

# Digitalisierung als Prozess

## Der philosophische Blick auf die Möglichkeit allmählicher Disruption

---

Armin Grunwald

**Abstract:** *This article analyzes the phenomenon of »gradual disruptions« at a societal level. These are upheavals with considerable to dramatic damage potential that do not occur unexpectedly and suddenly, but build up gradually until they finally lead to the sudden disruption of familiar constellations. This phenomenon can be observed in many areas of processes of digital transformation. The construction of »digital twins« of practices from the analogue world plays a special role in transformation through digitalization. This leads to considerable acceleration effects and increases both the possibility and the risk of disruptive change. The thesis is that the digital transformation that is currently being pursued and permitted is gradually creating and consolidating dependencies that are also gradually becoming more vulnerable. Dependencies on digital technology infrastructure that have become total are latent disruptions that can be described in epistemic, communicative, ethical and pragmatic dimensions. Particularly affected are practices that depend on the preservation of slow deliberation, correctable learning processes and trust in responsible design, for instance, democratic governance. The universal need to adapt to digital-technological dependency means a loss of future in the sense of a space that is amenable to creative shaping. This critical philosophical thesis counters the popular optimistic narrative that digitalization is the key to opening up the future.*

**Keywords:** *digital transformation; disruption; digital twin; technological dependency; loss of future*

### 1. Digitalisierung als Narrativ und Prozess

Die Digitalisierung überformt, so die gängige Redeweise, praktisch alle Bereiche der Lebensgestaltung, individuell wie kollektiv. Gemeinsam mit innovativen Nutzungsideen und Geschäftsmodellen eröffnet der digitaltechnische Fortschritt in rascher Folge neue Handlungsoptionen (z.B. Neugebauer 2018): Mustererkennung durch Big Data Analytik, Beschleunigung von Innovationsprozessen, individuali-

sierte Dienstleistungen, Roboter als künstliche Assistenten, lernende Algorithmen, autonome Entscheidungssysteme (ADM), selbst fahrende Autos, neue Formen der Schaffung von Kunstwerken und Texten, immer bessere Simulation menschlicher Fähigkeiten, und vieles mehr. Visionäre Erzählungen von erheblicher Reichweite über die bereits sichtbaren Effekte der Digitalisierung hinaus verleiten dazu, sie als Epochenbruch zu verstehen.

Der Begriff der Digitalisierung dient dabei auf zwei Ebenen unterschiedlichen Zwecken. Einerseits geht es deskriptiv um die Beschreibung empirisch beobachtbarer Phänomene, etwa den Einsatz digitaler Werkzeuge in Arbeitswelt, Freizeitgestaltung und öffentlicher Kommunikation sowie die Folgen dieses Einsatzes. Andererseits stellt er ein dominantes *Narrativ* zeitgenössischer Diagnostik mit überschießenden und teils visionären, teils fatalistischen und teils normativen Intentionen dar wie etwa in den Formulierungen eines ›digitalen Determinismus‹ (Mainzer 2016), der Rede von der ›digitalen Revolution‹, trans- und posthumanistischen Ideen der Überwindung defizitärer menschlicher durch eine vermeintlich perfekte digitaltechnische Zivilisation (Loh 2018; Grunwald 2019a) sowie politische Botschaften, die Gesellschaft müsse sich ›fit machen für die Digitalisierung‹.

Im Fokus dieses Beitrags<sup>1</sup> steht die Exploration des philosophischen Blicks auf die Digitalisierung als Prozess in spezifischer Hinsicht. Die zunehmende Ausstattung des individuellen und kollektiven Lebens, zusehends aber auch der gebauten Umwelt in Gebäuden, Ortschaften und Infrastrukturen mit digital funktionierenden Sensoren, Datenspeichern und Algorithmen ermöglicht diesen Prozess und stellt sozusagen die Grammatik dafür bereit. Digitalisierung *als Prozess* bezieht sich im Folgenden auf die auf diese Weise ermöglichte Transformation gesellschaftlicher Zusammenhänge, so etwa in Bezug auf Demokratie, durch die Umstellung der Wirtschaft auf eine Daten- und Wissensökonomie, neue Mensch/Maschine-Verhältnisse (Ethikrat 2023), Veränderungen der Arbeitswelt (Börner et al. 2018), der öffentlichen Kommunikation und in der Selbst- und Weltwahrnehmung von Menschen (Grunwald 2021).

Das wesentliche, die transformative Leistung erst ermöglichende Element der Digitalisierung liegt in der digitalen *Verdopplung der Welt*. Der analogen Welt aus Materie und Energie wird eine digitale Welt aus Daten, Modellen und Algorithmen zur Seite gestellt, in der Datenabbilder die Gegenstände der analogen Welt als ihre so genannten ›digitalen Zwillinge‹ in gewissen Hinsichten repräsentieren sollen. Konsumprofile sind genauso Elemente des digitalen Zwillings von Menschen wie verfügbare medizinische Daten, Daten aus Überwachungskameras oder Bewegungsprofile aus Handydaten. Diese Zwillinge sind speicher- und kopierbar, durch Algo-

1 Der vorliegende Text führt einen früheren Artikel (Grunwald 2019b) weiter und hat von Diskussionen in Workshops des CAIS-Projekts *Philosophische Digitalisierungsforschung* (2019–2022) profitiert.

rithmen zur Mustererkennung nutzbar, jedenfalls solange keine Regulierung dagegensteht, und durch Suchbefehle nach bestimmten Eigenschaften recherchierbar. Die dabei gewonnenen Erkenntnisse können in die analoge Welt rückübertragen und genutzt werden, etwa für individualisierte Werbung. Eine der Visionen vieler Digitalfirmen ist es, möglichst vollständige digitale Zwillinge aller analogen Objekte zu erzeugen, diese im Hintergrund mit schnellen Algorithmen auszuwerten und die Ergebnisse in der analogen Welt für geschäftliche oder politische Zwecke zu nutzen. Diese Operationen in der Welt der digitalen Zwillinge sind im Vergleich zur Auswertung in der analogen Welt dramatisch beschleunigt. Der Wegfall vieler Dämpfungsfaktoren der analogen Welt, etwa durch die räumlichen und zeitlichen Dimensionen vieler Prozesse, führt zu erheblichen Beschleunigungseffekten und erhöht sowohl Möglichkeit als auch Risiko disruptiver Veränderung. Dieser Thematik vor allem gelten die folgenden Überlegungen.

Philosophie und Ethik werden zur Orientierung in der und für die Digitalisierung nachgefragt. Einerseits geht es um Analysen zu und Antworten auf konkrete ethische Herausforderungen wie etwa in Computereethik, Datenethik, Internetethik oder Maschinenethik (Misselhorn 2018), behandelt häufig auch in Ethik-Kommissionen mit spezifischem Auftrag (z.B. Ethik-Kommission 2017). Andererseits stehen grundsätzliche philosophische Fragen in der Diskussion, so etwa zu Selbstverständnis und Zukunft des Menschen angesichts Künstlicher Intelligenz mit schnell wachsenden Fähigkeiten (z.B. Mainzer 2016; Nida-Rümelin/Weidenfeld 2018). Philosophische Herausforderungen der Digitalisierung erschöpfen sich also nicht in Aufgaben für Angewandte Ethik, sondern eröffnen auch Anfragen an handlungstheoretische, demokratietheoretische, bewusstseinsphilosophische, anthropologische und technikphilosophische Reflexion.

Im Folgenden behandle ich spezifische Konstellationen, die im Zuge der raschen Digitalisierung immer wieder Gegenstand wissenschaftlicher, philosophischer und öffentlicher Debatten sind. Das Thema, in dieser Formulierung vielleicht paradox klingend, sind *allmähliche Disruptionen* auf gesellschaftlicher Ebene. Damit sind Umbrüche mit erheblichem bis dramatischem Schadenspotential gemeint, die nicht unerwartet und plötzlich auftreten wie eine weltweite Pandemie oder ein Angriffskrieg, sondern die sich allmählich aufbauen, bis sie schließlich zum Zerschlagen vertrauter Konstellationen führen (Abschn. 2). Hintergrund ist die Beobachtung, dass zentrale Krisenphänomene der Gegenwart wie Klima- und Umweltprobleme, allmählich einkehrende und sich verfestigende Abhängigkeiten mit ebenso allmählicher Vulnerabilitätssteigerung und die Krise der Demokratien nicht plötzlich hereinbrechen, sondern sich mit vielen Vorzeichen langsam aufbauen und erst allmählich große bis dramatische Ausmaße annehmen. Dieser Typus potentieller und allmählicher Disruption lässt sich, so die These, in vielen Bereichen der Digitalisierung als Prozess erkennen (Abschn. 3). Schließlich werden Anfragen *allmählicher Disruption* an die Philosophie in den Blick genommen (Abschn. 4).

## 2. Disruption als Begriff der Zeitdiagnostik

Disruption ist in den letzten ca. zehn Jahren zu einem vielfach verwendeten Begriff geworden. Obwohl die Wortherkunft auf eher unangenehm klingende Bedeutungen verweist (lt. *disrumpere* = platzen, zerbrechen, zerreißen), geschieht dies einerseits in positiver Intention. So stehen disruptive Innovationen als technologische Sprünge oder paradigmatische Wechsel von Geschäftsmodellen hoch im Kurs der Innovationspolitik, vielfach motiviert durch die Veränderung wirtschaftlicher Verhältnisse in der Digitalisierung. Im Gegensatz zum inkrementellen, auf allmählichen Produktverbesserungen beruhenden Innovationsgeschehen zielt Disruption auf grundsätzliche Umwälzungen, in denen teils über Jahrzehnte bestehende Marktverhältnisse in kurzer Zeit umgestürzt werden. Den Gewinnern (oft »Disruptoren« genannt) stehen neue, oft globale Marktchancen offen. Auch gänzlich neue Märkte können entstehen, wie immer wieder im Rahmen der digitalen Transformation, etwa in den *social media*. Konkreter politischer Ausdruck dieses Denkens ist die im Jahre 2019 von der deutschen Bundesregierung gegründete *Bundesagentur für Sprunginnovationen* (SPRIND), die mit einem erheblichen Budget disruptive Technologie und Innovationen fördern soll, um die deutsche Volkswirtschaft zu stärken.

Der Beginn dieser Karriere des Disruptionsbegriffs liegt in der Theorie disruptiver Technologie (Bower/Christensen 1995). Diese wurde rasch auf den Bereich der Innovation ausgeweitet (z.B. Danneels 2004), indem auch von bahnbrechenden und disruptiven Innovationen gesprochen wird. Die Erwartungen an disruptive Innovation sind teils erheblich (de la Vera/Ramge 2021). Eine zentrale Annahme darin ist die fortschreitende Beschleunigung aller Innovationsvorgänge, die Regulierung und vorausschauende Politik obsolet werden lassen. Freilich sind sowohl die Begriffsbestimmungen als auch Voraussetzungen und Erwartungen umstritten (Gans 2017): »Disruption« is a business buzzword that has gotten out of control. Today everything and everyone seem to be characterized as disruptive – or, if they aren't disruptive yet, it's only a matter of time before they become so«. In dieser Kritik verflacht der Begriff der Disruption zu einem Synonym für Erfolg, was freilich an der bislang erfolgreichen Begriffskarriere nichts geändert hat.

Andererseits werden seit einigen Jahren auch *Krisenphänomene* in den begrifflichen Kontext der Disruption gestellt. Vor allem gelten Corona-Pandemie und Ukraine-Krieg als disruptive Ereignisse. Beide haben eine lange Zeit der weitgehenden Stabilität, jedenfalls im Globalen Norden, beendet und zeigen, nach verbreiteter Diagnose, den Übergang in eine Zeit der permanenten Krise an. Auf diese Weise wird der Begriff der Disruption zur Bezeichnung des Zerbrechens stabiler gesellschaftlicher Zustände mit der nicht selten befürchteten Folge dramatischer Auswirkungen eingesetzt. Indiz dafür ist die verbreitete Kommunikation katastrophischer Narrative in der öffentlichen Debatte, vor allem die Sorge vor einem Atomkrieg, der

Klimawandel als Ende der Bewohnbarkeit der Erde und das nachlassende Vertrauen in Demokratie in vielen Staaten.

In beiden erwähnten Bedeutungen meint Disruption den plötzlichen Abbruch vertrauter Konstellationen als, je nach Kontext, erwünschtes oder befürchtetes Ereignis. Stabilitätserwartungen, Kontinuitätsannahmen und Planungssicherheiten zerbrechen und lassen die Aussichten auf Zukunft in einem unsicheren Licht erscheinen. Das Platzen, Zerbrechen und Zerreißen der lateinischen Wortherkunft (s.o.) weist semantisch auf die Zeitform mehr oder weniger plötzlicher, abrupt auftretender Ereignisse hin. So gesehen erscheint die Rede von der *allmählichen* Disruption zunächst begrifflich falsch, widersinnig oder zumindest paradox.

Der nähere Blick erlaubt eine Differenzierung. Semantisch zeigen sich im Begriff der Disruption zwei Bedeutungsanteile: zum einen das *Zerbrechen* bislang stabiler Verhältnisse, zum anderen die *Schnelligkeit* dieses Zerbrechens. Während der erste Bedeutungsanteil dem Begriff etymologisch inhärent ist, kann der zweite flexibler gehandhabt werden. Zeitskalen des Zerbrechens sind dehnbar. So wird beispielsweise die Erfindung des Buchdrucks im späten Mittelalter gerne als disruptiv dargestellt – historisch gesehen erstreckte sich diese Disruption über viele Jahrzehnte der Diffusion in die damaligen Gesellschaften hinein. Zerbrechen und Abbruch können sich auch, und darauf kommt es mir jetzt an, *allmählich* über längere Zeiträume aufbauen, sich durch schwache Signale ankündigen und erst im zeitlichen Verlauf zu nur scheinbar plötzlich auftretenden Disruptionen im Sinne eines qualitativen Bruchs führen.

So wurde nach dem Beginn des Ukraine-Kriegs vielfach festgestellt, dass dieser nicht ohne Vorwarnung begann, sondern eine Vorgeschichte in Form russischer Angriffe auf Gebiete der ehemaligen Sowjetunion hatte. Viele Beispiele für Disruption mit erkennbaren, wenngleich oft nicht erkannten Vorzeichen sind aus der technischen Welt bekannt, vor allem Materialermüdung und Verschleiß. Die tägliche Belastung vieler technischer Objekte wie z.B. von Keilriemen in altmodischen Automobilen oder Brückenbauwerken, führt allmählich zu Verschleiß und Degradierung. Lange funktionieren sie dennoch verlässlich, bis der Verschleiß ein Ausmaß erreicht, bei dem das Bauteil von einem auf den anderen Moment ausfällt, dass also in dem gewählten Beispiel der Keilriemen reißt oder die Brücke einstürzt. Ein Beispiel aus der Klimadebatte sind die so genannten Kipp-Punkte (*tipping points*, s. Gladwell 2000). Bei weiterer Erwärmung könnten selbstverstärkende Rückkopplungseffekte einsetzen, die in kurzer Zeit dramatische Folgen, also eine disruptive Wirkung hätten. Das Disruptive ist in Vorgängen dieser Art also in inkrementellen und nur schwer erkennbaren Prozessen angelegt, kann lange unerkannt bleiben und dem frühzeitig intervenierenden und vorbeugenden Eingriff entgehen, jedoch potentiell weitreichende und plötzlich eintretende Folgen erzeugen. Die *Tragik* derartiger allmählicher Entwicklungen ist, so könnte man in leicht existenzialistischer

Emphase sagen, dass sich im inkrementellen Verlauf schwerwiegende Disruptionen zwar schleichend ankündigen, aber dann abrupt vollziehen können.

Mit dieser semantischen Differenzierung wird im Folgenden die Möglichkeit *allmählicher Disruption* in der Digitalisierung als Prozess betrachtet. Es soll dabei um mögliche Entwicklungen mit Schadens- oder sogar katastrophischem Potential gehen, nicht um die Frage nach (von manchen Akteuren) erwünschter disruptiver Innovation.<sup>2</sup>

### 3. Disruptive Potentiale der digitalen Transformation

Der teils sehr schnelle Erfolg vieler Entwicklungen im Rahmen der Digitalisierung, beispielsweise der globalen Expansion der *social media* innerhalb weniger Jahre ungefähr ab 2010 oder gegenwärtig die schnelle Diffusion von KI-Anwendungen, führt zu Verschiebungen in vielen Bereichen. Hierzu gehören etwa Mensch/Technik-Verhältnisse, Verantwortungsverteilungen, industrielle Produktion, Sicherheitspolitik und Überwachung, öffentliche Kommunikation und politische Meinungsbildung, Solidarität und Wettbewerb, Arbeitsmarkt- und Arbeitswelt sowie Freizeitverhalten und Medienkonsum. Einige dieser Verschiebungen bilden den Kern der teils aufgeregten öffentlichen Debatte zur Digitalisierung und vieler weitreichender Befürchtungen darin (Grunwald 2019). Diese Verschiebungen werden im Folgenden unter dem Aspekt potentieller und allmählicher Disruption diskutiert. Dabei lassen sich zwei häufig ineinander verflochtene Perspektiven unterscheiden: Verschiebungen in Mensch/Technik-Verhältnissen (3.1 und 3.2) sowie in gesellschaftlichen Konstellationen (3.3-3.5).

#### 3.1 Normierung menschlichen Handelns

Nach dem gängigen Verständnis sollen Technik und darauf aufbauende Dienstleistungen als Mittel für menschliche Zwecke dienen, Bedürfnisse befriedigen und Probleme lösen, um, so die Erzählung seit der Aufklärung, die Optionenvielfalt menschlichen Handelns zu erweitern und die Emanzipation des Menschen zu befördern. Dies ist jedoch nur die halbe Wahrheit. Denn während Technik menschliche Handlungsoptionen erweitert, führt sie simultan zu Anpassungsnotwendigkeiten unterschiedlichster Art bis hin zum Zwang (Grunwald 2022).

---

2 Freilich bedürfte auch diese Richtung einer philosophischen Kritik, ist doch die mangelnde ex ante-Einschätzbarkeit potentieller Folgen eine notwendige, mit dem begrifflichen Kern disruptiver Innovation verbundene Begleiterscheinung, was ernsthafte Fragen an die Verantwortbarkeit dieses Innovationstypus motiviert.

Technische Systeme strukturieren und regulieren menschliches Handeln, etwa durch Bedienungsanleitungen, Vorschriften und Benutzeroberflächen. In vielen Feldern ist dies trivial, wenn etwa in der Nutzung eines Spatens zum Umgraben bestimmte körperliche Bewegungen erforderlich sind oder wenn zur Bedienung einer Waschmaschine die Bedienungsanleitung zu beachten ist. Digitale Technik ändert jedoch subtil menschliches Handeln und Verhalten, möglicherweise ohne, dass dies bemerkt wird. Die Debatte um ›Software als Institution‹ (Orwat et al. 2010) hat darauf aufmerksam gemacht, dass Softwaresysteme regulierende Kraft haben können, z.B. durch die Regelung von Transaktionen oder von Zugangs- und Nutzungsrechten. So strukturieren privat geführte *social media*-Plattformen die öffentliche Kommunikation, sortieren Suchmaschinen mit von Privatfirmen entwickelten Algorithmen die Weltwahrnehmung ihrer Nutzer und strukturieren Online-Plattformen Geschäftsprozesse und Crowd-Sourcing.

Spezifische Anpassungsnotwendigkeiten entstehen im Zusammenwirken autonomer Software- und Robotiksysteme mit Menschen. In der Industrie 4.0-Welt, in der Roboter mit Menschen in der industriellen Produktion zusammenarbeiten sollen, muss aus Funktionalitäts- wie auch Sicherheitsgründen eine missverständnisfreie Kommunikation an diesen Schnittstellen gewährleistet werden, ähnlich beim autonomen Fahren im Mischverkehr mit menschlichen Verkehrsteilnehmern. Die Forderung

»Um eine effiziente, zuverlässige und sichere Kommunikation zwischen Mensch und Maschine zu ermöglichen und Überforderung zu vermeiden, müssen sich die Systeme stärker dem Kommunikationsverhalten des Menschen anpassen und nicht umgekehrt erhöhte Anpassungsleistungen dem Menschen abverlangt werden« (Ethik-Kommission 2017: 13)

ist leicht zu erheben und anthropologisch nachvollziehbar. Jedoch steht zu befürchten, dass es im realen Ablauf nicht so kommen wird, sondern dass die fortschreitende Digitalisierung das menschliche Handeln allmählich nach den Anforderungen technischer Systeme und technischer Kommunikation reguliert und normiert. Trotz allen Bemühens um menschliche Autonomie und Wahrung der Wahlfreiheit könnte es zum Gegenteil kommen, nämlich zum schleichenden und unbemerkten Verlust von Freiheiten. Als Indiz für diese Vermutung wird vor allem der Sog des Sicherheitsdenkens angeführt, wie etwa beim autonomen Fahren. Im Rahmen eines ›Schiefe-Ebene-Arguments‹ könne das Postulat der Sicherheit in der Folge ihrer technischen Durchsetzung menschliche Freiheit letztlich komplett aushebeln. Auch wenn

»[...] es dem Leitbild des mündigen Bürgers widersprechen [würde], würde der Staat weite Teile des Lebens zum vermeintlichen Wohle des Bürgers unentrin-

bar durchnormieren und abweichendes Verhalten sozialtechnisch bereits im Ansatz unterbinden wollen« (Ethik-Kommission 2017: 20),

bedeutet das dennoch nicht, dass sich das Leitbild in der Wirklichkeit durchsetzt. Ähnlich verhält es sich in der digitaltechnisch ermöglichten und durchsetzbaren Überwachung im privaten und öffentlichen Bereich, die immer wieder mit Sicherheitsargumenten gegenüber Freiheitsargumenten begründet wird. Die »allmähliche Disruption« in diesem Feld wäre ein unbemerktes Hinübergleiten in eine Welt, in der das Sicherheitsinteresse des Staates in der Hierarchie der abzuwägenden Aspekte ganz nach oben wandert und es in der Folge zu immer weiteren digitaltechnisch durchgesetzten Normierungen des menschlichen Handelns kommt, die letztlich zum Abschied vom freiheitsorientierten Individualismus hin zu einem gelenkten Kollektivismus führen könnten.

### 3.2 Abhängigkeit als latente Disruption

Moderne Gesellschaften sind bereits heute vollständig vom reibungslosen Funktionieren kritischer Infrastrukturen wie z.B. der Stromversorgung abhängig (Petermann et al. 2011). Dies gilt in zunehmendem Maß auch für digitale Infrastrukturen. Bei einem Ausfall des Internet würden Finanztransaktionen unmöglich, würde die Weltwirtschaft zusammenbrechen, wäre keine mediale Kommunikation mehr möglich, würde die medizinische Diagnostik vieler etablierter Verfahren beraubt, würden die internationalen Logistikketten stillstehen, und vieles mehr. Mit der zunehmenden Einführung von ADM-Systemen (*automated decision-making*) entsteht eine Abhängigkeit von KI-gesteuerten Systemen, die gemeinsam mit deren *black box*-Charakter und Intransparenz, aber auch aufgrund des psychologischen *automation bias* (Ethikrat 2023) eine zunehmende Abhängigkeit von diesen Systemen in entscheidungsrelevanten Kontexten wie Polizei und Sozialwesen bedeutet.

Die allmähliche Verdrängung des Bargelds ist ein aktuelles Beispiel für die Ambivalenz technischer Infrastrukturen. War zunächst der bargeldlose Zahlungsverkehr als Erleichterung für Wirtschaft und Privatpersonen eine *zusätzliche* Option neben dem Bargeld, findet ein allmählicher Übergang zu einer Welt ohne Bargeld statt (TAB 2020). Bargeld wird allmählich verdrängt, teils durch Konsumentenverhalten und Bequemlichkeit, teils durch Anreize und Regulierung aus Politik und Wirtschaft unter der Argumentation, so könnten Schwarzmarkt und Schwarzarbeit unmöglich gemacht werden, in der Pandemie verstärkt durch das Argument der prophylaktischen Kontaktlosigkeit. Sobald sich der bargeldlose Zahlungsverkehr vollständig durchgesetzt haben sollte, wie dies in einigen Ländern bereits weitgehend der Fall ist, ist die Wahlfreiheit der Bezahlpraxis verschwunden und bei Ausfall des Internets wäre kein Einkauf oder Zahlungsverkehr mehr möglich. War die bargeldlose Zahlung zunächst eine *zusätzliche* Option und erhöhte die Wahlmöglichkeiten, wurde



sie allmählich dominant, schließlich aufgrund des Verschwindens des Bargelds alternativlos und damit zum Zwang mit der Kehrseite der Abhängigkeit.

Abhängigkeiten sind für sich keine Disruption, aber sie tragen deren Keim in sich. Total gewordene Abhängigkeiten sind *latente Disruptionen*. Als Disruptionen auf Abruf bauen sie sich über wachsende Abhängigkeiten allmählich auf, können aber im Ernstfall, wenn z.B. die digitalen Techniken nicht mehr reibungslos funktionieren würden, abrupt eintretende und möglicherweise katastrophale Folgen haben. Auf ihr unbegrenztes reibungsloses Funktionieren zu setzen und Funktionsfähigkeit und Stabilität moderner Gesellschaften davon abhängig zu machen, ist allerdings eine Wette »ums Ganze« im Sinn von Hans Jonas (Jonas 1979) und entsprechend ethisch problematisch. Unerwartete Hacker-Ereignisse, ein Zusammenbruch der staatlichen Ordnung oder schwere wirtschaftliche Turbulenzen könnten auch Infrastrukturen wie das Internet betreffen und im schlimmsten Fall dysfunktional machen. Auch wenn es meist schwer, wenn nicht unmöglich ist, den Zeitpunkt in der allmählich ablaufenden Entwicklung zu erkennen, ab dem die vollständige Abhängigkeit einsetzt, dürfte dieser Punkt in Bezug auf viele digitale Infrastrukturen und Plattformen längst überschritten sein – was bedeutet, dass moderne Gesellschaften bereits im Modus dieser latenten Disruption sind.

Obwohl digitale Infrastrukturen als Mittel zu menschlichen Zwecken aufgebaut werden, verschiebt sich das Zweck/Mittel-Verhältnis allmählich: aufgrund steigender Abhängigkeit von digitaler Technik geraten Menschen in die Notwendigkeit, alles zu tun, um diese in gutem Zustand zu erhalten. Die technische Infrastruktur wird vom Mittel zum Zweck. Dieses allmähliche Umschlagen von Freiheit und Machtverhältnissen entspricht der Hegelschen Dialektik von Herr und Knecht (reformuliert nach Grunwald 2019a: 17):

»Ein Herr hat einen Knecht. Dieser Knecht muss alles für den Herrn tun. Dadurch verlernt der Herr die lebensnotwendigen Dinge. Der Herr wird abhängig vom Knecht, und schließlich wird aus dem Knecht der eigentliche Herr. Der Herr muss dann dafür sorgen, dass es dem Knecht gut geht. Fatal daran ist: Der Übergang vom Herrn zum Knecht geschieht unmerklich.«

### 3.3 Verlust der Zukunft

Digitale Techniken gelten als Synonym für Zukunft, ähnlich wie die Kernenergie im damals voller Optimismus so genannten Atomzeitalter der 1950er und 60er Jahre. Jedoch operieren digitale Techniken grundsätzlich auf Basis vergangener Daten. So bilden die digitalen Zwillinge (s.o.) immer nur eine Welt von gestern ab, z.B. indem Kundenprofile ausschließlich auf Basis vergangener Kaufakte und Konsumprozesse erstellt werden können. Digitale Zwillinge bilden grundsätzlich nur die *Vergangenheit* ihrer analogen Originale ab. Big-Data-Technologien können nur vergangene

Daten auswerten und ebenso vergangene Muster erkennen. KI-Systeme können nur an Daten aus der Vergangenheit trainiert werden, da Daten aus der Zukunft nicht verfügbar sind. Auch wenn mit Hilfe von KI und *Big Data* versucht wird, quantitative Prognosen zu erstellen, basieren diese auf Mustererkennung anhand vergangener Daten. Durch den unabdingbaren Datenbezug ist digitale Technik unentrinnbar auf vergangene Verhältnisse fixiert. Wenn Datensätze, digitale Zwillinge und durch KI aufgedeckte Korrelationen und Muster für Zukunftsaussagen genutzt werden, werden vergangene Verhältnisse auf die Zukunft übergewälzt, ihr sozusagen übergestülpt. Zukunft als ein zumindest teilweise offener Raum alternativer Pfade und Möglichkeiten wird durch eine datenbasierte Verlängerung der Vergangenheit ersetzt.

Die Digitalisierung bzw. einige ihrer Bereiche könnten auf diese Weise konservativ werden, indem sie Zukunftsgestaltung nicht an neuen Ideen, sondern an alten Daten ausrichten. Angesichts vielfacher anthropologischer Bestimmungen des Menschen als Wesen mit Zukunft und der Fähigkeit der vergegenwärtigenden Reflexion möglicher Zukünfte (z.B. Kamlah 1973), die nicht nur Verlängerung der Vergangenheit, sondern auch kreative Neuschöpfung im offenen Raum vieler Möglichkeiten sein und sogar kontrafaktischen und utopischen Charakter tragen kann, kann es hier zu einer allmählichen Disruption kommen, in der die grundsätzliche Offenheit der Zukunft zugunsten einer datengetriebenen Orientierung an der Vergangenheit in den Hintergrund tritt oder ganz verschwindet.

Verstärkt werden kann diese allmähliche Entwicklung durch den grassierenden digitalen Determinismus (Mainzer 2016). Danach ist die Digitalisierung von einer eigendynamischen Entwicklung gekennzeichnet und könne nicht nach Werten oder gesellschaftlichen Zielen gestaltet werden. Sie fahre wie ein Zug mit hoher Geschwindigkeit, den man weder aufhalten noch in seiner Richtung beeinflussen könne. Eine andere rhetorische Form beschreibt sie mit (vermeintlichen) Sachzwang-Argumenten und (ebenso vermeintlichen) Alternativlosigkeiten als ein unausweichliches Naturereignis wie etwa einen Tsunami oder ein Erdbeben. In der deterministischen Perspektive bliebe Mensch und Gesellschaft nur die Anpassung. Anpassung jedoch bedeutet ebenfalls einen Verlust von Zukunft im Sinne eines gestaltungsoffenen Raumes.

In Hinblick auf allmähliche Disruption ist hier das Phänomen selbsterfüllender Narrative zu erwähnen: wenn sehr viele Menschen davon überzeugt sind, dass die Digitalisierung eigendynamisch und nicht gestaltbar abläuft, dann werden mögliche Interventionen und Gestaltungsoptionen gar nicht erst wahrgenommen. Das technikdeterministische Narrativ würde durch diese fatalistische Haltung bestätigt und gewänne weitere Überzeugungskraft.

Der digitale Determinismus ist weder theoretisch noch empirisch haltbar, worauf vielfach hingewiesen wurde (z.B. bereits Ropohl 1982). Wissenschaftliche Erkenntnis und gesellschaftliche Haltungen oder Trends sind jedoch oft nicht im

Einklang miteinander, sondern im Widerspruch. Auch wenn von wissenschaftlicher, z.B. sozialkonstruktivistischer Seite immer wieder auf die Gestaltungs Offenheit der Digitalisierung und auf alternative Entwicklungspfade hingewiesen wird, auch wenn praktische Ansätze der Gestaltung wie etwa das Value Sensitive Design (vgl. Grunwald 2015) exploriert und erprobt werden, impliziert das nicht bereits einen Einfluss auf gesellschaftlicher Ebene. Dort kann es trotz wissenschaftlicher und philosophischer Widerlegung zum allmählichen Verlust der Zukunft als Gestaltungsraum kommen.

### 3.4 Verantwortungsdiffusion

An den ständig neu entstehenden Schnittstellen zwischen Menschen und digitalen Systemen werden Zuständigkeiten neu verteilt. Automatisierte bzw. autonome Entscheidungssysteme (ADM-Systeme), industrielle Produktion in der Kooperation zwischen Menschen und Robotern in der Industrie 4.0 und das autonome Fahren sind Beispiele. Einige soziologische Modelle sprechen angesichts dieser Konstellationen von verteilter Handlungsträgerschaft zwischen Mensch und Technik (Latour 2005; Rammert/Schulz-Schaeffer 2002), auch philosophische Theorien nehmen eine zwischen Mensch und Maschine verteilte Verantwortung an (Floridi 2016). Sie können für bestimmte Phänomene auf der empirischen Ebene durchaus einen Erkenntnisgewinn bedeuten. Über Verantwortungszuschreibung im normativen Sinne sagen sie jedoch nichts aus, da Verantwortung handlungstheoretisch nur Akteuren mit intentionalem Handlungsvermögen zukommen kann. Digitale, auch KI-gestützte Systeme verfügen nicht über Intentionen, sondern führen komplexe mathematische und statistische Operationen auf der Basis von Daten durch. Die Möglichkeit der Verantwortungszuschreibung und -trägerschaft bleibt daher, jedenfalls gegenwärtig und in der absehbaren Zukunft, Menschen vorbehalten (Ethikrat 2023).

Die Lokalisierung von Verantwortung und ihre Zuschreibung an spezifische Akteure wird allerdings durch Digitalisierung komplex. Zwar verbleiben Entscheidungen und damit Verantwortung beim Menschen, zusehends jedoch auf eine digital vermittelte Weise. Digitale Systeme und ihre Hersteller schieben sich zwischen intentional handelnde Menschen und realweltliche Effekte. Verantwortung wandert von individuellen Autofahrern oder, im Falle von militärischen Drohnen, von Soldaten zu Personen und Institutionen im Hintergrund, zu Firmen, Programmierern, Managern, Geheimdiensten, Generälen oder Regulierungsbehörden:

»Die dem Menschen vorbehaltene Verantwortung verschiebt sich bei automatisierten und vernetzten Fahrsystemen vom Autofahrer auf die Hersteller und Betreiber der technischen Systeme und die infrastrukturellen, politischen und rechtlichen Entscheidungsinstanzen.« (Ethik-Kommission 2017: 11)

Auf Basis der vorliegenden Erfahrungen mit Verantwortungszuschreibung in komplexen und arbeitsteiligen Zusammenhängen, beispielsweise in großen Unternehmen, erscheint die Aufgabe, Verantwortung bis hinein zu Schuld- und Haftungsfragen konkret festzulegen, als zwar anspruchsvoll aber machbar. Mit der Komplexität digital zwischen Menschen und Digitaltechnik verteilter Zuständigkeiten steigen allerdings die Gefahr einer allmählichen ›Verantwortungsdiffusion‹ ins Nichts und das Risiko intentionaler Verantwortungsverschleierung.

Eine wichtige Rolle hierbei spielt ein psychologischer Effekt, der spezifisch für digitale und insbesondere KI-Systeme ist: der »*automation bias*« (Saldar et al. 2020; vgl. auch Ethikrat 2023). Empirisch wurde gezeigt, dass viele Menschen algorithmisch erzeugten, auf großen Datenmengen beruhenden und mit KI-unterstützten Entscheidungsverfahren berechneten Ergebnissen stärker vertrauen als Menschen, vermutlich aufgrund von Objektivitätsunterstellungen gegenüber mathematischen Verfahren und einem Subjektivitätsverdacht gegenüber Menschen. Damit wird Verantwortung – zumindest unbewusst – den Algorithmen als »Quasi-Akteuren« zugeschrieben. Auch bei einer rechtlich, handlungstheoretisch und ethisch reflektierten Zuschreibung von Verantwortung in Entscheidungsprozessen, in denen ein KI-System normativ strikt auf *Entscheidungsunterstützung* begrenzt und menschliche Entscheider die Entscheidung treffen müssen, könnten KI-Systeme auf diese Weise allmählich in die Rolle der »eentlichen« Entscheider geraten und damit menschliche Verantwortung substantiell entleeren, sie zu einer bloß formalen Hülle degradieren.

Im Zusammenwirken mit dem Blackbox-Charakter von komplexen KI-Systemen könnten hier zusätzlich schleichend Abhängigkeiten von intransparenten und unverständlichen technischen Systemen entstehen, während die menschliche Urteilskraft aufgrund des Vertrauens in das technische System und mangelnder Nutzung der eigenen Fähigkeiten verkümmert. (Bainbridge 1983 spricht diesbezüglich von »ironies of automatization«.) Die allmähliche Disruption bestünde hier in der Kombination des Verlierens menschlicher Fähigkeiten in Bezug auf Urteilskraft und kritisches Denken mit zunehmender Abhängigkeit von den digitaltechnischen Entscheidungsverfahren, letztlich also von den Werten und Interessen der hinter diesen stehenden menschlichen Akteure in den großen Digitalkonzernen. Wenn in digitalen Beratungsangeboten, so etwa bei Gesundheits-Apps, algorithmischer Rechtsberatung oder in Finanzgeschäften nicht transparent ist, wem hier das Vertrauen entgegengebracht wird, worauf sich dieses gründet und wer welche Verantwortung trägt, können allmählich undurchschaubare und zunehmend intransparente Konstellationen entstehen, welche die Bedingungen der Möglichkeit ethischer Reflexion und nachvollziehbarer Verantwortungszuschreibung untergraben würden.

### 3.5 Abbruch von Reflexions- und Lernmöglichkeiten

Beschleunigung ist Teil des kapitalistischen Wirtschaftssystems. Sie setzt Kreativität und Innovation frei, vor allem durch Wettbewerb. Beschleunigung ist ein vielfach im Kontext der Digitalisierung diskutiertes Phänomen, war jedoch bereits vor hundert Jahren ein beachtliches Thema:

»Domestic life, political institutions, international relations and personal contacts are shifting with kaleidoscopic rapidity before our eyes. We cannot appreciate and weigh the changes; they occur too swiftly. We do not have time to take them in. No sooner we begin to understand the meaning of one such change than another comes and displaces the former.« (Dewey 1931: 54)

Die Erhöhung der Rechengeschwindigkeit, die Möglichkeit, Millionen von Optionen in kürzester Zeit durchzurechnen, die Verknüpfung kreativer Ressourcen über das Internet und die Beschleunigung von Datentransfer und Kommunikation, vieles vermittelt und weiter beschleunigt durch digitale Zwillinge, verkürzen die Innovationszyklen. Wie eingangs schon erwähnt, zieht daher seit einigen Jahren die »disruptive Innovation« als extreme Beschleunigung Faszination auf sich. Sie ist das Gegenteil allmählicher Innovationsprozesse und denkt die digital ermöglichte Beschleunigung im Extrem.

Allerdings kennt die klassische Wirtschaftstheorie auch zerstörerischen Wettbewerb. Die Beschleunigungsspirale ist nicht beliebig weit überdrehbar, sondern in Gefahr, die menschlichen und natürlichen Ressourcen zu übernutzen, aus denen sie sich speist. Eine Sorge in Bezug auf die Digitalisierung als Prozess bezieht sich auf negative und möglicherweise ruinöse Folgen immer weiterer Beschleunigung, insbesondere zur Frage, ob und wann die Beschleunigung wichtige Bedingungen der Reflexion grundsätzlich unterminieren könnte. Dies würde in Widerspruch mit den Prinzipien der Aufklärung, den Erkenntnissen der Technikfolgenabschätzung (Grunwald 2022) und den Anforderungen an nachhaltige Entwicklung stehen:

»[Für nachhaltige Entwicklung] sind institutionelle Bedingungen zu entwickeln, die eine über die Grenzen partikularer Problembereiche und über Einzelaspekte hinausgehende Reflexion von gesellschaftlichen Handlungsoptionen ermöglichen.« (Kopfmüller et al. 2001: 305)

Reflexivität meint die vorausschauende und vorsorgende Befassung mit Folgen von Handlungen und Entscheidungen auf den verschiedensten Ebenen (Beck 1986), impliziert also die Antizipation dieser Folgen bereits vor der Ausführung von Handlungen sowie die Berücksichtigung der Ergebnisse der Reflexion in den subsequenten Entscheidungsprozessen. Sie setzt das sorgfältige Bedenken und Beraten, das Ab-

wägen von Alternativen, die Suche nach dem rechten Maß und ethisch legitimierte Kriterien voraus. All dies benötigt in doppelter Weise Zeit: zum einen für die Beratungs- und Erwägungsprozesse selbst und zum anderen zur Umsetzung der Ergebnisse in praktische Entscheidungen. Allmähliche Disruption würde hier bedeuten, dass mit dem Argument kapitalistischen Wettbewerbs die gesellschaftlichen Strukturen und Fähigkeiten zur Reflexion schleichend ausgehöhlt würden. Im Narrativ eines innovationsorientierten Fatalismus unter dem Primat wettbewerblichen Denkens kann man sich Reflexion nicht mehr leisten, da ansonsten die Konkurrenz schneller ist und Marktvorteile gewinnt.

Unbegrenzte Beschleunigung stößt damit nicht nur an Grenzen der Ressourcen-Verfügbarkeit und menschlicher Gewöhnung, sondern auch an Grenzen von Vernunft und Verantwortung. Denn wie die Semantik der Disruption schon besagt, bringt Disruption es mit sich, dass über ihre Folgen im Vorhinein wenig oder nichts gewusst werden kann, da prospektives Wissen nur aufgrund von Kontinuitäten gewonnen werden kann. Komplette diskontinuierliche Vorgänge wären Sprünge in das komplett Unbekannte. Was uns berechtigt, der grenzenlosen Beschleunigungsrhetorik zu widersprechen, ist die Tatsache, dass wir Menschen als *zoon politicon* und als moralisches Lebewesen auf Nachdenken, Beratung und Dialog angewiesen sind.

#### 4. Philosophische Anfragen an potentiell disruptive Entwicklungen

Die genannten Beispiele verweisen auf die Phänomenologie allmählicher Disruptionen bzw. allmählicher Entwicklungen mit Disruptionspotential. Es zeigen sich übergreifende Muster vor allem in Bezug auf ihre (1) Epistemologie, (2) Beurteilung, (3) Pragmatik und (4) Kommunikation.

##### *(1) Epistemologische Dimension: wie allmähliche Disruption erkennen?*

Zur Phänomenologie allmählicher, sich schleichend aufbauender Entwicklungen gehört ihre meist schlechte Erkennbarkeit. Dies ist vor allem in ihren frühen Phasen eine Herausforderung, wenn die Datenlage schlecht ist und bestenfalls schwache Signale erkennen lässt. Aufgrund der schwachen Evidenz dieser Daten und fehlender Sensibilität für das erst langsam entstehende Disruptionspotential kann es schwer sein, systematische Forschung zur Abklärung des Sachverhalts zu motivieren und entsprechende Budgets zu mobilisieren. Vieles bleibt zunächst tendenziell spekulativ. Bloß *mögliche* Entwicklungen mit Disruptionspotential wechselwirken in einer unbekannten Zukunft mit anderen bloß *möglichen* Entwicklungen auf eine ebenfalls unbekannte Weise, sodass eine epistemologische Gemengelage hoher Komplexität und Unsicherheit entsteht. Erst im Fortschreiten der Entwicklung, also näher am Eintreten einer Disruption, wird die epistemische Evidenz größer,

weil die Effekte sichtbarer und das wissenschaftliche Verständnis der Zusammenhänge tiefergehend werden. Im Wissenszuwachs über den Klimawandel als einer derartigen allmählichen Disruption war dies in den letzten vierzig Jahren immer wieder zu beobachten.

Entsprechend ist in frühen Phasen möglicher Disruption der Auftrag von Hans Jonas (Jonas 1979) nicht oder kaum einlösbar, belastbares Wissen über die ›Fernwirkungen‹ der Digitalisierung zu gewinnen. Valides Folgenwissen mit der Möglichkeit konsequentialistischer Beurteilung gibt es zwar zu vielen Einzel- und Teilfragen der Digitalisierung, zu technischen Innovationen, Produkten, Dienstleistungen, Geschäftsmodellen und Regulierungsoptionen, jedoch noch kaum über langfristige Folgen allmählicher Verschiebungen. Erkenntnistheoretische Herausforderungen richten sich hier vor allem an die Epistemologie der Korrelation angesichts der Möglichkeiten, mit selbst digitalen Mitteln der Modellierung, des Data Mining und der KI schwachen Signalen disruptiver Entwicklungen möglichst belastbar auf die Spur zu kommen.

## (2) *Ethische Dimension: wie einordnen und beurteilen?*

Die epistemologische Gemengelage hat unmittelbare Folgen für die Bewertung und Einordnung der nur allmählich sichtbar werdenden Entwicklung. Die wissenschaftsimmanent naheliegende Schlussfolgerung, dass Ethik sich zurückhalten solle, bis besseres Wissen verfügbar ist (Nordmann 2007), verbietet sich angesichts der hohen Relevanz möglicher Disruption. Angesichts begrenzter Ressourcen müssen verschiedene langsam ablaufende Entwicklungen oder befürchtete Ereignisse miteinander verglichen und nach Dringlichkeit abgestuft werden. Priorisierungen und Dringlichkeitseinschätzungen, hinter denen normative Kriterien und Relevanzen stehen, z.B. hinsichtlich des Leitbilds nachhaltiger Entwicklung, hängen jedoch mit der Evidenz des Wissens zusammen. Ein bloßer Verdacht reicht auch dann nicht für eine hohe Priorisierung mit z.B. daraus resultierender Ressourcenallokation, wenn der Verdacht im Falle seiner Bewahrheitung in eine nach anerkannten Maßstäben verhängnisvolle und auf jeden Fall abzuwendende Entwicklung führen würde. Hier kommt es also zu schwierigen Aufgaben der Bewertung der Lage und ihrer Einordnung im Vergleich mit anderen Entwicklungen.

Entsprechend ist die Befassung mit allmählichen Verschiebungen im Rahmen der Digitalisierung der Angewandten Ethik methodisch vorgelagert. Sie kann als *explorative Philosophie* bezeichnet werden (Grunwald 2010), die sich technikphilosophisch, anthropologisch und gesellschaftstheoretisch den teils spekulativen Einschätzungen dieser Entwicklungen sowie möglichen Folgen für Mensch und Gesellschaft widmet, um kommende Debatten in begrifflicher, konzeptioneller und methodischer Hinsicht *vorzubereiten*. Analog zur hermeneutischen Erweiterung der Technikfolgenabschätzung (Grunwald/Hubig 2018) steht die Exploration



derjenigen Quellen technischer, sozial- und wirtschaftsethischer Fragen und Themen im Vordergrund, die mit allmählichen Verschiebungen in Gesellschaft und im Mensch/Technik-Verhältnis korrelieren und die eine bessere ethische Einordnung erlauben, z.B. durch Vergleich mit anderen Feldern.

### *(3) Pragmatische Dimension: wie handeln?*

In frühen Phasen möglicher Disruption kommt es zu Fragen nach Konsequenzen für das Handeln zwischen proaktiv intervenierender Prävention und dem Abwarten auf verbesserte Datenlagen und klarere Diagnosen. Hier ist das aus der Technikfolgenabschätzung (Grunwald 2022) bekannte Dilemma (Collingridge 1980) zu beachten: Zwar ist in frühen Phasen prinzipiell der weitere Gang der Dinge noch weit offen und daher besser zu beeinflussen als in den späteren, wenn die Konstellation durch Pfadabhängigkeiten bereits stark verfestigt ist. Allerdings ist dann das erforderliche Folgenwissen über die unter dem Verdacht allmählicher Disruption stehenden Entwicklungen zwangsläufig hochgradig unsicher oder fehlt ganz. Statt belastbarer Prognosen oder wenigstens plausibler Szenarien liegen üblicherweise nur mehr oder weniger spekulative Erwartungen oder auch Befürchtungen vor, deren epistemischer Gehalt oft nicht gut einschätzbar ist (Grunwald 2013).

In dieser Hinsicht stellen sich allmähliche Disruptionen als Radikalisierung, ja Extremform des Collingridge-Dilemmas dar: Vor dem disruptiven Ereignis ist nichts oder kaum etwas über die Folgen bekannt, so dass keinerlei vorsorgende oder proaktive Handlungen aktiviert werden können, danach jedoch ist das Ereignis geschehen und es ist zu spät für vorsorgende Maßnahmen. Es käme nur noch nachträgliche Reparatur in Frage. Das oben erwähnte Gebot der Reflexivität wäre somit verletzt. Angesichts dieser Situation stellt sich die Frage, wann die Evidenz eines Verdachts hinreichend groß ist, um intervenierende Maßnahmen zu legitimieren, Budgets zu mobilisieren, ggf. Freiheiten durch Regulation einzugrenzen etc. Das war die zentrale Konfliktthematik in den ersten Jahrzehnten der Debatten zum Klimawandel.

### *(4) Kommunikative Dimension: wie sprechen?*

Die schlechte Erkennbarkeit allmählicher Entwicklungen und die Schwierigkeiten ihrer Bewertung haben Folgen auch für ihre Kommunikation. Geringe Evidenz des Wissens macht Kommunikation anfällig für Ideologie und Spekulation (Grunwald 2002). Auf der einen Seite kommt es zu Verharmlosung und Abwiegeln mit dem Verweis, man solle doch warten, bis bessere Daten vorliegen und sich die Evidenz erhärtet hat, statt vorschnell Ressourcen für Maßnahmen zu verschwenden. Auf der anderen Seite werden schwache Signale als harte Entwicklung verstanden und in die Zukunft extrapoliert, mit oft dramatischen Ergebnissen bis hin zu Befürchtungen



einer raschen Disruption. Gegenseitige Vorwürfe um Übertreibung, Ideologie, Spekulation, Verharmlosung und Schönrednerei, Leichtsinn, Verantwortungslosigkeit oder permanente Bedenkenträgerei sind die Folge. Während der Corona-Pandemie ließen sich diese Kommunikationsprobleme mannigfaltig beobachten. Immer wieder schien es zwischen dramatisierender Übertreibung auf der einen und Verharmlosung der Risiken auf der anderen Seite kaum noch einen Weg der vermittelnden Vernunft zu geben.

## 5. Wider den Fatalismus

Das verbreitete Unbehagen an vielen Aspekten der Digitalisierung in der öffentlichen und insbesondere intellektuellen Debatte kontrastiert auf merkwürdige Weise mit dem oft schnellen und durchschlagenden Markterfolg neuer digitaler Angebote und Dienstleistungen. Hier könnte von einem gespaltenen öffentlichen Bewusstsein gesprochen werden: Auf der einen Seite werden Komfort und Bequemlichkeit digitaler Applikationen fast blindlings wertgeschätzt, andererseits werden Sorgen über die damit verbundenen Entwicklungen geäußert. Der Ursprung dieses Unbehagens liegt nach den vorgetragenen Überlegungen in der Kombination von zwei methodisch voneinander unabhängigen, in der Sache freilich verbundenen Einstellungen. Das Zusammenwirken technikdeterministischer Muster (s.o. Abschnitt 3.3) mit Extrapolationen allmählicher Verschiebungen in die Zukunft führt zu Fragen des Typs, wohin denn das alles führen solle und wer das noch beeinflussen könne. Die damit verbundene fatalistische Tendenz stellt in Frage, ob eine notwendige Bedingung für die Möglichkeit einer Ethik der *Digitalisierung als Prozess* überhaupt noch erfüllt ist, nämlich das Vertrauen, diesen Prozess durch ethische Reflexion mitgestalten zu können.

Ethik der Digitalisierung als Prozess muss, wenn sie ihrer Aufgabe nicht nur als akademische Disziplin, sondern auch in der öffentlichen Debatte wahrnehmen will, dieses Unbehagen (s.o. Abschnitt 3) als Sorge und damit *als Ressource* verstehen: als Quelle notwendiger Sensibilisierung auf dem Weg in eine immer stärker digitalisierte Welt. Es besteht zunächst Aufklärungsbedarf in zwei Richtungen. Zunächst muss der ›digitale Determinismus‹ (Mainzer 2016) überwunden werden. Aus digitaltechnisch ermöglichten Potentialen wird nicht in determinierter Weise eine zukünftige Gegenwart, sondern hierüber befinden Entscheidungen auf den unterschiedlichsten Ebenen, die wertbehaftet und der ethischen Explikation und Kritik zugänglich sind (van de Poel 2009). Statt vorauseilender Anpassung an die vermeintlich eigendynamische Entwicklung der Digitalisierung geht es um ihre Gestaltung im Hinblick auf ein ethisch reflektiertes, gesellschaftliches Wollen. In der Digitalisierung *als Prozess* müssen Mitgestaltungsmöglichkeiten eingefordert und umgesetzt werden. Des Weiteren ist Aufklärungsarbeit dahingehend zu leisten, dass Vi-

sionen und Extrapolationen Erzählungen in der Immanenz der Gegenwart sind, aber keine Tatsachen aus der Zukunft beschreiben (Luhmann 1990; Grunwald 2013).

Philosophie der Digitalisierung bedarf daher der Kooperation mit den empirischen Sozialwissenschaften, der Rechts- und Politikwissenschaft, der Psychologie und der Technikfolgenabschätzung. Die durch die Digitalisierung als Prozess erzeugten normativen Unsicherheiten hängen, jedenfalls was die in diesem Beitrag betrachteten allmählichen Disruptionen betrifft, untrennbar mit empirisch fundierten Einschätzungen ihrer Relevanz, Ausprägung und Dramatik zusammen. Zwar ist es wichtig und richtig, einen ›digitalen Humanismus‹ (Nida-Rümelin/Weidenfeld 2018) oder ›digitale Mündigkeit‹ (Grunwald 2019a) zu fordern – es darf aber mit der Forderung nicht sein Bewenden haben.

## Literatur

- Bainbridge, L. (1983): Ironies of Automatization, in: *Automatica*, 19(6), 775–779.
- Beck, U. (1996): Das Zeitalter der Nebenfolgen und die Politisierung der Moderne, in: Beck, U.; Giddens, A.; Lash, S. (Hg.): *Reflexive Modernisierung. Eine Kontroverse*, Frankfurt a.M.: Suhrkamp, 19–112.
- Börner, F.; Kehl, C.; Nierling, L. (2018): Chancen und Risiken mobiler und digitaler Kommunikation in der Arbeitswelt, Berlin: Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag. [[www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Arbeitsbericht-ab174.pdf](http://www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Arbeitsbericht-ab174.pdf)] (Zugriff: 25.04.2019).
- Bower, J.L.; Christensen, C.M. (1995): Disruptive Technologies. Catching the Wave, in: *Harvard Business Review*, 69, 19–45.
- Collingridge, D. (1980): *The Social Control of Technology*, New York: Pinter.
- Danneels, Erwin (2004): Disruptive Technology Reconsidered. A Critique and Research Agenda, in: *Journal of Product Innovation Management*, 21(4), 246–258.
- de la Vera, R.L.; Ramge, T. (2021): *Sprunginnovation. Wie wir mit Wissenschaft und Technik die Welt wieder in Balance bekommen*, Berlin: Econ.
- Dewey, J. (6. Auflage 1931): *Science and Society*, in: Ders. (Hg.): *The Later Works 1925–1953*, Carbondale (IL): Southern Illinois University Press, 53–63.
- Ehrenberg-Silies, S.; Peters, C.; Wehrmann, C.; Christmann-Budian, S. (2022): *TAB2020 Welt ohne Bargeld – Veränderungen der klassischen Banken- und Bezahlssysteme. TAB Kurzstudie Nr. 2*, Berlin: Büro für Technikfolgenabschätzung beim Deutschen Bundestag [<file:///C:/Users/wb9296/Downloads/TAB-Kurzstudie-kso02-1.pdf>].
- Ethik-Kommission autonomes und vernetztes Fahren (2017): Endbericht. [[https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf?__blob=publicationFile)] (Zugriff: 23.3.2022).

- Ethikrat – Deutscher Ethikrat (2023): Mensch und Maschine. Herausforderungen durch Künstliche Intelligenz, Berlin [<https://www.ethikrat.org/fileadmin/Publicationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>].
- Floridi, L. (2016): Faultless responsibility. On the nature and allocation of moral responsibility for distributed moral actions, in: *Philosophical Transactions of the Royal Society*, 374, 20160112 [<https://doi.org/10.1098/rsta.2016.0112>].
- Gans, J. (2017): The disruption dilemma, Cambridge (MA): The MIT Press. [Announcement [<https://mitpress.mit.edu/9780262533621/the-disruption-dilemma/>] (Zugriff: 22.03.2023)].
- Gladwell, M. (2000): The Tipping Point. How Little Things Can Make A Big Difference, New York u.a.: Little, Brown and Company.
- Grunwald, A. (2002): Zwischen Präventionsnotwendigkeiten und Alarmismus: Problemwahrnehmungen in der Nachhaltigkeitsdiskussion, in: Ministerium für Umwelt und Verkehr Baden-Württemberg (Hg.): Kommunikation über Umwelt Risiken zwischen Verharmlosung und Dramatisierung, Stuttgart/Leipzig: Hirzel, 87–101.
- Grunwald, A. (2010): From Speculative Nanoethics to Explorative Philosophy of Nanotechnology, in: *NanoEthics*, 4(2), 91–101.
- Grunwald, A. (2013): Modes of Orientation Provided by Futures Studies. Making Sense of Diversity and Divergence, in: *European Journal of Futures Studies*, 15:30 (2014) [<https://doi.org/10.1007/s40309-013-0030-5>].
- Grunwald, A. (2015): Technology assessment and design for values, in: van den Hoven, J.; Vermaas, P.E.; van de Poel, I. (Hg.), *Handbook of Ethics, Values, and Technological Design. Sources, Theory, Values and Application Domains*, Dordrecht: Springer, 67–86.
- Grunwald, A. (2017): Abschied vom Individuum – werden wir zu Endgeräten eines global-digitalen Netzes? In: Burk, S.; Hennig, M.; Heurich, B.; et al. (Hg.): *Privatheit in der digitalen Gesellschaft*, Berlin: Duncker&Humblot, 35–48.
- Grunwald, A. (2019a): Der unterlegene Mensch. Die Zukunft der Menschheit im Angesicht von Algorithmen, Robotern und Künstlicher Intelligenz, München: RI-VA-Verlag.
- Grunwald, A. (2019b): Digitalisierung als Prozess. Ethische Herausforderungen inmitten allmählicher Verschiebungen zwischen Mensch, Technik und Gesellschaft, in: *Zeitschrift für Wirtschafts- und Unternehmensethik*, 20(2), 121–145.
- Grunwald, A. (3. Auflage 2022): *Technikfolgenabschätzung. Einführung*, Baden-Baden: Nomos.
- Grunwald, A. (2023): Disruption in Zeitlupe, in: Brand, C.; Meisch, S.; Frank, D.; Ammicht Quinn, R. (Hg.), *Ich lehne mich jetzt mal ganz konkret aus dem Fenster. Festschrift für Thomas Potthast*, Tübingen: Tübingen Library Publishing, 447–456.

- Grunwald, A. (Hg.) (2021): Wer bist du, Mensch? Transformationen menschlicher Selbstverständnisse im wissenschaftlich-technischen Fortschritt, Freiburg: Herder.
- Grunwald, A.; Hubig, C. (2018): Technikhermeneutik. Ein kritischer Austausch zwischen Armin Grunwald und Christoph Hubig, in: Friedrich, A.; Gehring, P.; Hubig, C.; Kaminski, A.; Nordmann, A. (Hg.), *Jahrbuch Technikphilosophie*, 2018 [Arbeit und Spiel], Baden-Baden: Nomos, 321–352.
- Jonas, H. (1979): Das Prinzip Verantwortung. Versuch einer Ethik für die technologische Zivilisation, Frankfurt a.M.: Suhrkamp.
- Kamlah, W. (1973): Philosophische Anthropologie. Sprachkritische Grundlegung und Ethik, Mannheim: Bibliographisches Institut.
- Kopfmüller, J.; Brandl, V.; Jörissen, J.; Paetau, M.; Banse, G.; Coenen, R. (2001): Nachhaltige Entwicklung integrativ betrachtet. Konstitutive Elemente, Regeln und Indikatoren, Berlin: Edition Sigma.
- Latour, B. (2005): Reassembling the social. An introduction to actor-network-theory, Oxford: Oxford University Press.
- Loh, J. (2018): Trans- und Posthumanismus, Hamburg: Junius.
- Luhmann, N. (1990): Die Zukunft kann nicht beginnen. Temporalstrukturen der modernen Gesellschaft, in: Sloterdijk, P. (Hg.), Vor der Jahrtausendwende. Berichte zur Lage der Zukunft, Frankfurt a.M.: Suhrkamp.
- Mainzer, K. (2016): Wann übernehmen die Maschinen?, Heidelberg: Springer.
- Misselhorn, C. (2018): Grundfragen der Maschinenethik, Stuttgart: Reclam.
- Neugebauer, R. (Hg.) (2018): Digitalisierung. Schlüsseltechnologien für Wirtschaft und Gesellschaft, Heidelberg: Springer.
- Nida-Rümelin, J.; Weidenfeld, N. (2018): Digitaler Humanismus. Eine Ethik für das Zeitalter der Künstlichen Intelligenz, München: Piper.
- Nordmann, A. (2007): If and Then. A Critique of Speculative NanoEthics, in: *NanoEthics*, 1(1), 31–46.
- Orwat, C.; Raabe, O.; Buchmann, E.; et al. (2010): Software als Institution und ihre Gestaltbarkeit, in: *Informatik-Spektrum*, 33, 626–633.
- Petermann, T.; Bradke, H.; Lüllmann, A.; Poetzsch, M.; Riehm, U. (2011): What happens during a blackout. Consequences of a prolonged and wide-ranging power outage, Norderstedt: BoD – Books on Demand.
- Rammert, W.; Schulz-Schaeffer I. (2002): Technik und Handeln, in: Dies. (Hg.), Können Maschinen handeln? Soziologische Beiträge zum Verhältnis von Mensch und Technik, Frankfurt a.M.: Campus Verlag, 11–64.
- Ropohl, G. (1982): Zur Kritik des technologischen Determinismus, in: Rapp, F.; Durbin, P. (Hg.), *Technikphilosophie in der Diskussion*, Wiesbaden: Vieweg & Teubner, 3–18.

- Safdar, N.M.; Banja, J.D.; Meltzer, C.C. (2020): Ethical considerations in artificial intelligence, in: *European Journal of Radiology*, 122. [<https://doi.org/10.1016/j.ejrad.2019.108768>].
- van de Poel, I. (9. Auflage 2009): *Values in Engineering Design*, in: Meijers, A. (Hg.), *Philosophy of Technology and Engineering Sciences*, Amsterdam: North Holland, 973–1006.
- van den Hoven, J.; Vermaas, P.; van de Poel, I. (Hg.) (2015): *Handbook of Ethics, Values, and Technological Design*, Dordrecht: Springer, 67–86. DOI: 10.1007/978-94-007-6970-0.

