

BEAUFTRAGT VOM



Bundesministerium
für Bildung
und Forschung

DIE ZUKUNFT DES VERTRAUENS IN DIGITALEN WELTEN

Studie (Kurzfassung)



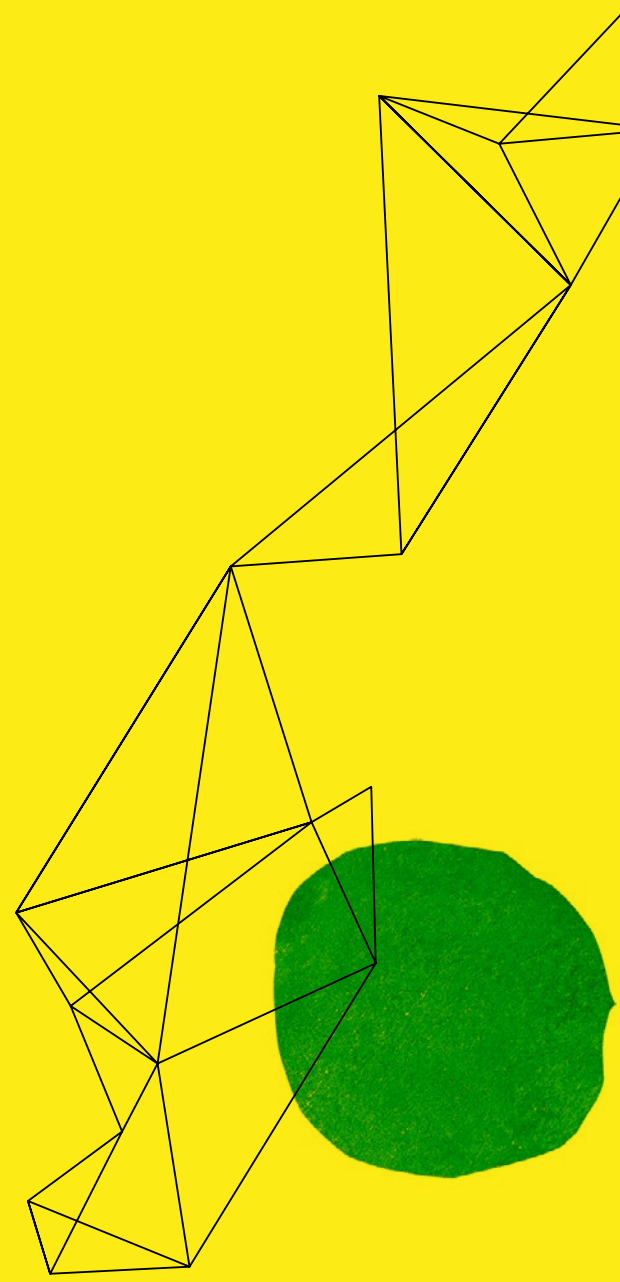
Autorinnen und Autoren:

Anna Hornik
Dr. Christian Grünwald
Daniel Bonin
Jan Reichert
Marie-Kristin Komendzinski
Julian Sachs
Holger Glockner
Michael Astor

Mai 2022

Für den Inhalt zeichnen die externen Autorinnen und Autoren verantwortlich. Das Bundesministerium für Bildung und Forschung (BMBF) übernimmt keine Gewähr für die Richtigkeit, die Genauigkeit und die Vollständigkeit der Angaben. Die in der Veröffentlichung geäußerten Ansichten und Meinungen müssen nicht mit der Meinung des BMBF übereinstimmen. Zudem sind die in der Studie skizzierten Szenarien nicht als Prognosen zu verstehen und stellen nicht unbedingt wünschenswerte Zukunftsbilder für die Bundesregierung oder das Zukunftsbüro des BMBF dar.

Diese Publikation entstand im Rahmen des Dienstleistungsauftrags „Zukunftsbüro des Foresight-Prozesses (Foresight III)“ der Prognos AG und der Z_punkt GmbH im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF), Referat – Strategische Vorausschau; Partizipation; Soziale Innovationen.



Inhaltsverzeichnis

Das Wichtigste in Kürze	2
1 WARUM BRAUCHT ES VERTRAUEN IM DIGITALEN?	4
1.1 Was ist Vertrauen?	12
1.2 Vertrauen in Technik	15
1.3 In welchen digitalen Welten vertrauen wir morgen?	17
1.4 Vertrauen, Risikoabschätzung und Sicherheit in den digitalen Welten von morgen	20
2 KERNEBESUNDE UND WEITERFÜHRENDE FRAGEN	24
2.1 Kernbefunde	24
2.2 Weiterführende Fragen	27
3 METHODIK	28
4 QUELLENVERZEICHNIS	30
Ansprechpartnerinnen und Ansprechpartner / Impressum	33

Abbildungsverzeichnis

Abbildung 1: Treiberlandkarte nach STEEP-Sektoren geordnet	11
Abbildung 2: Faktoren der Vertrauensbildung	14
Abbildung 3: Grundlegende technologische Treiber des Web 3.0 (Spatial Web)	18
Abbildung 4: Die drei Ebenen des Spatial Web	19
Abbildung 5: Methodische Schritte bei der Erstellung der Studie	29

Das Wichtigste in Kürze

Ziel und Aufbau der Studie

Die Studie „Die Zukunft des Vertrauens in digitalen Welten“ beleuchtet, welche Bedeutung Vertrauen für soziale Interaktionen in einer zunehmend digitalisierten Welt haben kann und wie sich darauf gestützte Handlungen im Digitalen wandeln könnten. Dazu nähert sie sich zunächst dem Phänomen „Vertrauen“ an und untersucht, inwieweit bekannte Mechanismen der Vertrauensbildung ins Digitale übertragbar sind. Parallel werden verschiedene Entwicklungspfade einer digitalen Sphäre von morgen beschrieben. Sechs Future Spotlights illustrieren die unterschiedlichen Ausprägungen in möglichen Alltagssituationen.

Das vorliegende Essay bildet die Kurzfassung der Studie und ermöglicht mit zwei von sechs Future Spotlights einen Einblick in die Zukunft des Vertrauens in digitalen Welten. Die Langfassung der Studie mit den übrigen Future Spotlights, weiteren Details zu den Ergebnissen sowie zur verwendeten Methodik findet sich zum Download auf vorausschau.de.

Warum braucht es Vertrauen im Digitalen?

Vertrauen ermöglicht erst das Miteinander in komplexen sozialen Zusammenhängen. Vertrauen erlaubt trotz mangelnder Kontrolle oder Vorhersehbarkeit einer Situation, Entscheidungen über die Interaktion mit anderen zu treffen und damit handlungsfähig zu sein.

Drei Faktoren tragen insbesondere dazu bei, dass Vertrauen im Digitalen anders als im direkten Miteinander gedacht werden muss:

- Der *rasche technologische Wandel* und der Einsatz neuer technologischer Errungenschaften erschweren es dem Einzelnen, belastbare Erfahrungswerte zu neuen Formen der Interaktion zu sammeln. Dadurch fällt es auch schwerer, Heuristiken zu bilden, auf Grundlage derer die Vertrauenswürdigkeit eines Gegenübers eingeschätzt werden kann. Bewertungsmaßstäbe darüber, ob in einer Situation vertraut werden kann, müssen immer wieder geprüft oder an veränderte Rahmenbedingungen angepasst werden.
- Interaktion im Digitalen findet immer mit *Technologie als Mittlerinstanz* statt. Dadurch werden Situationen komplexer: Es muss nicht nur die Vertrauenswürdigkeit des Interaktionspartners beurteilt werden, auch inwieweit die Betreiben-

den der genutzten technologischen Infrastruktur vertrauenswürdig sind, ist für die Bewertung der Gesamtsituation relevant.

- Ein stabiles Rechts- und Sanktionssystem schafft in einer Reihe von Situationen Sicherheit, die sonst allein auf Vertrauen basieren würden. Aufgrund des kontinuierlichen Wandels im digitalen Raum und seiner transnationalen Natur müsste die *rechtliche Regulierung* auf einer globalen Ebene stattfinden, wofür bislang sowohl Instanzen der Entscheidungsfindung und Koordination als auch der Durchsetzung des Rechts fehlen.

Was ist Vertrauen?

Vertrauen wird üblicherweise als positive Erwartungshaltung in das künftige Handeln eines Gegenübers definiert. Ob vertraut wird, hängt auf der einen Seite von persönlichen Erfahrungen, der Gesamteinschätzung der Situation sowie der individuellen Neigung, Vertrauen zu schenken, ab. Auf der anderen Seite ist die Wahrnehmung des Gegenübers entscheidend: Wichtige Faktoren sind hierbei, für wie kompetent, integer und wohlwollend das Gegenüber eingeschätzt wird.

Vertrauen in Technik

Nach dem heutigen wissenschaftlichen Verständnis von Vertrauen beruhen Vertrauensentscheidungen immer auch auf den vermuteten Intentionen des Gegenübers. Weil Technik selbst keinen Willen hat, kann man ihr selbst weder ver- noch misstrauen – wohl jedoch diejenigen, die ein technisches Gerät konstruiert oder Software entwickelt haben.

Zwei Entwicklungen könnten die bisherige Abgrenzung von Vertrauen zu verwandten Konzepten (wie z. B. sich auf etwas Verlassen) künftig aufweichen:

- Immer menschenähnlichere Technik – in Sprache, Aussehen und Haptik – könnte für eine Einzelperson ununterscheidbar machen, ob sie mit Menschen oder mit einer Software interagiert. Aus heutiger Perspektive bleibt streitbar, ob man vertraut, wenn man einem Gegenüber einen Willen zuschreibt, das keinen hat.
- Assistenzsysteme unterstützen menschliche Entscheidungen mittlerweile in vielen Situationen. Einige KI-Algorithmen, die hierfür genutzt werden (können), „lernen“, selbst Entscheidungen zu treffen und sind auch für die Programmierenden nicht – oder nur eingeschränkt – nachvollziehbar. Auch hier stellt sich die Frage, ob man in diesen Situationen KI „vertraut“.

In welchen digitalen Welten vertrauen wir morgen?

Die Studie geht davon aus, dass das Zusammenwirken von Technologien wie KI, intelligenten Sensoren, AR, VR und Ähnlichem die Grenzen von physischer und Online-Welt weitestgehend verschwimmen lassen wird. Ein räumliches Web – auch Spatial Web, Web 3.0¹ oder Metaversum genannt – entsteht.² In einem räumlichen Web entstehen neue Interaktionsformen – wenn beispielsweise Avatare miteinander kommunizieren –, die gegebenenfalls auch neue Muster und Formen der Vertrauensbildung hervorbringen.

1 Der Begriff Web 3.0 ist hier vom Begriff Web3, wie er seit 2021 verstärkt in der Diskussion ist, zu unterscheiden. Für eine genauere Begriffsdefinition siehe Kapitel 1, S. 12.

2 Die konkrete Ausgestaltung eines räumlichen Webs, wie sie Mark Zuckerberg mit dem Metaversum vorantreibt, ist nur eine mögliche Form, die die künftige digital-analoge Welt annehmen könnte.

Vertrauen, Risikoabschätzung und Sicherheit in den digitalen Welten von morgen

Ob und inwieweit vertraut wird, ist unter anderem dadurch bestimmt, wie groß das individuelle Risiko eingeschätzt wird, dass das geschenkte Vertrauen missbraucht wird. Diese Bewertung ist im digitalen Raum häufig verzerrt. Je nach Wissensstand werden Bedrohungen über- oder unterschätzt. In künftigen digitalen Welten könnten Risiken jedoch als realer wahrgenommen werden – auch weil sie durch ihre räumlichen Darstellungsformen stärker in der Offline-Welt wirken. Darüber hinaus hat sich die Entscheidung über die Nutzung eines Dienstes und die Vertrauenswürdigkeit der Betreibenden teilweise entkoppelt. Online-Dienste werden genutzt, obwohl den Betreibenden nicht vertraut wird.

- Netzwerkeffekte können dazu führen, dass Vorbehalte ausgeblendet werden: Überschreitet die Zahl der Nutzerinnen und Nutzer eine kritische Masse, scheint der Nutzen die Risiken zu überwiegen.
- Digitale Infrastrukturen ersetzen zunehmend ihre Vorläufer; dadurch wird es immer schwieriger, sich diesen zu entziehen, unabhängig davon, welche Risiken mit ihnen assoziiert werden.

Durch die Digitalisierung stellt sich die Frage von Vertrauen in Kommunikation und Interaktion neu. Neue Fähigkeiten, Kulturtechniken oder Instanzen der Verifizierung von Inhalten, aber auch der Echtheit und damit Vertrauenswürdigkeit unseres Gegenübers werden benötigt.



**1**

WARUM BRAUCHT ES VERTRAUEN IM DIGITALEN?

Ohne jegliches Vertrauen könnte der Mensch morgens sein Bett nicht verlassen. Mit diesem Gedanken leitet der Soziologe Niklas Luhmann seine bekannte Abhandlung über Vertrauen ein.³ Die Idee hinter dieser Feststellung: Das soziale Miteinander einer Gesellschaft ist zu komplex, um es in seiner Gänze begreifen, geschweige denn kontrollieren zu können. Niemand ist in der Lage, stets zu prüfen, ob sein soziales Umfeld ihn hintergeht, ob Häuser so gebaut wurden, dass sie einem nicht über dem Kopf zusammenstürzen, und ob die Mitmenschen sich an allgemeine Spielregeln und Gesetze halten, sodass man nicht fürchten muss, auf offener Straße ausgeraubt zu werden. Mit Vertrauen kann die Komplexität der Umwelt und des sozialen Miteinanders auf ein annehmbares Maß reduziert werden, sodass man trotz vieler Unsicherheiten handlungs- und lebensfähig bleibt. Vertrauen eröffnet Handlungsoptionen, ohne dabei in Kontrollressourcen investieren zu müssen. In der Ökonomie wird Vertrauen daher

gern als „Schmiermittel“ gesehen, das z. B. Kaufentscheidungen ermöglicht, ohne dass die Qualität eines Produkts im Detail überprüft werden kann.⁴ Auf gesamtgesellschaftlicher Ebene gibt es eine Reihe an Phänomenen, die nahelegen: Vertrauen ist gut für die Gesellschaft als Ganzes. In Staaten, in denen mehr vertraut wird, ist auch das Wirtschaftswachstum höher, es gibt mehr Innovationen, die Menschen sind gesünder und gebildeter.⁵

Auch jenseits wirtschaftswissenschaftlicher Befunde sorgt Vertrauen im Kleinen für soziale Netzwerke und Vertrauen im Größeren für gesellschaftlichen Zusammenhalt – und ist damit gestaltend für das soziale Miteinander. Jutta Allmendinger kommt im Rahmen der Vermächtnisstudie zu dem Schluss, dass der Zusammenhalt in der Gesellschaft bröckelt. Mit ihrem Kollegen Jan Wetzel führt sie den

3 Luhmann, N. (2014).

4 Speck, U. (2020).

5 Zak, P. J.; Knack, S. (2001), Beugelsdijk, S.; Smulders, S. (2009), Cardenas, J. C.; Carpenter, J. (2008), Dincer, O. C.; Uslaner, E. M. (2010).

schwindenden gesellschaftlichen Zusammenhalt auf mangelndes Vertrauen zurück: „In vielen Bereichen ist unsere Fähigkeit, mit großen Veränderungen umzugehen und sie zu gestalten, davon abhängig, ob wir anderen vertrauen.“⁶ Eine dieser großen Veränderungen ist schon länger im Gange – sie hat unser Miteinander bereits geprägt und wird es auch in Zukunft tun: die digitale Transformation, die Vernetzung von Daten, Menschen und Dingen. Das Verschmelzen von dem, was wir als „analoge Welt“ kennen und von dem, was wir heute als digital oder virtuell davon abgrenzen. Wie können wir mit dieser großen Veränderung umgehen? Ist Vertrauen nur Voraussetzung dafür, mit Veränderung umgehen zu können – oder wird Vertrauen selbst zum Gegenstand von Veränderung?

Die Omnipräsenz des Internets und das Verschmelzen technischer Systeme mit der „realen“ Welt verändern das soziale Miteinander nachhaltig. Spätestens durch die weltweite Verbreitung von Smartphones und sozialen Medien wuchsen auch die Bedeutung der digitalen Sphäre und die Zahl von Nutzerinnen und Nutzern so stark an, dass es kaum noch möglich war, sich digitalen Infrastrukturen zu entziehen. Die digitale Sphäre wurde zu einem festen Bestandteil des Alltags. Aus temporärem „Dial-on-Demand“ wurde „Always-on“. Wer nicht drin ist, ist raus aus vielem. Wirtschaft, Gesellschaft und Politik richten sich immer mehr an digitalen Infrastrukturen aus. Es stellt sich daher die Frage, ob und wie sich Vertrauen in diesem neuen digitalen Miteinander verändert. Greifen weiterhin die bekannten Heuristiken, die man im analogen Miteinander nutzt, um zu entscheiden, ob man seinem Gegenüber trauen kann, oder müssen wir neue Strategien entwickeln? Dabei gehen technologische Transformationen stets mit sozialen Latenzen einher – oder um es mit den Worten von Clay Shirky zu sagen: „Kommunikationsmittel werden erst dann sozial interessant, wenn sie technologisch langweilig werden.“⁷ Die sozialen Auswirkungen technologischer Innovationen spüren wir also immer erst zeitverzögert, wenn sich die Technologie in der Breitenanwendung befindet. Häufig bedienen neue Technologien bereits im Gang befindliche gesellschaftliche Veränderungsprozesse und verstärken diese⁸ – so wurde etwa der Wandel von der Industrie- zur Wissensgesellschaft durch die Digitalisierung beschleunigt und verstärkt.

Das Internet ist längst keine technologische Neuerung mehr, viele soziale Auswirkungen dieser weitreichenden Innovation manifestieren sich jedoch erst in jüngster Zeit. Dabei weist der Cyberspace von Beginn an zwei grundlegende Unterschiede zum analogen / physischen Raum auf, die wiederum maßgeblich Prozesse der Vertrauensbildung in der digitalen Sphäre prägen(ten).

Der *erste Unterschied* besteht darin, dass das Netz und viele der daraus erwachsenden Technologien zunächst weitgehend unreguliert waren und teilweise weiterhin sind. Eine Regulierung war zunächst deshalb schwierig, weil bestehende Instrumente nur bedingt auf das Internet angewendet werden konnten. Vieles im Cyberspace war schlichtweg juristisches Neuland. Hinzu kommt nach wie vor die Herausforderung, global grenzüberschreitende Vernetzungstechnologie durch Nationalstaaten zu regulieren, deren Regulierungskompetenzen eben weitestgehend an den Grenzen des Nationalstaats enden.

Der *zweite Unterschied* in der Frühphase der digitalen Sphäre resultierte daraus, dass wenige „Netz-Wissende“ einer großen Mehrheit von „Netz-Laien“ gegenüberstanden, welche die Funktionsweisen der Technologien nicht verstanden – und dies bis heute nicht tun.

Aus dieser Wissensasymmetrie ergibt sich auch eine Ungleichverteilung des Vertrauens. Wo sich neue unbekannte Räume öffnen, gibt es weder individuelle noch kollektive Erfahrungswerte. Eine Reziprozität des Vertrauens – sprich: je besser Menschen Vertrauensobjekte und -subjekte kennen, desto eher können sie das Verhalten einschätzen – ist in derartigen Räumen nicht gegeben. In neuen Umfeldern werden Erfahrungswerte durch Ausprobieren und natürlich auch Fehler gebildet. Vertrauen wirkt in diesem Kontext im Sinne Niklas Luhmanns als Mittel der Reduktion sozialer Komplexität ungleich stärker, weil nur durch einen enormen Vertrauensvorschuss eine Bewegung in einer nicht (be)greifbaren Sphäre überhaupt erst möglich wird.⁹

In diesem kollektiven Lernprozess verfestigen sich Muster – auch solche der Vertrauensbildung. Dabei scheint inzwischen Konsens zu sein, dass sich die netzutopistischen Träume der Internet-Pioniere nicht

6 Allmendinger, J.; Wetzels, J. (2020), S. 111.

7 Shirky, C. (2008), S. 105.

8 Stalder, F. (2016), S. 21 f.

9 Luhmann, N. (2014).

erfüllt haben, auch wenn diese versuchen, den Traum in realpolitischen Dimensionen weiterzuträumen.¹⁰ Anders gesagt: Der Glaube in die Selbstregulierungskraft der Netzwelt ist trotz vieler positiver Effekte der digitalen Welt nachhaltig erschüttert. Den Tech-Konzernen wird wenig Vertrauen entgegengebracht – weltweit vertraut über ein Viertel der Konsumentinnen und Konsumenten den Digitalgiganten wenig oder gar nicht.¹¹ Neue Phänomene wie Plattform-Ökonomie, Bots, Cyberkriminalität, Fake News oder auch Shitstorms haben die öffentliche Wahrnehmung der digitalen Sphäre verändert. Rufe nach umfassender Regulierung, insbesondere der sozialen Medien, werden von Jahr zu Jahr lauter.¹² So gibt es Stimmen, welche die Ansicht vertreten, dass ein offenes Netz langfristig nur durch stärkere Regulierung zu erhalten sei.¹³ Dynamische Entwicklungen bei Technologien – wie Augmented, Mixed oder Virtual Reality, Blockchain oder auch Hologramme, die steigende Nutzung von virtuellen Assistenzsystemen, Fortschritte beim Machine-Learning sowie am fernen Horizont sichtbar werdende Game-Changer wie Quantencomputer¹⁴ – lassen vermuten, dass die digitale Sphäre auch in den kommenden zehn bis fünfzehn Jahren von einer hohen Entwicklungsdynamik geprägt sein wird.

Wie also könnte in Zukunft Vertrauen in der digitalen Sphäre aussehen? Diese Frage kann nur zweistufig beantwortet werden. Im ersten Schritt gilt es herauszuarbeiten, wie und auf welchen Ebenen sich Vertrauen bildet. Dies muss dann in einem zweiten Schritt mit den Erkenntnissen, wohin sich die digitale Sphäre entwickeln könnte, abgeglichen werden. Konkret: Wie könnte ein Web 3.0 aussehen? Welche Technologien und Konvergenzen könnten es prägen? Wie sähen mögliche zwischenmenschliche Interaktionen und Interaktionen zwischen Menschen und Technik in dieser weiterentwickelten digitalen oder auch hybriden Sphäre aus?

Dabei sollte an dieser Stelle im Sinne einer Begriffsklärung darauf verwiesen werden, dass der Begriff „Web 3.0“ im Rahmen dieser Studie für eine Weiterentwicklung des Internets zu einem ubiquitären, von räumlichen 3D-Komponenten geprägten und semantischen Web mit hochleistungsfähiger

Datenverknüpfung durch KI-Systeme steht. Ein alternativer Begriff hierfür ist das „Spatial Web“, der sich jedoch kaum in den Debatten wiederfindet. Das „Web 3.0“ ist vom Begriff des „Web3“ abzugrenzen, der die Weiterentwicklung des Internets viel enger fasst und ein dezentralisiertes Web auf Basis von Distributed Ledger Technologien, Non-Fungible Tokens (NFTs) und Kryptowährungen anstrebt. Häufig werden jedoch in den Medien diese ähnlich klingenden Begrifflichkeiten miteinander vermischt.¹⁵

Um Wechselwirkungen sichtbar zu machen, sollten die alternative Zukünfte des Digitalen in Gesellschaftsszenarien eingebettet werden. Daher werden in der vorliegenden Untersuchung mögliche Zukünfte der digitalen Welt entlang der Rahmenbedingungen der Szenarien aus der Studie „Zukunft der Wertvorstellungen der Menschen in unserem Land“ gebildet.¹⁶ Um diese abstrakten Zukünfte greifbar zu machen, wurden die Szenarien sowohl als „Alltagsgeschichten“ einer Future Persona (Lana, 38, Bloggerin aus Deutschland) erlebbar gemacht als auch durch „Zukunftsartefakte“ angereichert.¹⁷

Die Studie „Die Zukunft des Vertrauens in digitalen Welten“ schaut bis zu 15 Jahre voraus in die 2030er-Jahre. Der hier vorliegende Essay fasst die wichtigsten Gedanken, Erkenntnisse und Befunde der Gesamtstudie komprimiert zusammen.¹⁸ Dabei stützen wir uns auf die These, dass die digitale Sphäre in 15 Jahren so wenig an die digitale Sphäre von heute erinnern wird, wie die digitale Sphäre von heute an die digitale Sphäre vor 15 Jahren erinnert. Das bedeutet, dass sich auch Vertrauensbildungsprozesse redefinieren. Es gilt also, den Cyberspace und mögliche Konvergenzen der maßgeblich zugrundeliegenden Technologien vorausschauend in alternativen Zukunftsbildern weiterzudenken, um daraus Ableitungen für die zukünftige Vertrauensbildung treffen zu können.

Wichtig ist dabei auch, sich von den Gegenwartsdebatten des Jahres 2021 (etwa zu Fake News oder der Debattenkultur in den sozialen Medien) zu lösen – und bestehende Zukunftskeime, Trends und schwache Signale in ihren Evolutionen in mögliche Zukünfte einzubetten. Manch ein Zukunftsbild für

10 Contract for the Web (2019).

11 Fleishmanhillard (2020).

12 Pörksen, B. (2021); Süddeutsche Zeitung (2021).

13 Rudelle, J. (2019).

14 BMBF (2020).

15 CNN (2022).

16 Prognos AG; Z_punkt GmbH The Foresight Company (2020).

17 Schaich, A.; Neef, A. (o. D.); Peter, S. et al. (2020).

18 Die Studie kann kostenlos auf der Webseite vorausschau.de heruntergeladen werden.

das Jahr 2035 mag dabei – gerade auch vor dem Hintergrund der jüngeren Debatten um die Weiterentwicklung des Internets hin zu einem Metaversum¹⁹ – im Jahr 2021 noch wie Science-Fiction anmuten, doch wer retrospektiv die Entwicklungssprünge der digitalen Sphäre seit dem Jahr 2006 analysiert, muss erkennen, dass vieles, was heute Normalität ist, damals noch schwer vorstellbar war. Zur retrospektiven Einordnung der vergangenen Dynamik: Facebook wurde 2004 gegründet, das erste iPhone kam erst 2007, WhatsApp zwei Jahre später auf den Markt.

Daher sollte „jede nützliche Idee über die Zukunft immer auch ein Stück weit lächerlich erscheinen“.²⁰ Denn neue Technologien ermöglichen neue Verhaltensweisen und stellen alte Überzeugungen infrage. Vieles, was für die Zukunft charakteristisch sein wird, ist zunächst neu und herausfordernd für den Gegenwartsdiskurs. Es scheint auf den ersten Blick typischerweise obszön, unmöglich, unlogisch, manchmal sogar grotesk. Und dann wird es doch vertraut und schließlich „normal“ – und damit auch ein Stück weit banal in der öffentlichen Wahrnehmung. So ist das, was heute oftmals als die „wahrscheinlichste Zukunft“ angesehen wird, bei Licht betrachtet eine der unwahrscheinlichsten Zukünfte.²¹ Ganz gleich, wie die Zukunft der digitalen Sphäre aussehen wird: Vertrauensbildung als impliziter sozialer Mechanismus wird sich an die veränderten Umfeldbedingungen anpassen. Dafür ist es für eine Bestimmung der Zukunft des Vertrauens in einer digitalen Welt vorab notwendig, eine Begriffsschärfung des im öffentlich Diskurs geradezu inflationär benutzten Begriffs „Vertrauen“ vorzunehmen.

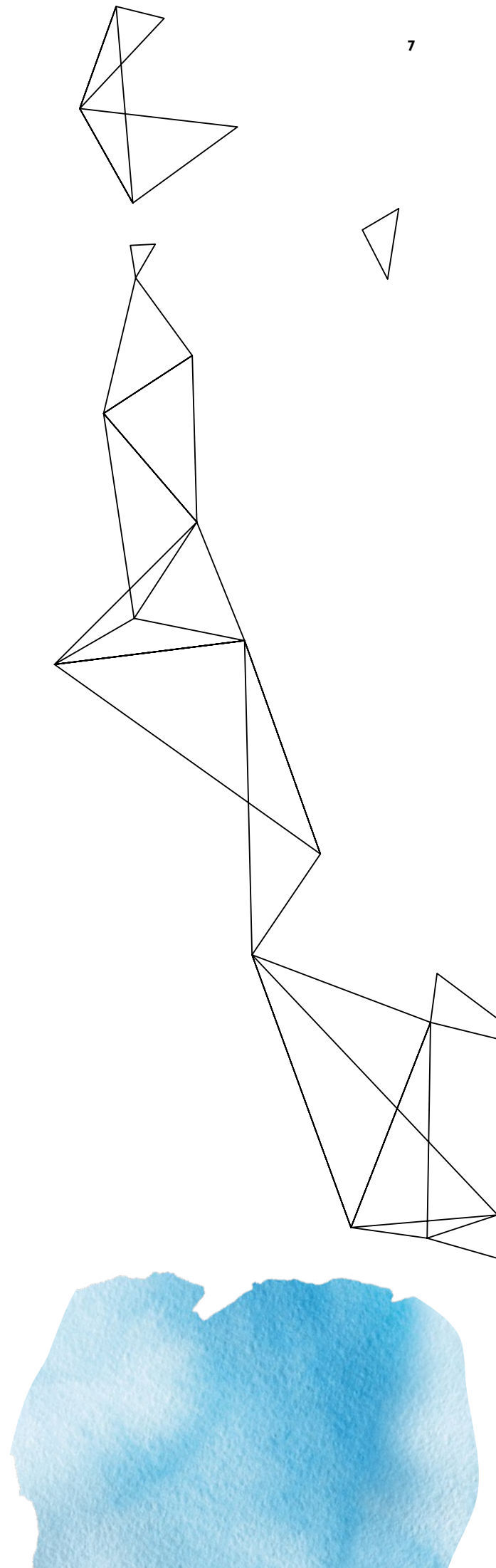
Hinweise zur Langfassung der Studie

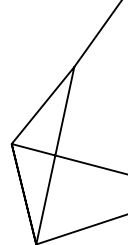
Eine umfassende Darstellung von Einflussfaktoren auf das Vertrauen von morgen findet sich in der Langfassung der Studie in Kapitel 2.4 und Kapitel 2.5. Die Langfassung der Studie findet sich zum Download auf vorausschau.de.

19 Merten, M. (2021).

20 So das in der Zukunftsforschung bekannte „Zweite Gesetz über die Zukunft“ von Jim Dator. Dator, J. (o. D.).

21 Dator, J. (o. D.).





Alltagsgeschichte aus der Zukunft Future Spotlight „Der europäische Weg“

Donnerstag, 17. Mai 2035.

6:15 Uhr. „I'm walking on sunshine ...“ Wie jeden Morgen spielt das EU-zertifizierte Assistenzsystem PAT – kurz für: Personal Assistant Technology – die Lieblings-Playlist von Lana zum Wachwerden. PAT ist mit 4,8 von 5 Punkten das virtuelle Assistenzsystem mit der höchsten Anti-Bias-Bewertung der EADA (European Algorithmic Debiasing Agency), das derzeit erhältlich ist. Firmen dürfen im Personalwesen seit einiger Zeit nur noch Assistenzsysteme mit einem Mindestscore von 4,65 einsetzen.

Ziemlich müde murmelt Lana: „PAT, gibt's Nachrichten?“ Die Musik stoppt – und PAT antwortet: „Lana, du hast zwei neue Bildsprachnachrichten“. Aus der ultradünnen, transparent anmutenden Smart Card auf Lanas Nachttisch wird vertikal ein dreidimensionales, sprechendes Hologramm projiziert. Es ist das Abbild ihres Mannes João, der mit dem gemeinsamen vierjährigen Sohn Luis gerade bei ihren Schwiegereltern in Portugal weilt. „Bom dia, Querida ...“. Lana liebt es, morgens durch das Hologramm begrüßt zu werden, wenn João nicht da ist. Offensichtlich hat er gerade durch den Newsfeed gescreent, denn er berichtet ganz begeistert von einem vielversprechenden Forschungsprojekt im europäischen Exzellenzcluster an der Universität seiner Heimatstadt Coimbra, wo Schlaganfallpatientinnen und -patienten mittels nicht invasiver Brain-Brain-Interfaces zwischen gesunden Angehörigen und Betroffenen therapiert werden sollen. João liebt spannende Forschungsprojekte. „Mein kleiner Nerd ...“, denkt Lana. Nachdem João's Botschaft zu Ende ist, verschwindet auch sein Hologramm wieder in der Smart Card. Die Smart Cards dienen bei den allermeisten Menschen in der EU als Multi-Kommunikationstool, aber auch zum Bezahlen, als Impfpass oder auch als digitaler Personalausweis. Ein klassisches All-in-one-Device eben. In ihrem Freundeskreis nutzen beinahe alle inzwischen nur noch die Smart Card. „Kaum zu glauben, wie unhandlich Smartphones früher waren“, denkt Lana. Voraussetzung für die Marktzulassung einer Smart Card ist, dass alle Daten ausschließlich auf in der EU befindlichen und zertifizierten Servern gespeichert werden.

Ein neues Hologramm erscheint, mit einer neuen Nachricht. Sie ist von Jan, ihrem belgischen Blogger-Kollegen aus Antwerpen. Er und Lana gehören zu einem investigativen europäischen Blogger-Netzwerk, das sich gegenseitig bei Recherchen unterstützt. Die Gruppe trifft sich auch in den sozialen Medien, so etwa im Web-3.0-Netzwerk EUTOPIA, innerhalb dessen sie einen eigenen virtuellen Redaktionsraum in Brüssel für ihre Avatare erstellt haben. Nachdem die chinesischen und US-amerikanischen Netzwerke infolge von Datenschutzskandalen in Europa massiv an Boden verloren haben, gilt zumindest in der Europäischen Union die französische Plattform EUTOPIA als der neue „virtual place to be“. Insbesondere die effiziente KI-basierte Identifikation von Hatespeech und Cyber- bzw. Avatarmobbing, die es Opfern ermöglicht, schnell Anzeige zu erstatten, hat dazu beigetragen, dass EUTOPIA als „nice network“ gilt. Eine befreundete Virtual-Interior-Designerin hatte die Ausgestaltung der Räumlichkeiten für die Gruppe übernommen. Lana und Jan arbeiten gerade gemeinsam an einer Story, die sich mit Deepfake-Echtbild-Avataren befasst. Diese Deepfakes wirken so echt, dass sie in den sozialen Medien selbst für erfahrene Cyber-Forensikerinnen und -Forensiker erst nach zahlreichen, von Künstlicher Intelligenz unterstützten Prüfverfahren identifizierbar sind.

Da Jans Nachricht auf Flämisch ist, gibt Lana den Sprachbefehl für Echtzeit-Übersetzung: „PAT, übersetze bitte auf Deutsch.“ Die Nachricht wird auf Deutsch wiedergegeben. Dahinter steht das von einem europäischen Hochschulverbund entwickelte, auf Deep Learning basierende Echtzeit-Übersetzungssystem EUROtranslate. Dieses hat auch dabei geholfen, eine europäische Öffentlichkeit entstehen zu lassen. Beim Launch von EUROtranslate im Jahr 2027 waren die Übersetzungen bei einigen Amtssprachen noch ziemlich holprig. Doch inzwischen ist selbst die maltesische Version dank starker Nutzung – und entsprechend vielen Trainingsdaten – auf einem hohen Niveau. Jan berichtet von seinen Recherchen zu den jüngsten Identitätsbetrugsfällen mit gehackten Echtbild-Avataren. Die Avatare der nichts ahnenden Besitzerinnen und Besitzer haben online Luxusgüter im hohen fünfstelligen Bereich in



den virtuellen Zwillingen der Münchner Edel-Stores in der Maximiliansstraße geshoppt. Dabei müssen die Täter auf sämtliche Daten der Betroffenen Zugriff gehabt haben, da selbst die von der EU vorgegebene verpflichtende Drei-Faktor-Authentifizierung umgangen werden konnte. Wahrscheinlich wurden die Smart Cards der betroffenen Echtbild-Avatare gehackt, denkt Lana. Oder die Täter hatten Zugriff auf die Stimmproben. Europol warnt jedenfalls davor, sich vorschnell auf Cybersöldnerinnen und -söldner aus Zentralasien festzulegen, berichtet Jan. Während Jan erzählt, fragt sich Lana kurz, ob das wirklich der „echte“ Jan ist, dessen Hologramm gerade zu ihr spricht. Sie lächelt und wischt den Gedanken schnell beiseite. Das dürfte Rufe nach dezentralen und mehrstufigen Identitätsnachweisen im digitalen Raum wieder lauter werden lassen, denkt Lana. Der cybersicherheitspolitische Ausschuss des EU-Parlaments fordert dies seit Jahren, konnte sich bislang aber noch nicht durchsetzen. Skeptische Stimmen verteidigen die zentrale Datenerfassung und befürchten ein „neues Bürokratiemonster“.

„PAT, bitte Kaffeemaschine an!“, ruft Lana aus dem Badezimmer, und sofort beginnt ihre italienische Siebträgermaschine sich aufzuheizen. Noch etwas müde läuft Lana in die Küche. Dort setzt sie ihre smarte Brille auf, per Iris-Scan wird ihr individuelles virtuelles Assistenzsystem freigeschaltet. Sie greift zu ihrer Smart Card. Das Assistenzsystem PAT weiß, dass Lana jeden Morgen zum Kaffee die wichtigsten europäischen Medien studiert. Die Europäische Zentrale für politische Bildung stellt jeden Morgen gebündelt relevante Medieninhalte aus der EU in allen gängigen Sprachen zur Verfügung. Heute empfiehlt ihr PAT unter anderem Artikel im dänischen „Politiken“, in der spanischen „El Pais“ sowie eine Kurzreportage eines in Toulouse lebenden, slowakischen Video-Bloggers über Korruptionsvorwürfe beim Bau von neuen Water-Harvesting-Anlagen in Südfrankreich.

Da fällt Lana ein, dass sie unbedingt noch eine potenzielle heikle Quelle für ihre Reportage kontaktieren muss. Die Uhr auf der Smart Card zeigt 7:52 Uhr. Sie gibt den Befehl „Bildschirm und Tastatur“, wor-

aufhin die Smart Glass das Sichtfeld als transparenten Bildschirm gestaltet und aus der Smart Card horizontal eine Tastatur auf den Tisch projiziert wird. Lana beginnt zu tippen und stoppt unmittelbar wieder. Ihr fällt ein, dass die Quelle – anders als sie selbst – digitale Kommunikation für grundsätzlich unsicher hält. Normalerweise bietet sie in solchen Fällen an, das mehrstufige NGO-Verifizierungstool TRUST zu nutzen, das bewusst auf staatliche Schnittstellen bei der Verifizierung verzichtet. Aber in diesem Fall scheint das Misstrauen so groß, dass es vermutlich keine gute Idee wäre, jetzt eine Message zu schreiben, selbst wenn sie doppelt verschlüsselt ist. Sie geht also zum Küchenschrank, holt Briefpapier und einen Umschlag heraus und beginnt, mit ihrem recycelten Kugelschreiber einen Brief zu schreiben ...

Neue Interaktionsmuster und Elemente der Vertrauens- und Misstrauensbildung im Future Spotlight „Der europäische Weg“

Neue Interaktionsmuster	› Virtuelle Assistenzsysteme sind extrem weit verbreitet, der Umgang noch normaler als heute ohnehin schon › Kommunikation mit Hologrammen, die als kommunikativer Transmitter zwischen Menschen dienen, ist alltäglich › Räumliche, dreidimensional erlebbare Social-Media-Foren › Tagtäglicher Umgang mit Echtbild-Avataren › Smart Card als neues persönliches Datenspeicher-Tool
Vertrauensbildende Elemente	› Anti-Bias-Bewertung der EADA (European Algorithmic Debiasing Agency) › Drei-Faktor-Authentifizierung › KI-basierte Hatespeech-Identifikation im sozialen Medium EUTOPIA › Stimmproben und andere biometrische Sicherungsverfahren
Misstrauensbildende Elemente	› Deepfake-Echtbild-Avatare › Avatarmobbing und Hatespeech

Quelle: © Eigene Darstellung Prognos AG und Z_punkt 2021.

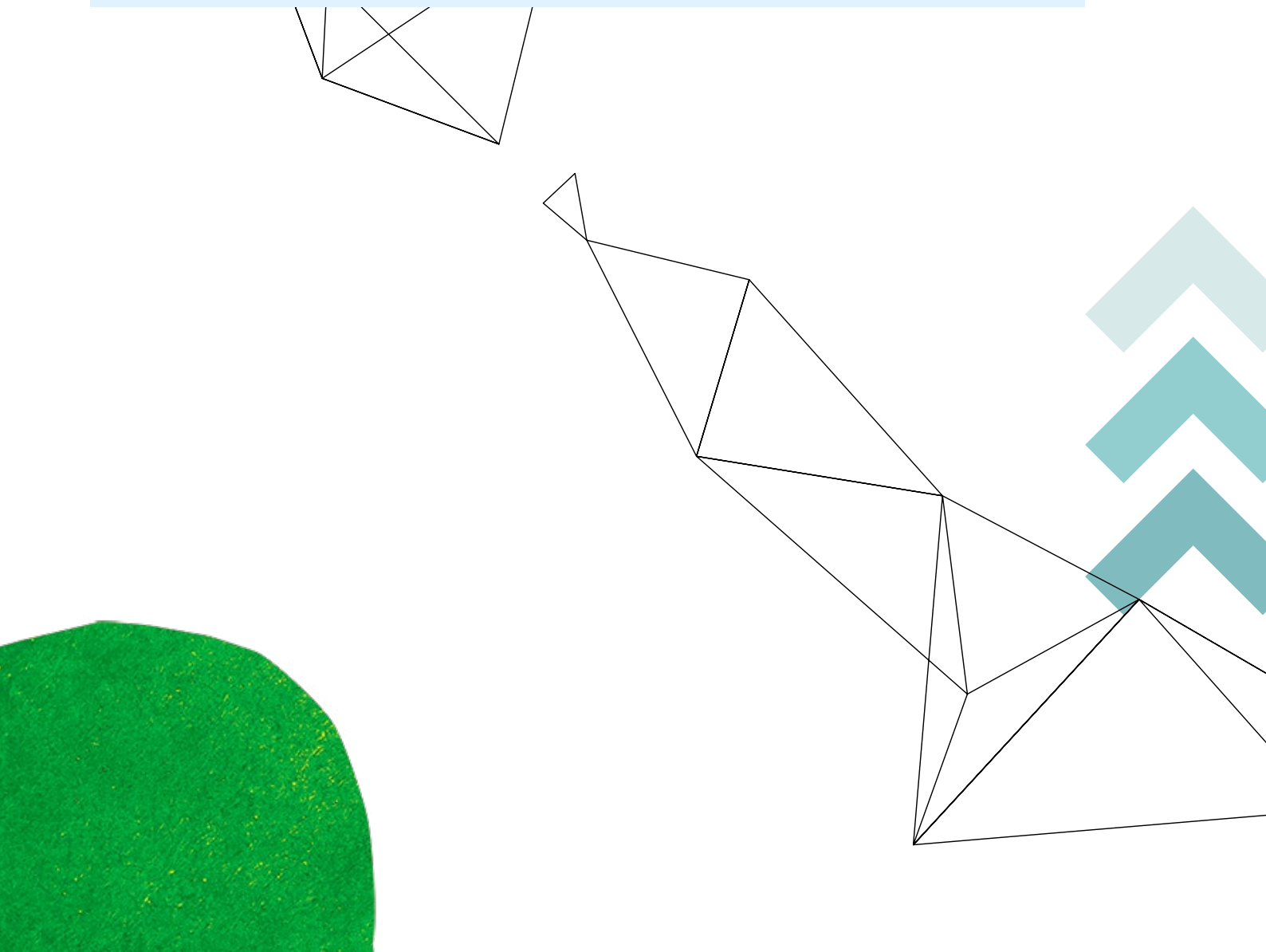
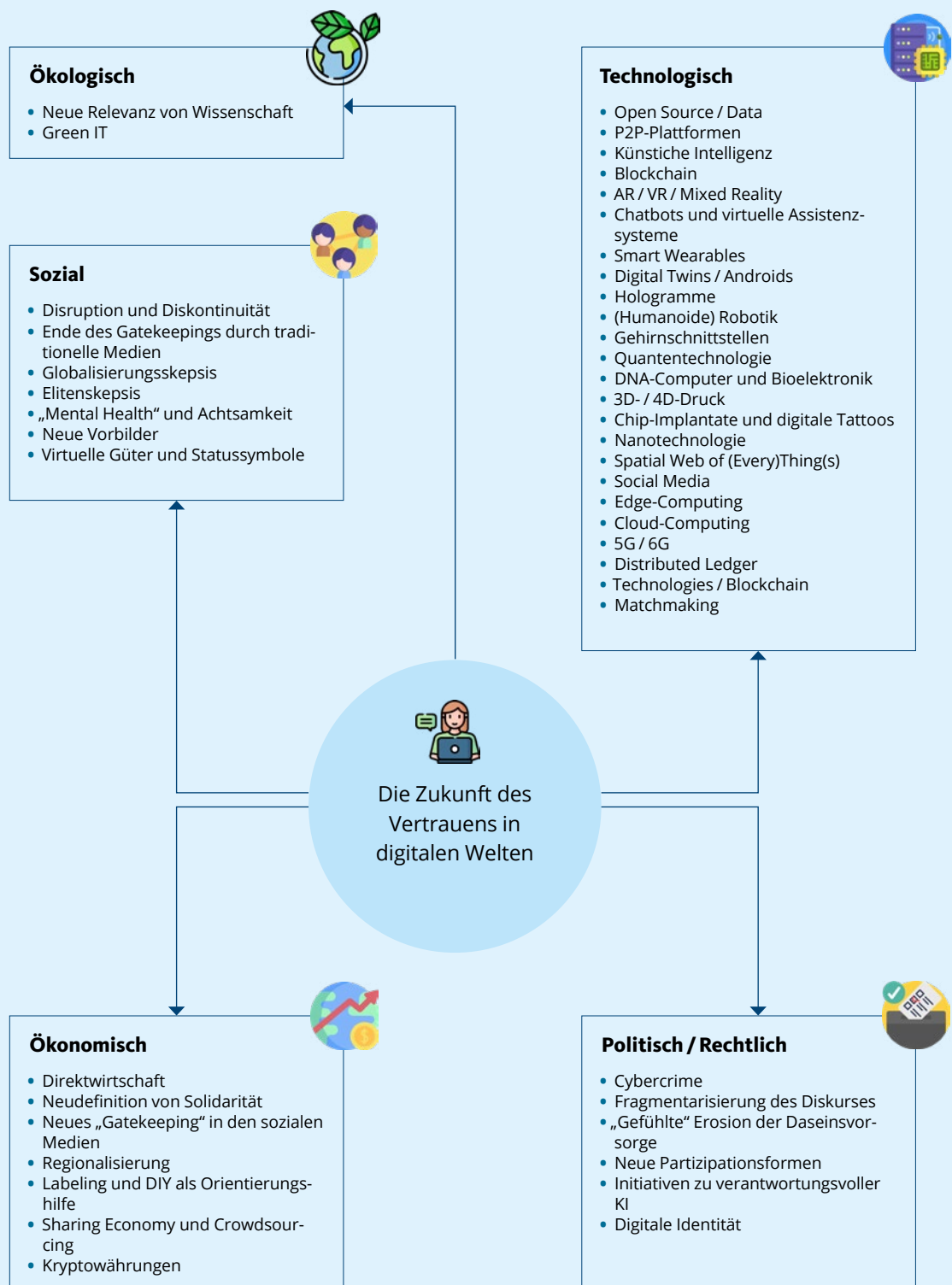


Abbildung 1: Treiberlandkarte nach STEEP-Sektoren geordnet, Tabelle 1: siehe Studie

1.1 Was ist Vertrauen?

Der zu Beginn angeschnittene Ansatz Luhmanns zur Bedeutung von Vertrauen in unserer Gesellschaft zeigt nicht nur, dass Vertrauen Handeln in komplexen Gesellschaften überhaupt erst ermöglicht, er zeigt auch, dass Vertrauen alle Bereiche des Lebens durchzieht. Hunderte von Ratgebern beschäftigen sich mit Vertrauen und Eifersucht in Beziehungen. Die Politik wirbt um das Vertrauen der Bürgerinnen und Bürger, Banken um das Vertrauen ihrer Kundinnen und Kunden, selbst Impfstoffe verlangen Vertrauen. Der Begriff und die Forderung nach Vertrauen scheint omnipräsent zu sein. Die Vielfalt der Kontexte, in denen Vertrauen eine Rolle zu spielen scheint, gibt einen ersten Hinweis darauf, dass Vertrauen nicht nur Komplexität reduziert, sondern selbst ein komplexes Phänomen ist. In der Auseinandersetzung mit dem Vertrauensbegriff stößt man auf zwei Probleme. Zum einen ist das alltagssprachliche Verständnis von Vertrauen ein anderes als das wissenschaftliche, zum anderen definiert jede wissenschaftliche Disziplin Vertrauen mit ihrem fachspezifischen Fokus, sodass es schlichtweg nicht die eine Definition von Vertrauen gibt.²² Wir versuchen also, uns an dieser Stelle dem Begriff anzunähern und eine Arbeitsdefinition von Vertrauen zu erarbeiten. Ziel ist es, auf dieser Basis besser zu verstehen, welche besondere Rolle Vertrauen in seinen verschiedenen Formen für das gesellschaftliche Zusammenleben spielt, wie sich Vertrauen bildet und was dazu beitragen kann, dass mehr oder weniger vertraut wird.

Was in der alltagssprachlichen Verwendung des Begriffs Vertrauen häufig fehlt, ist das, was Luhmann oben beschreibt: Vertrauen eröffnet Handlungsspielräume. Vertrauen sollte daher etwas konkreter als positive Erwartungshaltung in das zukünftige Handeln eines Gegenübers definiert werden. Dabei kann die Erwartungshaltung durchaus variieren und hängt vom jeweiligen Kontext ab. Vertraut man Ärztinnen und Ärzten, Pilotinnen und Piloten oder Banken, so beschränkt sich die positive Erwartungshaltung in der Regel auf deren (professionelle) Rolle. Vertrauen ist also nicht gleich Vertrauen – es hängt auch von verschiedenen Kontextfaktoren ab.²³

Vertrauen ist immer mit Unsicherheit verbunden. Es erlaubt zu agieren, obwohl man keine (absolute) Kontrolle bzw. Sicherheit hat. Man agiert also im Wissen, dass das eigene Vertrauen enttäuscht werden könnte. Das kann auch bedeuten, dass das Gegenüber das ihr oder ihm geschenkte Vertrauen zum eigenen Vorteil ausnutzt. Vertrauen hat daher auch immer einen Wagnischarakter. Einige Autorinnen und Autoren charakterisieren Vertrauen daher als Kosten-Nutzen-Abwägung²⁴, andere stellen eher die Tiefe der Verletzung, die man erfährt, wenn Vertrauen gebrochen wird, in den Vordergrund und fassen den Wagnischarakter daher als „akzeptierte Verletzlichkeit“.

Ob und wie man vertraut, hängt von zwei grundsätzlichen Aspekten ab. Zum einen sind Menschen verschieden, dies gilt auch für die Neigung zu vertrauen. Individuelle Disposition und Charaktermerkmale beeinflussen die Bereitschaft, Vertrauen gegenüber anderen aufzubringen.²⁵ Zum anderen ist die Wahrnehmung der Situation und des Gegenübers entscheidend.²⁶ Diese Wahrnehmung ist wiederum von eigenen Erfahrungen, Wissen und Werten geprägt, gleichzeitig ist sie nichts Statisches: Das Wissen über und die Erfahrungen mit der oder dem (potenziell) Vertrauten können sich mit der Zeit wandeln. Ebenso kann sich die eigene Erwartungshaltung entwickeln und sich je nach (zeitlichem) Kontext und Lebensbereichen der involvierten Individuen bzw. Institutionen verändern. Mit diesen Veränderungen können sich also auch Qualität und Tiefe des Vertrauens wandeln.²⁷

Für unsere Betrachtungen sind zwei Faktoren besonders interessant (siehe Abbildung 2). Erstens wie Risiken wahrgenommen werden bzw. für wie verletzlich man sich hält. Und zweitens anhand welcher Faktoren wir zu erkennen glauben, ob eine Person (oder Institution) vertrauenswürdig ist. Wir konzentrieren uns vorerst auf die Eigenschaften, die darüber bestimmen, ob man geneigt ist, seinem Gegenüber zu vertrauen.

22 Hatak, I. (2011).

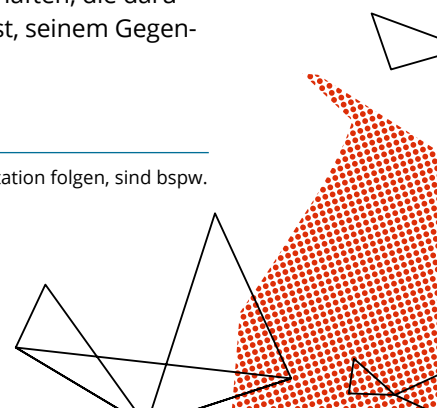
23 Die Bereitschaft, Informationen eines Gegenübers als vertrauenswürdig, generalisierbar und relevant zu akzeptieren, wird als epistemisches Vertrauen bezeichnet.

24 Klassische Autoren, die dieser Argumentation folgen, sind bspw. James Colemann und Diego Gambetta.

25 Enste, D.; Suling, L. (2020).

26 Mayer et al. (1995).

27 Lewicki, R. J.; Bunker, B. B. (1995).



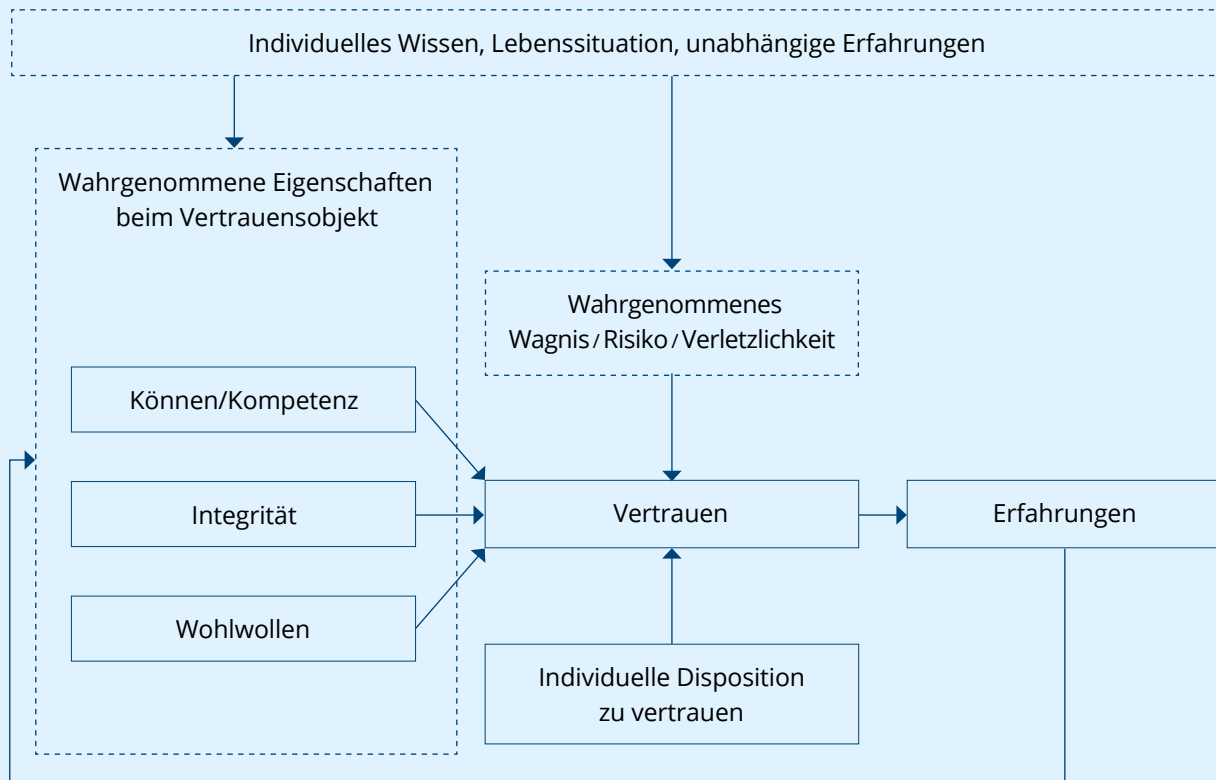
Diese lassen sich in drei Kategorien zusammenfassen:

- Können bzw. Kompetenz
- Integrität / Ehrlichkeit / Authentizität
- Wohlwollen / positive Absicht

Diese drei Eigenschaften sind nicht in allen Vertrauensbeziehungen gleichermaßen wichtig. Gerade gegenüber Freundinnen und Freunden wie auch Bekannten spielen „Kompetenz und Können“ wohl eine eher untergeordnete Rolle. Umso wichtiger ist dieser Faktor bei Personen und Institutionen, von denen man eine konkrete (professionelle) Rollenerwartung hat. Es fiel wahrscheinlich den meisten Menschen schwer, einer Polizei zu vertrauen, die nicht in der Lage ist, die öffentliche Ordnung aufrechtzuerhalten, auch wenn sie in bester Absicht handelt. Gerade beim Faktor Können sollte man sich vor Augen führen, dass all die hier genannten Eigenschaften wahrgenommene Eigenschaften sind. Es ist also letztlich nicht entscheidend, wie kompetent die Polizei tatsächlich ist, sondern in erster Linie, für wie kompetent sie gehalten wird.

Die vertrauensbildenden Eigenschaften von „Integrität / Ehrlichkeit / Authentizität“ erklären sich über die Bestimmung des Vertrauens als „positive Erwartung“. Sie alle vermitteln das Gefühl, das zukünftige Handeln des Gegenübers besser einschätzen zu können. Der letzte Punkt ist wahrscheinlich der für das Vertrauen spezifischste Faktor: die wahrgenommenen „positiven Absichten des Gegenübers“. In persönlichen Beziehungen ist dieser Aspekt recht eingängig. Der Eindruck, dass das Gegenüber einem wohlgesonnen ist, scheint eine fast unabdingbare Bedingung dafür zu sein, eine Freundschaft zu diesem aufzubauen. Aber auch bei abstrakteren Beziehungen hilft der Faktor Wohlwollen, Vertrauen von anderen sozialen Phänomenen abzugrenzen. Die Politik dient hier als Beispiel: Egal für wie kompetent und integer man eine Politikerin oder einen Politiker hält, solange diese andere politische Ziele verfolgen als die eigenen, wird man ihnen kaum vertrauen, die eigenen Absichten zu vertreten. Häufig werden daher in der Literatur wie auch von den im Rahmen der Studie interviewten Expertinnen und Experten gemeinsame Werte oder wahrgenommener Respekt und Anerkennung als weitere vertrauensbildende Faktoren genannt.²⁸

28 Interviews mit Expertinnen und Experten.

Abbildung 2: Faktoren der Vertrauensbildung

Quelle: © Eigene Darstellung Prognos AG und Z_punkt, in Anlehnung an Mayer et al. (1995).

Hinweise zur Langfassung der Studie

Eine umfassende Darstellung zum aktuellen Forschungsstand zu Vertrauen findet sich in den Kapiteln 2.1–2.3 in der Langfassung der Studie. Die Langfassung der Studie findet sich zum Download auf vorausschau.de.

1.2 Vertrauen in Technik

Gibt es überhaupt Vertrauen in Technik? Stand 2021 lautet die weit überwiegende Meinung der Expertinnen und Experten, mit denen wir im Rahmen der diesem Essay zugrunde liegenden Studie gesprochen haben: Nein.²⁹ Nun mag ein Einwand lauten, dass viele Menschen doch jeden Morgen darauf vertrauen, keinen Stromschlag zu bekommen, wenn sie den Fön in die Hand nehmen. Mit Blick auf die oben stehende Definition erfüllt ihre Situation alle Kriterien. Sie handeln in der positiven Erwartung, dass der Fön genau das tut, was sie von ihm erwarten – nämlich die Haare zu trocknen. Es kann vorausgesetzt werden, dass die meisten Menschen weder den technischen Sachverstand und die Muße haben, jeden Morgen ihren Fön zu zerlegen, um zu kontrollieren, ob der Fön noch intakt ist. Sie handeln also unter Unsicherheit. Sie gehen also tatsächlich ein gewisses Risiko ein, sich zu verletzen, sollte wider Erwarten ein technischer Defekt vorliegen. Dennoch behaupten die Expertinnen und Experten, man könnte Technik nicht vertrauen. Einen ersten Hinweis, warum das so ist, gibt schon die Formulierung. Wenige Menschen würden die Formulierung wählen „ich vertraue meinem Fön“, sondern wohl eher „ich vertraue darauf, dass mein Fön funktioniert“. Die Formulierung „auf etwas vertrauen“ macht deutlich, dass es an dieser Stelle vielmehr um etwas geht, das man als „Systemvertrauen“ beschreiben könnte. Man vertraut also weniger dem Fön selbst als vielmehr dem System aus Zertifizierungen, Kontrollen und eigenen Erfahrungen um den Fön herum. Noch deutlicher wird es, sollte jemand nun doch einen Stromschlag bekommen. In diesem Fall wird die- oder derjenige den Fön wahrscheinlich nicht mehr benutzen, vielleicht sogar mit der Begründung, dass er oder sie dem Fön nicht mehr vertraut. Aber würden Menschen auch sagen, der Fön habe ihr Vertrauen enttäuscht, gebrochen oder gar missbraucht? Wohl eher nicht. Denn damit würden sie dem Fön ein Motiv, letztendlich sogar einen Willen unterstellen. Und damit sind wir beim Kern dieses kleinen philosophischen Exkurses: Vertraut man, erkennt man die Motive des anderen an – in der Erwartung, dass das Gegenüber seine Motive nicht über die eigenen stellt, das geschenkte Ver-

trauen also ausnutzt und einen Vorteil daraus zieht. Weil Technik keinen Willen bzw. keine eigene Agenda hat, kann sie aber Vertrauen nicht ausnutzen. Wenn sie es doch tut, so nicht, weil sie es „will“, sondern weil sie so konstruiert wurde, es zu tun. Wenn wir davon sprechen, Technik zu vertrauen, sprechen wir in der Regel davon, den Entwicklerinnen und Entwicklern, den Ingenieurinnen und Ingenieuren oder den Instanzen, die Technik kontrollieren, und Standards zu vertrauen.³⁰

Warum schreiben wir nun so explizit, dass es laut Expertinnen und Experten Stand 2021 kein Vertrauen in Technik gibt? Algorithmen helfen uns, als digitale Assistenten im eigenen Zuhause, als Fahrerassistenz im Auto oder als Chatbots im Kundenservice, Entscheidungen zu treffen oder den Alltag zu organisieren. Viele der bisher genutzten Algorithmen basieren dabei auf klassischen Entscheidungsbäumen. Hier ist die Reaktion der Technik auf das eigene Handeln klar nachvollziehbar. Verschiedene KI-Algorithmen funktionieren jedoch nicht mehr nach diesem Prinzip. Sie haben selbst „gelernt“, Entscheidungen zu treffen. Nach welchen Prinzipien die Entscheidung getroffen wird, bleibt teilweise selbst den Programmierenden verborgen. Aus komplizierten linearen Systemen werden komplexe nonlineare Systeme, in denen sich der Output nicht mehr linear aus dem Input vorausberechnen lässt.³¹ Beurteilt werden kann nur die Performanz des Algorithmus – also wie gut die Ergebnisse sind, die er hervorbringt. Wir als Laien wissen somit nichts darüber, warum eine Entscheidung getroffen wurde. Oder anders gesagt: wir kennen die Motive hinter der Entscheidung nicht. Die oben stehenden Modelle zur Vertrauensbildung gehen jedoch davon aus, dass die Entscheidung zu vertrauen auch immer mit der eigenen Auffassung über die Absichten des Gegenübers zusammenhängt. Wenn das Gegenüber jedoch keine Absichten, keinen eigenen Willen hat – und davon ist beim heutigen Stand der Technik noch auszugehen –, so weist die heutige Verwendung des Konzepts „Vertrauen“ hier eine Unschärfe auf, denn auch wenn keine Annahmen über die Absichten einer KI getroffen werden können, so setzt man sich dennoch Entscheidungen aus, die man

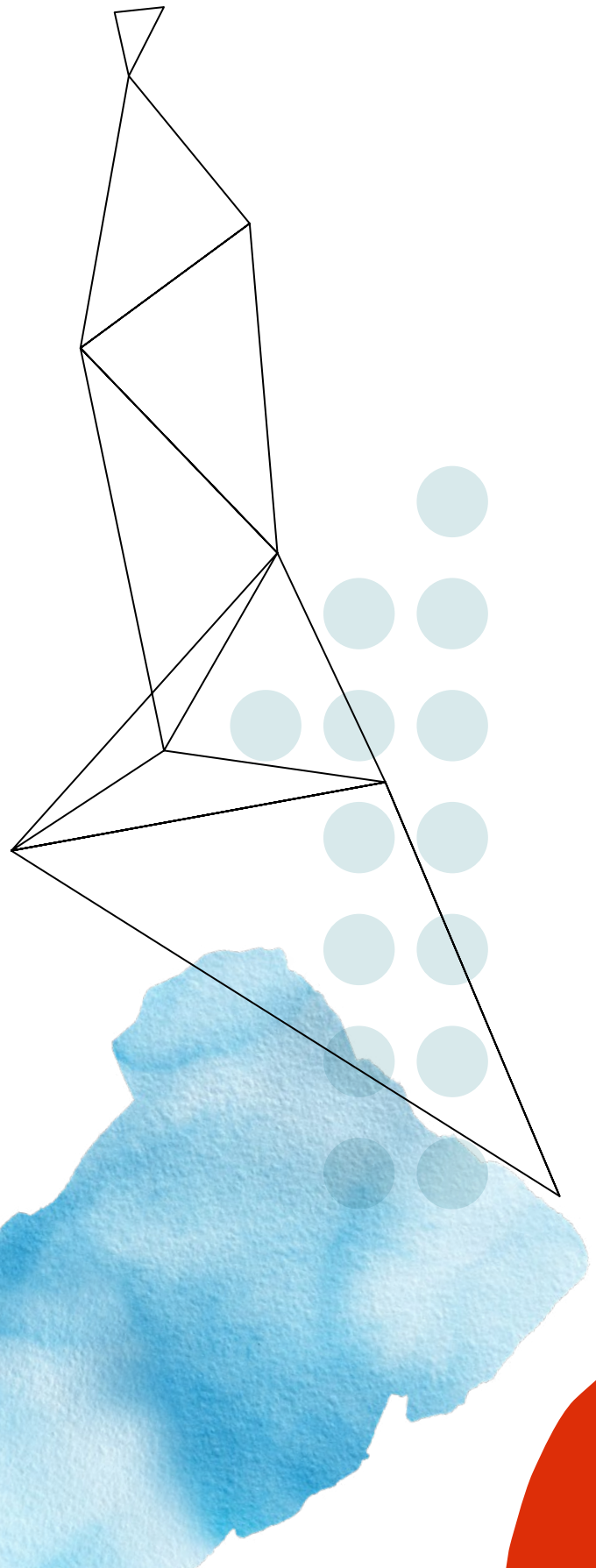
29 Gespräche mit Expertinnen und Experten im Rahmen dieser Studie. Eine Liste der Personen, mit denen gesprochen wurde, findet sich im technischen Bericht.

30 Hartmann, M. (2020).

31 Ramge, T. (2020).

nicht nachvollziehen, nicht kontrollieren kann, und begibt sich damit in eine Situation, die klassischerweise von Vertrauen geprägt wäre.

Davon unabhängig kündigt sich eine zweite Entwicklung an: Technik wird menschlicher. Algorithmen lernen unter anderem, sich in natürlicher Sprache auszudrücken, und Roboter sehen immer menschlicher aus. Es stellt sich also die Frage: Was passiert, wenn wir nicht mehr in der Lage sind, Technik und Menschen zu unterscheiden? Selbst wenn die Technik noch immer keinen Willen hat, würden wir ihr nicht dennoch Motive zuschreiben, wenn wir nicht wüssten, dass sie nicht menschlich ist? Können wir dann auch vertrauen? An dieser Stelle kündigen sich neue Formen des Vertrauens – oder schlicht eine Neudefinition dessen, was wir unter Vertrauen verstehen – an. Dieses neue Vertrauen könnte jedoch weit fragiler sein als das in unsere Mitmenschen. Denn während wir Menschen Fehler verzeihen – in der Annahme, dass Fehler unbeabsichtigt waren und daraus gelernt wird – und Vertrauen dadurch wiederhergestellt werden kann, bleibt bei Maschinen nur ein neuer Versuch, eine neue Programmierung und die Hoffnung, dass es diesmal funktioniert.³²



32 Dietvorst et al. (2015).

1.3 In welchen digitalen Welten vertrauen wir morgen?

In der langfristigen Vorausschau sollte sich die Perspektive auf das richten, was nach dem Web 2.0 kommen könnte. Während die Menschen in drei Dimensionen leben, ist das heutige Web weitestgehend zweidimensional. Das Web 2.0 wurde für die gemeinsame Nutzung von Informationen konzipiert, die über den Flachbildschirm der Endgeräte verfügbar sind. Aktuelle Bestrebungen zielen darauf ab, die digitale Sphäre wortwörtlich in eine neue Dimension zu überführen. Für diese Visionen existiert eine Vielzahl von Begrifflichkeiten: das Spatial Web, das Web 3.0 oder auch das Metaversum.³³ Obwohl er aus einem eher dystopischen Science-Fiction-Kontext entstammt³⁴, scheint sich der Terminus „Metaversum“ sukzessive in der öffentlichen Debatte als Synonym für die räumliche Weiterentwicklung des Internets zu etablieren. Dies hat sicher auch mit der großen weltweiten Aufmerksamkeit zu tun, welche die Umbenennung von Facebook in Meta sowie die Pläne von Facebook-Gründer Mark Zuckerberg für ein Metaversum auf sich zogen.

Die konkrete Ausgestaltung des Metaversums, wie sie sich Mark Zuckerberg vorstellt, ist aber nur ein mögliches Szenario unter vielen für eine mögliche Weiterentwicklung der digitalen Sphäre.³⁵ Allen Begriffen von Spatial Web über Web 3.0 bis hin zu Metaversum gemeinsam ist aber die Vorstellung eines räumlichen Webs – einer Computeratmosphäre, die in einem umfassenden dreidimensionalen Raum existiert. Die Grenzen von analoger und virtueller Realität verschwimmen dabei.³⁶ Was zum Cyberspace gehört und was nicht, könnte zukünftig nicht mehr so leicht zu kategorisieren sein. Ermöglicht werden soll dieses räumliche und semantische Web durch die Konvergenz einer Vielzahl von Technologien, Infrastrukturen, Anwendungen und soziotechnologischen Phänomenen (siehe Abbildung 3). Das Zusammenwirken von Technologien wie KI, intelligenten Sensoren, Distributed-Ledger-Systemen, Digital Twins, 5G (und 6G), AR, VR und miteinander verbundenen Objekten hebt die Grenzen von physischer und Online-Welt weitestgehend auf.

In dieser neuen digitalen Sphäre könnte ein Großteil der Interaktion mit digitalen Informationen nicht mehr auf den heute allgegenwärtigen Bildschirmen, Tablets und Smartphones stattfinden. Neue Schnittstellen entstehen. Parallel zur neuen Struktur könnten sich also auch neue Endgeräte durchsetzen. Denkbar wären hier AR- und VR-Applikationen, IoT-Wearables sowie smarte Brillen und Kontaktlinsen, die sich nahtlos mit der physischen Umgebung verbinden. Im Spatial Web könnte langfristig jedes physische Element in der realen Welt vollständig digitalisiert werden. Auch könnte es für jeden Menschen virtuelle Avatare geben, die an virtuellen Arbeits- oder Treffpunkten mit anderen Avataren zusammentreffen. Diese Verfügbarmachung von technophysischer Präsenz könnte die menschlichen Interaktionsmuster um neue Elemente erweitern. Während im Web 2.0 über ein Domain-Name-System URLs vergeben wurden, werden im Web 3.0 / Spatial Web Adressen für Räume vergeben. Dadurch können die Nutzerinnen und Nutzer lokalisierbare Orte, physische Objekte, Personen oder digitale Inhalte im Cyberspace identifizieren und besuchen.

33 Rene, G. (2019), Deloitte (2020), The Economist (2020).

34 Der Begriff wurde 1992 durch den Science-Fiction-Autor Neal Stephenson in seinem Cyberpunk-Roman „Snow Crash“ geprägt.

35 Tagesschau.de (2021), Merten, M. (2021), Ball, M. (2021), The Economist (2021).

36 Deloitte (2020).

Abbildung 3: Grundlegende technologische Treiber des Web 3.0 (Spatial Web)

Technologien, technologische Infrastrukturen und soziotechnologische Phänomene

Technologien

 Künstliche Intelligenz	 Chatbots und virtuelle Assistenzsysteme	 Smart Wearables	 Digitale Androiden / Digital Twins
 Augmented, Mixed und Virtual Reality	 Hologramme	 Humanoide Robotik	 Gehirnschnittstellen (BCI und BBI)
 Quantentechnologien	 DNA-Computer und Bioelektronik	 3D- / 4D-Druck	 Chip-Implantate und digitale Tattoos
 Nanotechnologie			

Technologische Infrastrukturen

 Spatial Web und Internet of (Every-)Thing(s)	 Social Media und Plattformen	 Edge-Computing
 5G / 6G	 Distributed Ledger, Technologies / Blockchain	 Cloud-Computing

Soziotechnologische Phänomene

 Open and Free Data	 Digitale Identität	 Virtuelle Güter und Statussymbole	 Sharing Economy und Crowdsourcing
 Matchmaking	 Green IT	 Kryptowährungen	

Quelle: © Eigene Darstellung Prognos AG und Z_punkt 2021.

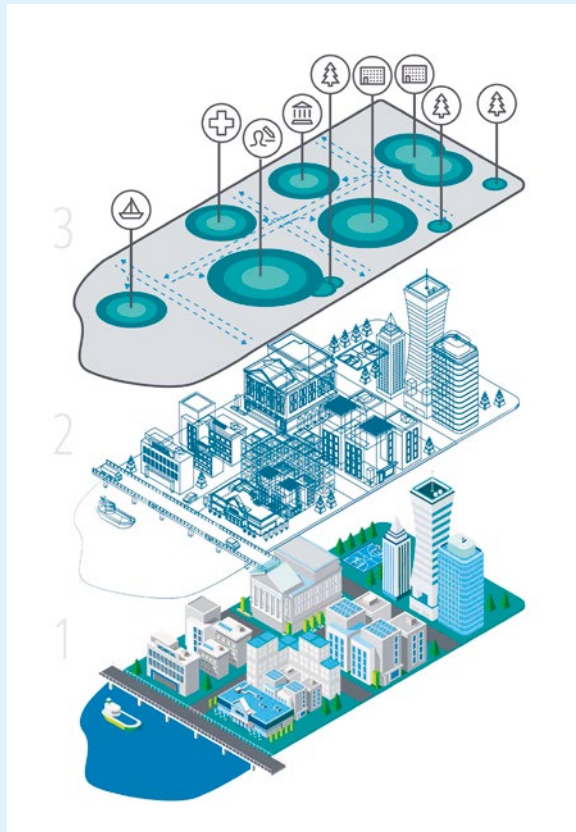
Menschliche Interaktionen mit realen und virtuellen Versionen anderer Menschen könnten sich nachhaltig verändern. Dies bedeutet auch, dass neue Muster der Vertrauensbildung notwendig würden, da diese erweiterte Dimension des Digitalen mit den bisherigen Kulturtechniken nicht zu erfassen ist. Ein Web 3.0 / Spatial Web / Metaversum könnte in vielfacher Hinsicht unsere Sicht auf die Welt verändern (siehe Abbildung 4). Eine räumliche digitale Sphäre könnte aber auch dazu führen, dass das Vertrauen in die digitale Sphäre grundsätzlich steigt, da hier ein sinnliches Erleben der virtuellen Sphäre viel umfänglicher als in der „flachen“ 2D-Welt des Web 2.0 möglich ist. Neue Formen der interpersonellen und institutionellen Vertrauensbildung dürften in diesen neuen Realitäten entstehen, etwa Vertrauensbil-

dungsprozesse zwischen Avataren, zwischen Menschen und Avataren oder auch von Menschen in die neuen Technologien.

Durch ein Spatial Web / Web 3.0 / Metaversum könnten also bestehende Vertrauensmuster um neue Elemente erweitert werden. Eine räumliche digitale Sphäre könnte dabei helfen, das Vertrauen zu steigern, indem Webseiten durch erlebbare Räume mit „echten“ Avataren, die das Äußere ihrer realen Vorbilder haben, ersetzt würden. In Kombination mit KI-basierter Echtzeitübersetzung könnten so auch kulturelle Verständnisbarrieren leichter überwunden werden. Zugleich könnte bei Personen mit negativen Erfahrungen das Misstrauen dahingehend steigen, dass Ängste nun auch in der digitalen

Abbildung 4: Die drei Ebenen des Spatial Web

Technologien, technologische Infrastrukturen und soziotechnologische Phänomene



Räumliche Interaktionsebene

Durch die Schnittstellen der nächsten Generation wie Smart Glasses oder Stimme werden wir in der Lage sein, mit kontextbezogenen Echtzeitinformationen zu interagieren, die durch intuitive und sensorische Trigger wie Geolocation, Computer Vision und Sprach-, Gesten- oder biometrische Befehle aufrufbar sind. Im Endeffekt verschmelzen so die digitale und die physische Ebene für Nutzerinnen und Nutzer.

Digitale Informationsebene

Durch ubiquitäre Sensorik und digitale Kartierung der physischen Welt könnte theoretisch ein digitaler Zwilling von jedem Objekt an jedem Ort entstehen. Heute wird auf diese Art von digitalen Informationen hauptsächlich über Bildschirme und Dashboards zugegriffen. In Zukunft könnten sie vor allem über die räumliche Interaktionsebene abgerufen werden.

Physische Ebene

Die physische „reale“ Welt, welche Menschen durch ihre fünf Sinne wahrnehmen.

Quelle: © Deloitte (2020), Beschreibungen adaptiert von Prognos AG und Z_punkt 2021.

Sphäre konkreter werden. Die Offenheit dieser Fragen zeigt, dass sich die Zukunft des Vertrauens in der digitalen Sphäre nicht nur entlang technologischer Fortschritte herausbilden wird, sondern auch im Kontext größerer gesellschaftlichen Entwicklungen. Die weitere Entwicklung der digitalen Sphäre beruht maßgeblich auf dem komplexen Zusammenspiel von relevanten Technologien, technologischen Infrastrukturen und soziotechnologischen Phänomenen – die in unterschiedlichen sozialen, ökonomischen und politischen Rahmenbedingungen auch unterschiedliche Wirkungen entfalten können.

Hinweise zur Langfassung der Studie

Kapitel 3 der Langfassung der Studie stellt die Ergebnisse zur digitalen Sphäre von morgen dar. Die Langfassung der Studie findet sich zum Download auf vorausschau.de.

1.4 Vertrauen, Risikoabschätzung und Sicherheit in den digitalen Welten von morgen

Nicht erst seit der Corona-Krise haben sich digitale Infrastrukturen als unverzichtbar erwiesen. Inzwischen sind sie aus dem Alltag und dem Wirtschaftsleben nicht mehr wegzudenken. Was auf der einen Seite einen Zugewinn an Bequemlichkeit, Vernetzung und niedrigschwelliger Kommunikation bedeutet, bringt auf der anderen Seite einen indirekten Zwang zur Nutzung der digitalen Infrastruktur mit sich. Viele Dienstleistungen sind heute nur noch online verfügbar. Eine bewusste Nichtnutzung der digitalen Infrastrukturen bedeutet auch einen massiven Verlust an sozialer Teilhabe – und in letzter Konsequenz eine Exklusion von zivilisatorischen Leistungen.

Wie oben bereits beschrieben, wäre das Entstehen eines Cyberspace ohne Vertrauen nicht möglich gewesen. Dazu beigetragen hat auch eine verzerrte Wahrnehmung von potenziellen Risiken in der digitalen Sphäre: Da virtuelle Risiken nicht greifbar sind, werden diese nicht als wirklich bedrohlich eingeschätzt. Vielmehr fürchten sich Menschen am meisten vor Bedrohungen, die unbekannt sind, sich rasant ausbreiten, katastrophale Ausmaße annehmen können und potenziell für jeden tödlich sind.³⁷ Virtuelle Risiken stellen aber (noch) keine unmittelbare Gefahr für Leib und Leben dar. Hinzu kommt ein Gewöhnungseffekt: Je vertrauter uns ein Risiko wird, desto mehr lässt die anfängliche Besorgnis nach.³⁸ Verstärkend wirken dabei Netzwerkeffekte: Das Erreichen einer kritischen Masse an Nutzerinnen und Nutzern einer App oder einer Software führt dazu, dass Vorbehalte ausgeblendet werden – und die Vorteile der Nutzung in den Vordergrund rücken.³⁹

Wie oben bereits beschrieben, könnte das Web 3.0 / Spatial Web / Metaversum unsere Sicht auf virtuelle Risiken durch die räumliche, haptisch anmutende Optik verändern. Risiken könnten „realer“ wahrgenommen werden. Dabei ist davon auszugehen, dass Cyberkriminalität und -mobbing auch im

räumlichen Web eine große Rolle spielen werden.⁴⁰ Durch das dreidimensionale Erleben könnten Diskriminierung und Kriminalität deutlich greifbarer werden, sodass aus bislang eher abstrakten Ängsten in der digitalen Sphäre realere Bedrohungen werden. Erste Anzeichen hierfür zeigen sich bereits in Berichten zu sexuellen Übergriffen in VR-basierten Computerspielen oder VR-Welten.⁴¹ Ein wichtiger Aspekt dürfte hier auch sein, inwieweit bestehende Vorurteile und Diskriminierungsmuster aus der realen Welt im räumlichen Web ihre Fortführung finden – oder ob in diesem Kontext neue Diskriminierungsmuster entstehen. Grundsätzlich denkbar wären hier Formen des Avatar-Mobbings, welches wohl das interpersonelle Vertrauen im räumlichen Web schwächen dürfte. Ähnlich wie beim bereits bestehenden Phänomen des Cybermobbings, sind auch bei neuen Varianten psychischer Gewalt Auswirkungen in der analogen Welt in Form von psychischen Problemen denkbar. Sowohl neue Cyberphänomene als auch das grundsätzliche Vertrauensniveau dürften dabei abhängig von den regulatorischen Rahmenbedingungen sein. Dabei sind die weiteren Entwicklungen – sowohl technologisch als auch gesellschaftlich-politisch – mit hoher Unsicherheit behaftet. Fragen hinsichtlich einer möglichen zukünftigen Offenlegungspflicht digitaler Identitäten und Klarnamenpflicht sind noch lange nicht abschließend diskutiert.⁴²

Zugleich ist davon auszugehen, dass die Bedeutung digitaler Infrastrukturen weiter zunehmen wird. Dies zeigt sich unter anderem in den großen Investitionen in diesem Bereich in den kommenden Jahren⁴³, die auch auf kollektiven Erwartungshaltungen hinsichtlich der weiteren Bedeutung beruhen.⁴⁴ Dies könnte die de facto bereits bestehende „Unverzichtbarkeit“ der digitalen Sphäre weiter verstärken. Zugleich ist aufgrund technologischen Fortschritts auch eine dynamische Weiterentwicklung der digitalen Sphäre zu erwarten. Digitalskeptikerin-

37 Slovic, P. (1987).

38 Die Zeit (2021).

39 TH Köln (2015).

40 Security Affairs (2021).

41 Süddeutsche Zeitung (2022).

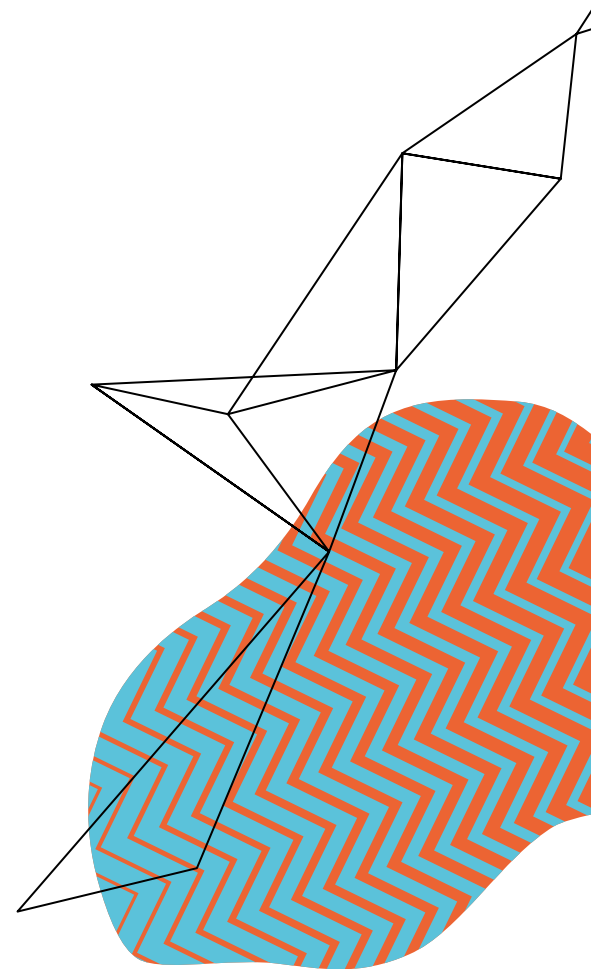
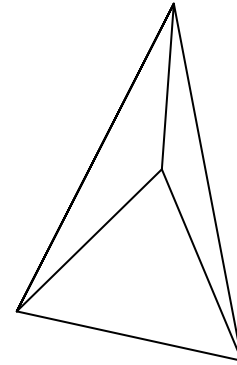
42 Marx, I. (2020).

43 Bundesministerium der Finanzen (2020).

44 Beckert, J. (2018).

nen und -skeptiker hätten immer weniger die Möglichkeit eines Opt-outs. Es könnte also sein, dass es zu einer Ausdifferenzierung der Vertrauensgrundlage hinsichtlich der Nutzung digitaler Dienste und Infrastrukturen kommt. Auf der einen Seite die, die sich auch in Zukunft mit einer Art „Grundvertrauen“ in einer sich dynamisch weiterentwickelnden digitalen Sphäre bewegen. Und auf der anderen Seite die, die sich gegen ihre Überzeugung in digitalen Welten bewegen (müssen), weil ihnen sonst soziale Teilhabe oder der Zugang zu relevanten Dienstleistungen verwehrt bleiben.

Für Letztere könnte das Vertrauen in dem Maße erleichtert werden, indem die Cybersicherheit erhöht wird. Für die erstgenannte Gruppe gilt hingegen das, was seit Anbeginn des Internets gilt: Auch das Web der Zukunft kann sich nicht ohne Vertrauen entwickeln. Es braucht also auch in Zukunft Menschen, die bereit sind, in immer neue Räume vorzustoßen – ganz gleich, welche technologischen Hilfsmittel sie dabei nutzen. Bisher verlief der Weg auf einer Einbahnstraße vom Analogen ins Digitale. Das Web 3.0 mit der Erzeugung von hybriden Wirklichkeiten könnte nun dafür sorgen, dass die Fahrbahn auch in die andere Richtung verläuft. Um mit dieser neuen soziodigitalen Komplexität umgehen zu können, braucht es Vertrauen. Gestern, heute, morgen und übermorgen.



Alltagsgeschichte aus der Zukunft Future Spotlight „Ökologische Regionalisierung“

Donnerstag, 17. Mai 2035.

11:00 Uhr. Lana kommt von der morgendlichen Gemeinschaftsarbeit zurück. Ihren Sohn hat sie gerade in der Kita abgegeben. Zweimal die Woche hilft sie Pasquale beim Aufbau eines Stadtgartens. Pasquale ist Rentner und möchte in einer alten Lagerhalle eine Vertical Farm für die Gemeinde errichten und damit anderen Freiwilligen einen intuitiven Zugang zum digitalen Gärtnern ermöglichen. Lana unterstützt ihn bei der Programmierung eines humanoiden Roboters, der Kindern die Versorgung der unterschiedlichen Pflanzentypen und die Funktionsweise der Farm nahebringen soll. Beide wissen, dass es eine Herausforderung sein wird, die Eltern von dem Projekt zu überzeugen. Viele sehen zwar den wirtschaftlichen Nutzen von digitalen Anwendungen, möchten ihre Jüngsten aber noch nicht in Kontakt mit der immersiven und virtuellen Welt bringen. Lanas Sohn ist jedenfalls begeistert von dem Projekt, und sie ist überzeugt, dass er so deutlich mehr Wissen mitnimmt als durch irgendwelche Videos.

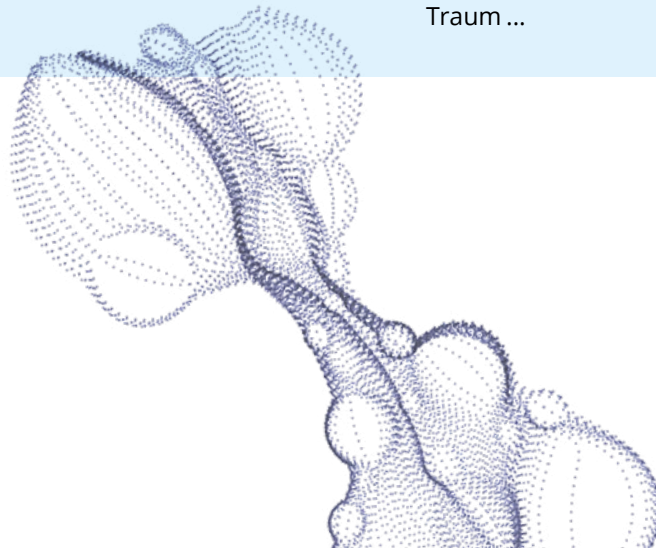
Zu Hause angekommen, füllt sich Lana ein Glas Wasser in der Küche. Ihre Smartwatch vibriert. Lana schließt den Hahn und schaut auf die Uhr: „Heute ist Freitag. Sie haben erst die Hälfte des üblichen wöchentlichen Verbrauches erreicht. Glückwunsch!“ Lana ist Teil eines Digital-Twin-Pilotprojekts für Haushalte. Ihre Wohnung ist mit smarten Sensoren ausgestattet, um Daten über Strom-, Wasser- und Abfallmengen zu analysieren. Gleichzeitig erhält Lana regelmäßig personalisierte Tipps, um ihre Verbräuche zu optimieren, oder wird über erste Anzeichen einer Erkrankung informiert. Wie viele andere Bürgerinnen und Bürger sieht sie digitale Daten als wichtigen Beitrag zur Gemeinschaft, vor allem zum Klimaschutz. Green IT ist zur Selbstverständlichkeit geworden. Dabei kann man explizit auswählen, welche Informationen geteilt werden sollen und für welchen Zweck. Die erhobenen Daten fließen anonymisiert in die regionale Open-Data-Plattform ein und dienen vor allem der Verbesserung der öffentlichen Verwaltung.

Über Mittag hat sie ihren monatlichen Austausch mit den Kolleginnen und Kollegen von ShareHub, einer Plattform für Wissenstransfer zwischen verschiedenen europäischen Regionen. Lana verantwortet die Pressearbeit und den offiziellen Blog. Viele Kommunen und Organisationen in diesem Verbund haben sich mit dem Aufkommen des „Spatial Web“ der strikten Trennung von virtueller und physischer Welt verschrieben. Für den ShareHub-Blog arbeitet Lana gerade an einem Stück, in dem sie die Umsetzbarkeit und Hindernisse dieser Trennung beschreibt. So werden beispielsweise in Schulen, der Pflege oder der medizinischen Versorgung häufig reale menschliche Kontakte virtuellen Begegnungen vorgezogen. Ein Kollege aus Tschechien berichtet, dass in seiner Region die Nutzung von Virtual-Reality-Anwendungen auf berufliche Treffen und die Steuerung von automatisierten Produktionssystemen beschränkt ist. Allerdings führe dies zu der erheblichen Herausforderung, dass immer weniger Menschen die intuitiven, aber dennoch komplexen Anwendungen nutzen können und wollen. Eine andere Kollegin aus Deutschland beschreibt eine neue Kampagne namens heART, die vor allem auf AR-Anwendungen statt immersiver, virtueller Realität für soziale Medien setzt. Auch Lana hat davon schon erfahren. Sie trifft sich am Nachmittag mit Maria, einer Gründerin, die Teil der Initiative ist.

Maria wartet schon auf Lana in einem kleinen Café in der Nachbarschaft. Mit ihrem Start-up entwickelt sie eine Anwendung, die Informationen kontextbasiert in den Alltag einbindet und so die lokale Wirtschaft und Gemeinschaft stärken soll. Für solche Angebote wurden erst vor Kurzem die neuen 6G-Netzfrequenzen freigeschaltet. Über die vernetzte Brille oder den Ohrhörer erhalten z. B. Spazierende im Vorbeigehen Informationen über Cafés oder Geschäfte, über Freundinnen, Freunde und Bekannte in der Gegend, über mögliche Ausflugsziele oder über aktuelle Entwicklungen in der Stadt. Jede und jeder soll Inhalte beisteuern und kommentieren können. Die Idee ist ein soziales Netzwerk, das auf der realen Umgebung der Menschen aufbaut und so die Echtheit und Transpa-

renz von Informationen und Verhalten sicherstellt. Maria gibt Lana eine Brille. Lana blickt sich im Café um, sofort erscheint die Karte mit den Tagesempfehlungen. Ihr Blick schweift auf eine Baustelle auf der Straße. In ihrem Sichtfeld erscheint „Hier entsteht das neue Stadtbad. Baustand 80 %. Eröffnung Oktober 2035. Sie möchten dabei sein? Melden sie sich hier an.“ Lana blinzelt zweimal, und die Brille registriert ihre Anmeldung. Echt cool, Lana muss die Anwendung unbedingt mit ihrer eigenen Brille testen!

Wieder zu Hause lässt sich Lana auf das Sofa fallen. Heute war wieder viel los und gleich kommt ihr Sohn nach Hause ... etwas Zeit für sich könnte ihr definitiv nicht schaden. Eine Freundin redet derzeit nur noch von ihrer neuen "Find-yourself-Bubble: ein persönlicher, virtueller Raum, der ihr dabei hilft, den Arbeitstag abzuschließen und einen klaren Schnitt zum Privatleben zu machen. Dabei lässt sich auswählen zwischen Achtsamkeits- und Meditationsangeboten oder immersiven Naturerlebnissen. So ein kurzer Urwaldspaziergang wäre jetzt natürlich ein Traum ...



Neue Interaktionsmuster und Elemente der Vertrauens- und Misstrauensbildung im Future Spotlight „Ökologische Regionalisierung“

Neue Interaktionsmuster	➤ Humanoide Roboter als Lernhilfen und „Lehrende“ ➤ Digital Twins von Wohnungen zur Steigerung der Energieeffizienz und des Abfallmanagements ➤ Kommunikation mit der Umgebung durch Smart Glasses und Smart Earplugs möglich ➤ Weite Verbreitung von AR-Anwendungen und Smart Glasses
Vertrauensbildende Elemente	➤ Individuelle Begrenzung von Data-Sharing möglich ➤ Erhobene Daten fließen anonymisiert in Open-Data-Plattform ein ➤ Wo möglich, werden menschliche Begegnungen virtuellen vorgezogen
Misstrauensbildende Elemente	➤ Weitverbreitete Grundskepsis gegenüber digitalen Anwendungen

KERNBEFUNDE UND WEITERFÜHRENDE FRAGEN

2.1 Kernbefunde

» **Vertrauen ist ein zentraler Mechanismus, um Komplexität und Unsicherheit zu bewältigen und in neuen Situationen und Räumen handlungsfähig zu bleiben. Es war damit Voraussetzung, um den digitalen Raum zu erobern – und wird auch wegen der konstanten Weiterentwicklung der digitalen Sphäre weiterhin von zentraler Bedeutung bleiben.**

Ohne Vertrauen gäbe es keinen Cyberspace, kein Internet und auch keine sozialen Medien. Die rasend schnell voranschreitende digitale Vernetzung durch das Internet und Entwicklungssprünge bei Technologien wie Smartphones, Künstlicher Intelligenz, VR und Big-Data-Anwendungen stehen sinnbildlich für die konstante Weiterentwicklung der digitalen Sphäre. Eine Verlangsamung der Entwicklung ist zunächst nicht absehbar; vielmehr zeigt sich, dass digitale Technologien, Virtualisierung und

Vernetzung immer stärker mit der analog-physischen Welt verflochten sein werden. Hybride Wirklichkeiten zeichnen sich bereits am Horizont ab. Für die Einzelne und den Einzelnen führt dies (scheinbar) zu Erleichterungen und mehr Bequemlichkeit im Alltag. Informationen sind schneller und einfacher verfügbar, digitale Assistenten unterstützen bei Routinetätigkeiten und organisieren den Alltag. Die zugrunde liegenden Technologien und Algorithmen zu verstehen, nachzuvollziehen oder gar zu kontrollieren vermag die oder der Einzelne jedoch nicht. Um sich in dieser komplexen Welt weiterhin bewegen zu können, müssen Bürgerinnen und Bürger vertrauen. Das Vertrauen muss dabei nicht zwangsläufig in die Akteure hinter den Technologien und der Software gesetzt werden. So wie heute das Vertrauen in die Instanzen öffentlicher Ordnung jeder und jedem das Gefühl gibt, sich ohne Gefahr für Leib und Leben im öffentlichen Raum

bewegen zu können, könnten auch zukünftig ordnende Instanzen – neue oder bereits etablierte – das Systemvertrauen fördern und zur Interaktion und Bewegung in einer digital vernetzten Welt befähigen.

Wissensasymmetrien zwischen Nutzern und Bereitstellern digital vernetzter Technologien verzerren die Risikowahrnehmung in der digital vernetzten Welt. Das Verhältnis zwischen Mehrwert und Risiko digitaler Technologien wird falsch eingeschätzt. Diese Wissensasymmetrie könnte in Zukunft sogar noch größer werden.

Die digitale Sphäre ist von einer grundlegenden Wissensasymmetrie geprägt. Wenige Unternehmen und Expertinnen und Experten haben Einblick in Funktionsweisen, Entscheidungsprozesse und Geschäftsmodelle der digitalen Welt, die für die Mehrheit der Nutzerinnen und Nutzer weitgehend intransparent bleiben. Die Überwindung dieser Wissensasymmetrie treibt auch die Debatte um verantwortungsvolle KI.⁴⁵ Rechtliche Rahmenbedingungen werden häufig von technologischen Neuerungen überholt. Gleichzeitig sind die Risiken und Missbrauchspotenziale, die durch die Nutzung von Technologien und deren Anwendungen eingegangen werden, schwerer einschätzbar: Risiken wie der Missbrauch oder die Manipulation von Daten sind noch vergleichsweise neu. Konsequenzen und mögliche persönliche Schäden können in der Regel noch nicht anhand von Erfahrungswerten eingeschätzt werden. Kollektive Erfahrungswerte liegen nur in anekdotischer Form vor. Hinzu kommt, dass die Risiken abstrakter und weniger unmittelbar sind als Risiken für Leib und Leben im physischen Raum. In der Folge kommt es zu Fehleinschätzungen von Mehrwert und Risiko der Nutzung. Vor dem Hintergrund eines sich abzeichnenden dynamischen technologischen Wandels durch neue Technologien und Technologiekonvergenzen dürfte sich diese Wissensasymmetrie auch in der digitalen Sphäre der Zukunft nicht verringern. Im Gegenteil: Komplexe technologische Anwendungen wie Deep Learning könnten diese sogar noch vergrößern, weil bei verschiedenen KI-Algorithmen selbst für die Programmierenden die Funktionsweise nur begrenzt

nachvollziehbar ist. Die Bedeutung von Initiativen zu verantwortungsvoller KI oder Ethical Coding dürften zukünftig weiter zunehmen.

Die digitale Sphäre wird von einer Art „Convenience-Paradigma“ dominiert. Die Nutzungsbereitschaft gegenüber neuen Technologien hängt in hohem Maße von deren Mehrwert ab. Ein hoher wahrgenommener Mehrwert und Netzwerkeffekte treiben die Verbreitung voran. Nutzungsparadoxe entstehen: Technologien werden genutzt, obwohl den Betreibern und deren Absichten misstraut wird.

Das Leistungsversprechen vieler technologischer Anwendungen in der digitalen Sphäre ist ein einfacheres, bequemer Leben. Verstärkt wird dieses Phänomen durch Netzwerkeffekte. Technologien und Anwendungen, deren Mehrwert in der Vernetzung von Menschen oder Daten liegt, profitieren wiederum von Monopol- oder Oligopolstrukturen. Je mehr Personen oder Dinge vernetzt sind, desto größer ist auch für Einzelne der (wahrgenommene) Nutzen, die gleiche Lösung oder Plattform zu verwenden. Gleichzeitig sind Alternativen mit immer größerem Aufwand und Kosten verbunden. Die Folge sind Nutzungsparadoxe: Obwohl die positiven Absichten der Betreibenden infrage gestellt werden und auch Datenschutzbedenken aufkommen, ihnen also aktiv Misstrauen entgegengebracht wird, überwiegt der wahrgenommene Nutzen (oder die kritische Masse der Nutzerinnen und Nutzer) die wahrgenommenen Risiken. Die zunehmende Verbreitung von Algorithmen, die menschliches Entscheiden ersetzen oder unterstützen (z. B. digitale Assistenten), könnten diesen Prozess vorantreiben.

Auch wenn es zunächst paradox wirkt: Sicherheit schafft nicht mehr Vertrauen. Je größer die Sicherheit in einer Situation, desto weniger Vertrauen ist notwendig, um mit der verbleibenden Unsicherheit umzugehen. Je sicherer die Rahmenbedingungen, desto eher ist man jedoch geneigt, die verbleibende Unsicherheit durch Vertrauen zu „überbrücken“.

Sicherheit und Vertrauen werden häufig in einem Atemzug verwendet. Sicherheit schaffe Vertrauen.

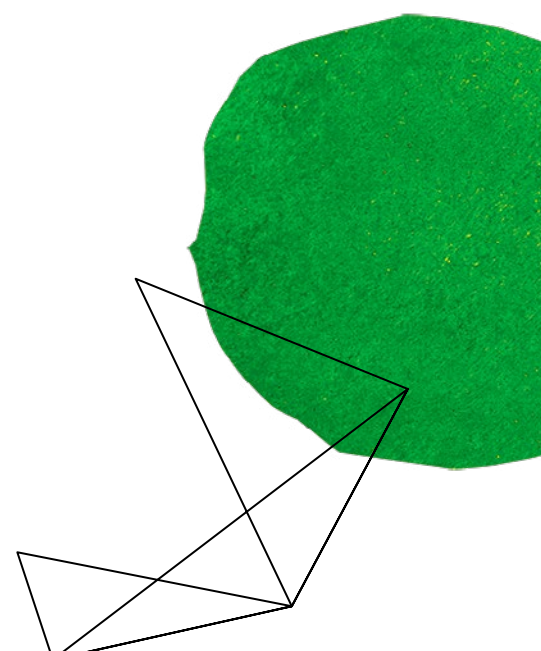
45 Siehe Kapitel 2.4.4 in der Langfassung der Studie. Verfügbar zum Download auf vorausschau.de.

Tatsächlich sind die Wechselwirkungen zwischen Sicherheit und Vertrauen hochgradig komplex. So wäre bei vollkommener Sicherheit kein Vertrauen mehr nötig. Wer sich in vollkommener Sicherheit wähnt, sei es z. B. durch das Schließen von Verträgen, durch Kontrollmechanismen oder Möglichkeiten der Sanktionierung, muss nicht vertrauen. Nichtsdestotrotz kann ein Gefühl von Grundsicherheit gegenüber größeren Risiken dazu beitragen, dass verbleibende Unsicherheiten mit einem Vertrauensvorschuss überbrückt werden. Als Sicherheit schaffende Elemente können technische Sicherheitsmaßnahmen (Cybersecurity), Intermediäre und Regulierungen Risiken in der digitalen Sphäre reduzieren. Durch mehr Sicherheit wird also weniger Vertrauen als „Schmiermittel“ benötigt. Je sicherer jedoch die Gesamtsituation ist, umso mehr „Grundvertrauen“ ist auch vorhanden.

Umgang mit humanoiden Robotern. Am stärksten gilt dies für die komplexe Interaktion mit Avataren, seien es Echtbild-Avatare oder animierte Grafiken, und Hologrammen, die vollkommen neue Formen der Begegnungen im Web 3.0 / Spatial Web / Metaversum hervorbringen. Wird es einen parallel ablaufenden Avatar-Check auf Vertrauenswürdigkeit durch virtuelle Assistenzsysteme geben? Gibt es Scoring-Systeme für Avatare? Werden wir Echtbild-Avataren mehr Vertrauen schenken als grafischen Animationen? Wie werden wir Avatare und Hologramme als vertrauenswürdig einschätzen? Diese neuen Interaktionsmuster könnten zukünftig neue Kulturtechniken der Vertrauensbildung hervorbringen, bei denen wir paradoxerweise wiederum technologische Hilfe in Form von Verifizierungs-KI-Systemen nutzen könnten.

Vertrauen ist in erster Linie abhängig von der Wahrnehmung des Gegenübers. Die Verschmelzung von Analogem und Virtuellem verändert diese Wahrnehmung und damit auch das Vertrauen. Neue Formen und neue Qualitäten des Vertrauens könnten entstehen.

Eine zunehmende Digitalisierung und Virtualisierung reduziert auch die Unmittelbarkeit von Begegnung und schränkt Wahrnehmung im Vergleich zu physischen Begegnungen ein. Wie das Gegenüber wahrgenommen wird, ist jedoch entscheidend dafür, ob man es für vertrauenswürdig hält. Technologien werden immer häufiger zum Mittler von Kommunikation und Interaktion. Dadurch werden die bestehenden Strategien, auf die man seine Entscheidung über die Vertrauenswürdigkeit eines Gegenübers gründet, immer weniger zuverlässig. Dies ist dadurch bedingt, dass man sein Urteil auf weniger Informationen gründen kann als in einer unmittelbaren Begegnung, weil visuelle, haptische und sensorische Reize reduziert werden. Gleichzeitig kann die Authentizität der wahrgenommenen Reize nicht mehr geprüft werden. Vertrauen wird also komplexer, denn auch allen Technologien und Mittlerinstanzen, die die Interaktion ermöglichen, muss Vertrauen entgegengebracht werden. Dies gilt etwa für die in den Alltagsszenarien, den Future Spotlights, beschriebenen, dort längst im Alltag angekommenen virtuellen Assistenzsystemen, die weitreichende Entscheidungen vornehmen. Oder den



2.2 Weiterführende Fragen

Aus der Studie ergeben sich viele weiterführenden Fragestellungen für den zukünftigen Möglichkeitsraum. Viele dieser Fragen sind implizit in den alternativen Zukunftsbildern, den Future Spotlights, behandelt worden. Es lohnt sich jedoch, diese nochmals explizit zu benennen.

Zunehmende Vernetzung ermöglicht es, Räume Gleichgesinnter zu finden, was wiederum subjektiv das Vertrauen in andere Menschen erhöht. Gleichzeitig droht eine verstärkte Abgrenzung gegenüber anderen Gruppen. Wie können neue digitale Begegnungsräume geschaffen werden, die die Gesellschaft zusammenführen?

Wenn Technologie zunehmend individualisiert (digitale Assistenzsysteme etc.) und humanoider wird, wird sie dann auch subjektiviert? Welchen Einfluss haben algorithmische Entscheidungen (Algorithmic Decision Making, ADM) oder KI-gestützte Systeme auf unsere Wahrnehmung der Welt, und wie könnten sie unser Handeln steuern? Inwieweit könnte der Mensch seine Handlungsautonomie verlieren?

Befördern immer menschlicher wirkende Technologien eine neue Variante von Vertrauen? Erfordern nicht mehr nachvollziehbare algorithmische Entscheidungen diese Form des Vertrauens sogar?

Hartmann spricht von „warmem Vertrauen“, und die Jugendlichen der Fokusgruppen sehen „menschliche Nähe“ als Voraussetzung für tiefe zwischenmenschliche Beziehungen. Können Avatare und holographische Darstellungen ein Äquivalent zu menschlicher Nähe werden oder die Nähe anderer Menschen vermitteln?

Durch ubiquitäre Digital Twins könnten „digitale Simulationsräume“ zur Grundlage von Entscheidungen werden. Wird sich in der Folge die Gewichtung von Empirie und „digitalen Simulationen“ verschieben?

Technologische Entwicklungen, wie neue Verschlüsselungen und Blockchains zeugen von einem verstärkten Streben nach Sicherheit in der digitalen Welt. Läuft dieses Bedürfnis nach immer mehr Sicherheit einer gewachsenen Netzkultur zuwider?

In wie weit muss Sicherheit (z. B. auch Datensicherheit) stärker regulatorisch eingefordert werden, weil Bürgerinnen und Bürger zunehmend auf digitale Infrastrukturen angewiesen sind? Müssen zukünftige Entwicklungen in der Gesetzgebung stärker antizipiert werden – also verstärkt das regeln, was sein könnte, und nicht nur das, was bereits ist? Kann dies auf nationaler Ebene überhaupt gelingen, ohne ein Splinternet zu erzeugen?

Wie könnten in einem Spatial Web neue Formen des Hatespeech oder der psychischen Gewalt (z. B. Avatarmobbing wie etwa Hasskommentare zum optischen Erscheinungsbild eines Avatars) vermieden werden?

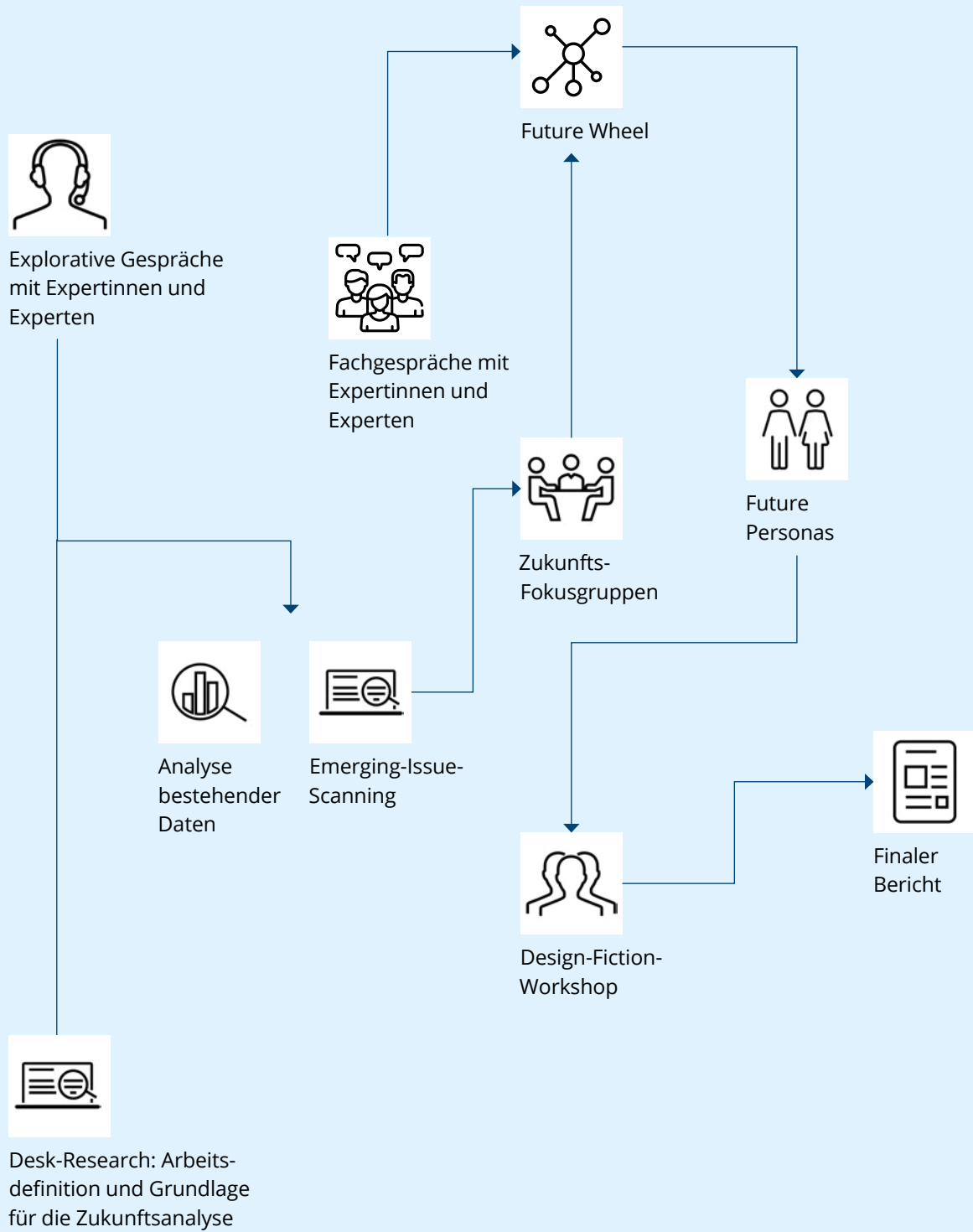


Die Studie stützt sich auf ein umfassendes Repertoire aus klassischen, wissenschaftlich fundierten Methoden wie den Interviews mit Expertinnen und Experten, der Desk Research oder der Sekundärdatenanalyse. Ergänzend kommen etablierte Foresight- und Beteiligungsmethoden zum Einsatz, wie beispielsweise ein Futures-Wheel-Workshop und die Erarbeitung von Anwendungsszenarien im Rahmen der Future Spotlights sowie ein Design-Fiction-Workshop mit jungen Designerinnen und Designern zur Diskussion potenzieller Designansätze, die auf die Bildung (oder aber die Zerstörung) von Vertrauen abzielen. Explorative und zukunfts offene sowie dialogorientierte Foresight-Methoden stützen sich folglich auf eine solide empirische Basis, die die Forschungsdynamik nachzeichnet und zugleich Entwicklungsperspektiven erfasst.

Die Hauptarbeiten an der Studie wurden zwischen Oktober 2020 und Mai 2021 durchgeführt. Zudem wurden im Winter 2021 / 22 nochmals kleinere Ergänzungen aufgenommen, um die aktuellen Debatten über die Weiterentwicklung der digitalen Welten entsprechend zu berücksichtigen.

Hinweise zur Langfassung der Studie

Weiterführende Informationen zu den beteiligten Expertinnen und Experten sowie zu den verwendeten Datensätzen und deren Analyse finden sich in der Langfassung der Studie in Kapitel 7. Die Langfassung der Studie findet sich zum Download auf vorausschau.de.

Abbildung 5: Methodische Schritte bei der Erstellung der Studie

QUELLENVERZEICHNIS

Allmendinger, J. (2019). Das Vermächtnis – Wie wir leben wollen und was wir dafür tun müssen. Ergebnisse 2019. DIE ZEIT, infas, WZB. Online unter: https://live0.zeit.de/infografik/2019/Vermaechtnis-Studie_Broschue-re_2019.pdf (Stand: 28.04.2022).

Allmendinger, J.; Wetzel, J. (2020). Die Vertrauensfrage: Für eine neue Politik des Zusammenhalts. 1. Edition. Berlin: Duden.

Ball, M. (2021). A Framework for the Metaverse. Online unter: <https://www.ballmetaverse.co/research/a-framework-for-the-metaverse> (Stand: 15.11.2021).

Beckert, J. (2018). Woher kommen Erwartungen? Die soziale Strukturierung imaginierter Zukünfte. Jahrbuch für Wirtschaftsgeschichte / Economic History Yearbook, Band 59, Heft 2. Berlin: De Gruyter Oldenbourg.

Beugelsdijk, S.; Smulders, S. (2009). Bonding and bridging social capital and economic growth. Tilburg: Tilburg University.

Bundesministerium für Bildung und Forschung (2020). Karliczek: Wir starten strategische Initiative zum Quantencomputing. Online unter: <https://www.bmbf.de/bmbf/shareddocs/pressemitteilungen/de/karliczek-wir-starten-strategi-nitiative-zum-quantencomputing.html> (Stand: 30.07.2021).

Bundesministerium der Finanzen (2020). Investitionen in ein modernes Land. Online unter: https://www.bundesfinanzministerium.de/Monatsberichte/2020/03/Inhalte/Kapitel-2b-Schlaglicht/2b-Investitionen_pdf.pdf?jsessionid=6E08353D2D76944288D4EE1691F11587.delivery1-master?__blob=publicationFile&v=3 (Stand: 11.05.2021).

Cardenas, J. C.; Carpenter, J. (2008). Behavioural development economics: Lessons from field labs in the developing world. The Journal of Development Studies. 44(3): 311–338.

CNN (2022). What it is and why does it matter? Online unter: <https://edition.cnn.com/videos/tech/2022/02/05/web-3-explainer-magic-wall-nobilo-vpx.cnn> (Stand 08.02.2022).

Contract for the Web (2019). Online unter: <https://contractfortheweb.org/> (Stand 21.04.2021).

Dator, James (o. D.). What Futures Studies is, and is not. University of Hawaii. Online unter: <http://www.futures.hawaii.edu/publications/futures-studies/WhatFSis.pdf> (Stand 25.01.2021).

Deloitte (2020). The Spatial Web and Web 3.0. Online unter: https://www2.deloitte.com/content/dam/insights/us/articles/6645_Spatial-web-strategy/DI_Spatial-web-strategy.pdf (Stand: 25.11.2020).

Die Zeit (2021). „Gewöhnung macht leichtsinnig“ (Interview mit dem Risikoforscher Ortwin Renn). Online unter: <https://www.zeit.de/wissen/gesundheit/2020-08/ortwin-renn-coronavirus-risikoforschung-angst-ansteckung-leichtsinn/komplettansicht> (Stand: 11.05.2021).

Dietvorst, B. J.; Simmons, J.; Massey, C. (2015). Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err. *Journal of Experimental Psychology: General*. 144(1): 114–126.

Dincer, O. C.; Uslaner, E. M. (2010). Trust and growth. *Public Choice* 142, 59. Berlin/Heidelberg: Springer Science+Business Media.

Enste, D.; Suling, L. (2020). Vertrauen in Wirtschaft, Staat, Gesellschaft 2020: Vertrauensindex: Europäische Länder im Vergleich. No. 5. IW Policy Paper.

Fleishmanhillard (2020). Techlash 2020. Online unter: <https://fleishmanhillard.com/wp-content/uploads/2021/03/Techlash-2020-Why-the-Technology-Sector-Needs-to-Lean-in-Now-on-Consumer-Expectations.pdf> (Stand: 25.01.2021).

Hartmann, M. (2020). Vertrauen – die unsichtbare Macht. Berlin: Fischer Verlag.

Hatak, I. (2011). Kompetenz, Vertrauen und Kooperation: Eine experimentelle Studie, Forschungsergebnisse der Wirtschaftsuniversität Wien. Berlin: Peter Lang International Academic Publishers.

Lewicki, R. J.; Bunker, B. B. (1995). Trust in relationships. *Administrative Science Quarterly*. 5(1): 583–601. Berlin: Peter Lang International Academic Publishers.

Luhmann, N. (2014). Vertrauen. Ein Mechanismus der Reduktion sozialer Komplexität. 5. Auflage. Stuttgart: UTB.

Marx, I. (2020). Klarnamenpflicht. Mit offenem Visier. Tagesschau.de. Online unter: <https://www.tagesschau.de/inland/schaeuble-klarnamenpflicht-soziale-netzwerke-101.html> (Stand 04.02.2022).

Mayer, R. C.; Davis, J. H.; Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*. 20(3): 709–734.

Merten, M. (2021). Alles wird meta. ada. Online unter: <https://ada-magazin.com/de/alles-wird-meta> (Stand: 24.11.2021).

Peter, S.; Riemer, K.; Hovorka, D. (2020). Artefacts from the Future: Engaging Audiences in possible Futures with Emerging Technologies for better Outcomes. In: *Proceedings of the 28th European Conference on Information Systems (ECIS)*. An Online AIS Conference, June 15–17, 2020.

PrognosAG; Z_punkt GmbH The Foresight Company (2020). Studie: Zukunft von Wertvorstellungen der Menschen in unserem Land. Online unter: https://www.vorausschau.de/SharedDocs/Downloads/vorausschau/de/BMBF_Foresight_Wertestudie_Langfassung.pdf?__blob=publicationFile&v=1 (Stand: 17.01.2022).

Ramge, T. (2020). Augmented Intelligence. Wie wir mit Daten und KI besser entscheiden. Ditzingen: Reclam.

Rene, G. (2019). An Introduction to The Spatial Web. Medium. Online unter: <https://medium.com/swlh/an-introduction-to-the-spatial-web-bb8127f9ac45> (Stand: 01.02.2021).

Rudelle, J. (2019). Warum das offene Internet Grenzen für die Big Techs braucht. Horizont. Online unter: <https://www.horizont.net/medien/kommentare/regulierung-warum-das-offene-internet-grenzen-fuer-die-big-techs-braucht-176915> (Stand: 21.04.2021).

Schaich, A.; Neef, A. (o. D.). Future Personas. Den Kunden der Zukunft erlebbar machen. Whitepaper Z_punkt. Online unter: <https://z-punkt.de/de/themen/den-kunden-der-zukunft-erlebbar-machen> (Stand: 25.01.2021).

Shirky, C. (2008). Here comes Everybody. The Power of Organizing without Organizations. New York: Penguin Press.

Security Affairs (2021). US Govt kicked off 'Hack the Army 3.0' bug bounty program. Online unter: <https://securityaffairs.co/wordpress/113116/security/hack-the-army-3-0.html> (Stand: 03.02.2021).

Slovic, P. (1987). Perception of Risk. Science. 236(4799): 280–285.

Speck, U. (2020). Vertrauen ist das Schmiermittel der Globalisierung. Neue Zürcher Zeitung. Online unter: <https://nzz.ch/meinung/kolumnen/vertrauen-ist-das-schmiermittel-der-globalisierung-ld.1544415> (Stand: 05.05.2021).

Stalder, F. (2017). Kultur der Digitalität. 2. Auflage. Berlin: Suhrkamp.

Süddeutsche Zeitung (2022). Am eigenen Leib. Online-Belästigung im Metaverse. Online unter: <https://www.sueddeutsche.de/kultur/metaverse-vr-virtual-reality-microsoft-sexuelle-belaestigung-1.5519527> (Stand: 15.02.2022).

Tagesschau.de (2021). Zukunftsideen der Tech-Konzerne. „Metaversum“ als Internet-Nachfolger? Online unter: <https://www.tagesschau.de/wirtschaft/technologie/metaversum-silicon-valley-internet-101.html> (Stand: 02.12.2021).

The Economist (2020). The Metaverse is coming. Technology Quarterly. Online unter: <https://www.economist.com/technology-quarterly/2020/10/01/the-metaverse-is-coming> (Stand: 01.02.2021).

The Economist (2021). What is the metaverse? Online unter: <https://www.economist.com/the-economist-explains/2021/05/11/what-is-the-metaverse> (Stand: 10.11.2021).

TH Köln (2015). „Nutzungszahlen sind die neue Währung“ (Interview mit Isabel Zorn). Online unter: https://www.th-koeln.de/hochschule/nutzungszahlen-sind-die-neue-waehrung_22617.php (Stand: 10.05.2021).

Wittenhorst, T. (2019). Gegen Hetze im Netz. Schäuble fordert Klarnamen-Pflicht. Heise.de. Online unter: <https://www.heise.de/newsticker/meldung/Gegen-Hetze-im-Netz-Schaeuble-fordert-Klarnamen-Pflicht-4425451.html> (Stand: 19.10.2021).

Zak, P. J.; Knack, S. (2001). Trust and growth. The economic journal. 111(470): 295–321.

Ansprechpartnerinnen und Ansprechpartner Kampagnenbüro Strategische Vorausschau

Torstraße 49, 10119 Berlin

Telefon: +49 30 818777158

E-Mail: kontakt@vorausschau.de
presse@vorausschau.de

Internet: vorausschau.de

Ansprechpartnerinnen und Ansprechpartner Zukunftsbüro

Michael Astor

Supervision und Gesamtprojektleitung
bei der Prognos AG

Anna Hornik

Projektleitung bei der Prognos AG

Dr. Christian Grünwald

Projektleitung bei der Z_punkt GmbH

Impressum

Herausgeber

Prognos AG
Europäisches Zentrum für Wirtschaftsforschung
und Strategieberatung
Goethestraße 85, 10623 Berlin

Z_punkt GmbH
The Foresight Company
Schanzenstraße 22, 51063 Köln

Stand

Mai 2022

Druck

BMBF

Text

Prognos AG
Z_punkt GmbH

Gestaltung

familie redlich AG – Agentur für Marken und Kommunikation
KOMPAKTMEDIEN – Agentur für Kommunikation GmbH

Diese Publikation wird von der Prognos AG und der Z_punkt GmbH kostenlos herausgegeben. Sie ist nicht zum Verkauf bestimmt und darf nicht zur Wahlwerbung politischer Parteien oder Gruppen eingesetzt werden.

