befragten, gegen eine verstärkte Wissenschaftskommunikation und sehen keine Anreize innerhalb ihrer Organisationen und im Wissenschaftsbetrieb. Es lohne sich nicht Wissenschaftskommunikation zu betreiben, wenn man sich in der akademischen Welt auf den Weg zur Professur macht.

Über manche Forschende sagen die Interviewten, dass es „teilweise auch nicht gerne [ge]sehen [ist], wenn ihre Projekte so einfach kommuniziert werden, weil sie dann so einen gewissen Knowhow-Verlust befürchten.“ Gleichzeitig betonen aber viele Kommunikator*innen, dass es auch einen anderen, neueren Typ unter den For-

schenden gibt, der vermehrt Wert darauflegt und Freude daran hat, sich in der Wissenschaftskommunikation zu engagieren.

Die Vorstellung und der Wille, in einen ernstgemeinten Dialog mit der Bevölkerung oder einzelnen Gruppen zu treten, bestimmt das Selbstverständnis der Kommunikator*innen im Feld der Digitalisierungsforschung. Diese Auffassung und die Ausrichtung auf innovative, digitale Kanäle und Veranstaltungen (siehe oben) treffen zum Teil allerdings auf Strukturen im Wissenschaftsbetrieb, die Aktivitäten in der Wissenschaftskommunikation nicht belohnen oder fördern.

Zwischen Quereinsteigenden und Profis

Bei der Professionalisierung der gesamten Wissenschaftskommunikation erscheint – zumindest zum Zeitpunkt der Interviews – noch Luft nach oben zu sein. Ebenso unterscheiden sich die Organisationen in ihrer strategischen Ausrichtung. So stehen fehlende Planungen ausdifferenzierten Strategien gegenüber.

Ein frühes Indiz dafür erhielten wir bereits bei der Rekrutierung der Expert*innen für unsere Studie: Drei von den 15 Personen begründeten ihre Absage damit, dass sie (noch) keine Expert*innen in dem Gebiet seien – und das, obwohl sie für die Kommunikationsaktivitäten der Organisationen verantwortlich waren (Abbildung 1). Fünf von sieben interviewten Personen bezeichneten sich als berufliche Quereinsteigende. Entweder waren die Kommunikator*innen früher selbst in der Forschung tätig und sind nun in die Kommunikation gewechselt, oder die Verantwortlichen

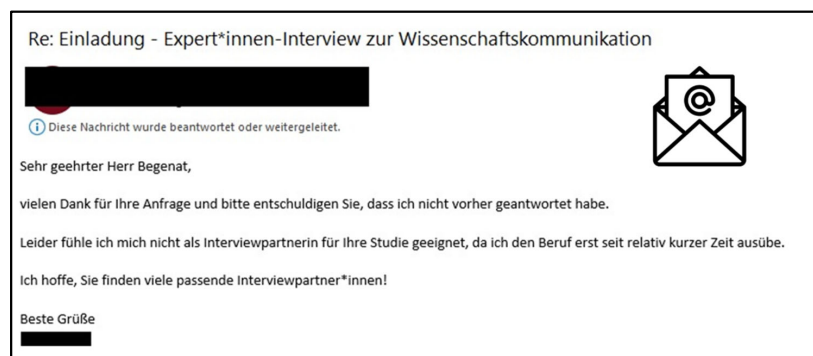


Abbildung 1 Absage einer angefragten Person mit der Begründung von mangelnder Expertise.

kamen aus anderen Kommunikationsberufen (z.B. aus dem Marketing) und haben den Sprung in die Wissenschaftskommunikation vollzogen. Unter den Interviewten war niemand, der oder die eine spezifische Ausbildungsform für Wissenschaftskommunikation absolviert hat. Viele gaben an, sich fort- und weitergebildet und/oder ein „Training on the Job“ durchlaufen zu haben.

» **Ein Personal, das sich erst langsam über die Tätigkeiten im Kommunikationsbereich professionalisiert.** ‹

Ein weiterer Hinweis dafür, dass sich die Wissenschaftskommunikation noch nicht als eigenständiger Bereich etabliert hat, zeigt der Umstand, dass vier von 15 angefragten Personen keineswegs *nur* für die Kommunikation verantwortlich sind. Diese Personen übernehmen neben der Kommunikationsarbeit auch noch weitere wichtige Aufgaben (bis hin zur Geschäftsführung).

Die Interviews verdeutlichen, dass die strategische Wissenschaftskommunikation je nach Organisationskontext sehr unterschiedlich ausgestaltet ist. Es existieren zwar professionalisierte Abteilungen, die mit umfangreichen Ressourcen und kommunikativ geschultem Personal ausgestattet sind und die langfristig und nach klar definierten Zielen planen.

„Wir versuchen nicht unseren Nachbarn und Nachbarinnen zu erklären, was wir tun, sondern wir sind eher darauf ausgelegt lukrative Partnerschaften für unsere Forscherinnen und Forscher zu finden – und die finden wir halt eher in der Wirtschaft.“

Demgegenüber stehen aber auch kaum professionalisierte Strukturen, mit wenig

Ressourcen und einem Personal, das sich erst langsam über die Tätigkeiten im Kommunikationsbereich professionalisiert. Hier erscheint konzeptionelles und strategisches Kommunizieren kaum möglich. Dies trifft vor allem auf neuere Organisationen zu.

„Gibt es etwas wie eine Strategie? Einen langfristigen Plan mit Zielen, den Sie verfolgen?“
– *„Nein, gibt es nicht. Leider nicht. Das ist eine gute Idee, das zu machen. [...] Das ist aber auch nicht so einfach.“*

Die Unterscheidung zwischen Wissenschaftskommunikation, die auf die öffentliche Vermittlung von Wissenschaft ausgerichtet ist, versus einer interessengeleiteten Kommunikation, die z.B. nach Fördermitteln strebt, ist fast allen Kommunikator*innen durchaus bewusst.

„Ja, wir betreiben Wissenschaftskommunikation, aber ich würde auch sagen, dass ein Teil unserer Arbeit auch PR für unser Institut ist. Das ist dann auch eine Art politische Kommunikation oder Imagekommunikation. Es gibt Diskussionen, ob man das nicht mehr trennen sollte.“

Während aufgrund mangelnder Ausbildungsmöglichkeiten Wissenschaftskommunikator*innen über alle Themenfelder hinweg recht häufig Quereinsteigende sein dürften, erscheint die noch nicht abgeschlossene strukturelle und strategische Professionalisierung vor allem ein Merkmal des noch jungen Alters der Digitalisierungsforschung und der betrachteten Organisationen zu sein. Interessant und unbedingt notwendig wäre es, die mittlerweile zwei Jahre alten Aussagen in neuen Interviews zu überprüfen. Es scheint plausibel, dass sich bis heute in den jungen Organisationen die Professionalisierung und strategische Planung stark weiterentwickelt hat.

Wie sind die Ergebnisse einzuordnen: Was kann die Studie und was nicht?

Mit unserer Studie wollten wir nachvollziehen, wie sich Organisationen, die zum digitalen Wandel forschen, in die öffentliche Debatte einbringen. Unsere Ergebnisse, die sich ausschließlich auf Interviews mit Personen aus der Digitalisierungsforschung stützen, gelten daher nur für das Feld der Digitalisierungsforschung. Bei der Einordnung sollte man zudem bedenken, dass die Interviews zur Veröffentlichung dieses Textes rund zwei Jahre zurückliegen und sich die Wissenschaftskommunikation weiterentwickelt haben dürfte.

Unsere Studie zeigt dennoch, dass Wissenschaftskommunikation in der Digitalisierungsforschung insgesamt besondere Merkmale aufweist. Die Auseinandersetzung mit digitalen Phänomenen überträgt sich in die Ausgestaltung der Wissenschaftskommunikation. Dies gilt ...

- für Veranstaltungsformate, die mit digitalen, partizipativen Elementen angereichert werden,
- für die starke Nutzung digitaler Verbreitungswege (z.B. Social Media, Podcasts oder Newsletter),
- für die überwiegend direkte Kommunikation, die weniger stark auf die Vermittlung durch Journalist*innen setzt,
- für Inhalte und Phänomene, die häufig als neuartig für die Öffentlichkeit und ohne eine lange Historie in der Forschung beschrieben werden,
- und für die vielfältigen fachlichen Perspektiven, die in der Forschung angelegt sind, sich aber eben auch in der Wissenschaftskommunikation zeigen.

Die Analyse der Interviews macht nachvollziehbar, wie schwierig es in der praktischen Umsetzung ist, die ausgerufenen Kommunikationsziele zu verfolgen. Die Kommunikator*innen stehen vor vielfältigen Herausforderungen. Neben Forschenden, die sich – meist intrinsisch motiviert – auch an Zielgruppen außerhalb der Wissenschaft richten, treffen sie zumindest teilweise auf Forschende, die aus verschiedenen Gründen Wissenschaftskommunikation reserviert gegenüberstehen. Darüber hinaus sind die Ressourcen der Kommunikationseinheiten z.T. kaum ausreichend, zuweilen existieren sie noch nicht einmal als eigenständige Einheit. Dies alles erschwert es, Bürger*innen und andere Gruppen außerhalb der Wissenschaft zu erreichen.

Dennoch kommunizieren die Institute für Digitalisierungsforschung ausgiebig über ihre Forschung und ihre Ergebnisse. Sie erstellen vielfältige, verständlich aufbereitete Angebote, mit denen sich Bürger*innen etwa über den Einsatz von Künstlicher Intelligenz in unterschiedlichen Bereichen informieren und Orientierung verschaffen können. Sie bieten damit häufig genau die Inhalte und Expertise, die es braucht, um die öffentliche Debatte *sachlicher* zu gestalten. In den zu Beginn skizzierten unsicheren, digitalen Zeiten ist dieser Beitrag als sehr wertvoll einzustufen. Dass sich Akteure aus Wirtschaft und Politik in der öffentlichen Auseinandersetzung laut und reichweitenstark äußern, sollte die Wissenschaftskommunikation und die Digitalisierungsforschung insgesamt nicht entmutigen, sondern vielmehr anspornen, sich weiter zu professionalisieren, Strukturen in diesem Bereich zu stärken und ihre Kommunikationsaktivitäten auch strategisch zu planen – ohne sich allerdings von dem Ziel zu verabschieden, evidenzbasiert zu kommunizieren.

Ich danke Laura Kill und Kevin Bremken für ihre Unterstützung bei der Datenerhebung und -auswertung.

Literatur

- Beck, K. & Dogruel, L. (2016). Ökonomische Perspektiven des Wissenschaftsjournalismus und der Wissenschaftskommunikation - Expertise für die Akademien-Arbeitsgruppe.
- Blöbaum, B., Nölleke, D. & Scheu, A. M. (2016). Das Experteninterview in der Kommunikationswissenschaft. In S. Averbeck-Lietz & M. Meyen (Hg.), *Handbuch nicht standardisierte Methoden in der Kommunikationswissenschaft* (S. 175–190). Springer VS. https://doi.org/10.1007/978-3-658-01656-2_11
- Blöbaum, B. (2012). Professionalisierung. In G. Bentele, H.-B. Brosius & O. Jarren (Hg.), *Lexikon Kommunikations- und Medienwissenschaft* (S. 278). Springer VS. https://doi.org/10.1007/978-3-531-93431-0_1
- Broer, I. & Hasebrink, U. (2022). Wissenschaftskommunikation als kommunikative Figuration. Ein konzeptioneller Rahmen für die empirische Untersuchung von Domänen der Wissenschaftskommunikation. In *Medien & Kommunikationswissenschaft*, 70(3), 234–255. <https://doi.org/10.5771/1615-634X-2022-3-234>
- Marcinkowski, F. (2014). Mediatisation of politics. Reflections on the state of the concept. *Javnost – The Public*, 21(2), 5–22. <https://doi.org/10.1080/13183222.2014.11009142>
- Mayring, P. (2020). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Hg.), *Handbuch Qualitative Forschung in der Psychologie, Band 2: Designs und Verfahren* (2. Aufl., S. 495–511). Springer VS. https://doi.org/10.1007/978-3-658-26887-9_52
- Mieg, H. (2018). Professionalisierung. In F. Rauner & P. Grollmann (Hg.), *Handbuch Berufsbildungsforschung* (3. Aufl., S. 452–361). UTB.
- Neuberger, C., Weingart, P., Fähnrich, B., Fecher, B., Schäfer, M. S., Schmid-Petri, H. & Wagner, G. G. (2021). *Der digitale Wandel der Wissenschaftskommunikation*. Berlin-Brandenburgische Akademie der Wissenschaften.
- Raupp, J. (2017). Strategische Wissenschaftskommunikation. In H. Bonfadelli, B. Fähnrich, C. Lüthje, J. Milde, M. Rhomberg & M. S. Schäfer (Hg.), *Forschungsfeld Wissenschaftskommunikation* (S. 143–163). Springer VS. https://doi.org/10.1007/978-3-658-12898-2_8
- Renn, O. (2017). Kommunikation zwischen Wissenschaft und Politik. In H. Bonfadelli, B. Fähnrich, C. Lüthje, J. Milde, M. Rhomberg & M. S. Schäfer (Hg.), *Forschungsfeld Wissenschaftskommunikation* (S. 183–205). Springer VS. https://doi.org/10.1007/978-3-658-12898-2_10
- Schäfer, M. S. (2017). How Changing Media Structures are Affecting Science News Coverage. In K. H. Jamieson, D. Kahan & D. Scheufele (Hg.), *Oxford Handbook on the Science of Science Communication*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190497620.013.5>
- Schäfer, M. S. & Fähnrich, B. (2020). Communicating science in organizational contexts: toward an “organizational turn” in science communication research. *Journal of Communication Management*, 24(3), S. 137–154. <https://doi.org/10.1108/JCOM-04-2020-0034>
- Schäfer, M. S., Kristiansen, S. & Bonfadelli, H. (2015). Wissenschaftskommunikation im Wandel: Relevanz, Entwicklung und Herausforderungen des Forschungsfeldes. In M. S. Schäfer, S. Kristiansen & H. Bonfadelli (Hg.), *Wissenschaftskommunikation im Wandel* (S. 10–42). Herbert von Halem.
- Weingart, P. (2001). *Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft*. Velbrück.
- Weingart, P. (2005). *Die Wissenschaft der Öffentlichkeit. Essays zum Verhältnis von Wissenschaft, Medien und Öffentlichkeit*. Velbrück.

Matthias Begenat

Center for Advanced Internet Studies (CAIS)

E-Mail matthias.begenat@cais-research.de

Dr. Matthias Begenat leitet die Wissenschaftskommunikation am Center for Advanced Internet Studies (CAIS). Er gestaltet Formate, in denen Forschung anschaulich präsentiert wird. Er forscht zur Wissenschaftskommunikation und zu politischen Milieus.



Published by

GESIS – Leibniz Institute for the Social Sciences
Knowledge Exchange & Outreach (KEO)
Unter Sachsenhausen 6-8
50667 Cologne
easy@gesis.org ▪ www.gesis.org/easy

easy Editors

Dr. Johannes Breuer, Dr. Philip Jost Janßen,
Dr. Lydia Repke, Dr. Sophie Zervos

Editorial Office

Dr. Philip Jost Janßen (Team Publications)
Dr. Sophie Zervos (Team Transfer)

Layout

Bettina Zacharias
© Foto Cover, Adobe Stock

GESIS is member of the Leibniz Association

ISSN 2749-2850 (Online)
