DISCUSSION

Values and Ethics in Information Systems

A State-of-the-Art Analysis and Avenues for Future Research

**Sarah Spiekermann · Hanna Krasnova · Oliver Hinz · Annika Baumann ·
Alexander Benlian · Henner Gimpel · Irina Heimbach · Antonia Köster ·
Alexander Maedche · Björn Niehaves · Marten Risius · Manuel Trenz**

1 Motivation

Sarah Spiekermann, Hanna Krasnova and Oliver Hinz

In late 2019 about a dozen BISE chairs from the German-speaking community met around ICIS to discuss the ethical challenges arising from the current construction, deployment, and marketing of Information Systems (IS). It turned out that many were and are concerned about the negative implications of IS while at the same time being convinced that digitization also supports society for the better. The questions at hand are what the BISE community is contributing in terms of solutions to the societal challenges caused by IS, how it should handle politically and socially ambiguous developments (i.e., when teaching students), and what kind of relevant research questions should be addressed. In the aftermath of the initial get-together, an online retreat took place in the late summer of 2020, during which all colleagues presented their current research projects. It turned out that BISE scholars have a very strong interest and track record in this area, and consequently, the plan was born to publish this discussion paper as well as a BISE Special Issue dedicated to the issues of "Technology for Humanity" (Spiekermann-Hoff et al. 2021).

In the following, 12 colleagues interested in this community effort have contributed their reflections and viewpoints on fostering technology in humanity's interest. Hence, this discussion paper is a collection of individual views and contributions. Starting from the design perspective, Alexander Maedche reminds us that one of the core interests of IS is to improve the well-being of users, and describes how he and his team are using machine learning techniques to support the adaptiveness of IS. He notes, however, that at a higher level of abstraction, well-being is a broad concept. Hence, "when designing IS for well-being it is not straightforward to define the actual

S. Spiekermann (✉)
Vienna University of Economics and Business, Vienna, Austria
e-mail: sspieker@wu.ac.at

H. Krasnova · A. Köster
University of Potsdam, Potsdam, Germany

O. Hinz
Goethe University Frankfurt, Frankfurt am Main, Germany
e-mail: ohinz@wiwi.uni-frankfurt.de

A. Baumann
Weizenbaum-Institut, Berlin, Germany

A. Benlian
TU Darmstadt, Darmstadt, Germany

H. Gimpel
University of Hohenheim, Stuttgart, Germany

I. Heimbach
WHU — Otto Beisheim School of Management, Vallendar, Germany

A. Maedche
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

B. Niehaves
University of Siegen, Siegen, Germany

M. Risius
University of Queensland, Brisbane, Australia

M. Trenz
University of Göttingen, Göttingen, Germany

design goal and measure specific well-being outcomes." The question of design goals is the one that many scholars in the field of ethical and social computing may seek to answer from the standpoint of human values. Values are conceptions of the desirable and principles of the ought-to-be that can and should be identified in the early phases of system requirements analysis (as well as business model development). In her contribution, Sarah Spiekermann argues that these values can be the "design goals" sought for humanity. Hence, IS innovators should strive to foster positive values through solutions beyond technical quality (e.g., reliability or security) and the achievement of economic goals. Examples are the values of health, trust, and transparency that some BISE colleagues work on and present here. Friendship, dignity, knowledge, and freedom are other high intrinsic values that are worth protecting. However, they are currently undermined by some instances of IS which instead provide a breeding ground for hate speech and fake news, which fuel envy, limit human autonomy, and expose users to surveillance capitalism.

Building on the idea of value-based system design advanced by Alexander Maedche and Sarah Spiekermann, the following contributions describe the values that the authors deem important in their work and on which they have already published extensively. In particular, health (Alexander Benlian and Henner Gimpel), trust (Annika Baumann and Björn Niehaves), and transparency (Irina Heimbach, Oliver Hinz, and Marten Risius) are discussed. These individual papers define the problem space of each of these values, give hints to relevant literature sources, and outline research questions that they believe are worth tackling.

In the next step, four contributions address the grand value-related challenges of an IT-enabled society: Alexander Benlian and Henner Gimpel outline how the "gig economy" can lead to social challenges and value destruction in digitally transformed work environments. Manuel Trenz presents the challenges surrounding surveillance capitalism. He argues that IS researchers should be at the forefront of guiding and monitoring the development of ethical personal data markets, informing regulatory bodies and facilitating an informed, consent-based release and use of personal data for the social good. Antonia Köster and Marten Risius describe what happens when data is used for voter manipulation and targeting. They further describe the processes that empower online extremism. Finally, Annika Baumann, Irina Heimbach, and Hanna Krasnova end this discussion paper by reminding us that we are seeing an evolutionarily influential transition of human beings into "digitized individuals." Despite an array of positive implications, this transition also implies changes in individual behavior and perceptions about oneself, others, and the world at large, which can be

unintended and potentially detrimental. Beyond personal harm, adversarial micro-changes at an individual level may accumulate and ultimately "collectively contribute to major issues affecting society at large."

2 Designing Information Systems for Well-being

Alexander Maedche

"Ensuring healthy lives and promoting well-being for all at all ages" is the third United Nations Sustainable Development Goal. Health is not only defined here by the absence of illness or diseases but also considers physical, psychological, and social factors linked to well-being. Well-being is a complex, multi-dimensional construct and is grounded in different schools of thought: First, the subjective well-being perspective follows a hedonic approach and emphasizes happiness, positive emotions, and the absence of negative emotions, as well as life satisfaction (Diener 1984; Diener et al. 1999; Kahneman et al. 1999). Second, the eudaimonic perspective on well-being draws on Aristotle's definition of happiness as being in accordance with virtue. Thus, eudaimonic well-being focuses on optimal psychological functioning through experience, development, and having a meaningful life (Ryff and Keyes 1995; Ryan and Deci 2001). Third, these two core perspectives can be complemented by a social dimension of well-being that emphasizes such aspects as social acceptance, contribution, and integration (Keyes 1998).

With the rapid digitalization of all areas of life and work, designing IS for well-being has become increasingly important. However, in this context, IS should be seen as a double-edged sword: they can have positive as well as negative impacts on individual well-being. For example, online games or streaming services aim at triggering positive emotions and user experiences (UX), potentially contributing to hedonic well-being. Furthermore, these services enable new forms of social connectedness that may contribute to social well-being. Modern IS in the workplace follow the same or similar principles. They enable the virtualization of work independent of time and space, personal development, and globally connected employee networks. Thus, one may argue that IS are a key facilitator of well-being in the workplace and at home. However, the underlying business model of digital service providers for private life consumption is often advertisement-based and therefore focuses on maximizing user attention, use, and time on site. Reflecting on this development, scholars have called for attention to be treated as a scarce commodity (Davenport and Beck 2001). Similarly, virtualized workplaces erase previous boundaries between work and private life and enable 24/7 availability of the

workforce. Furthermore, multi-tasking and overuse of IS in private and work life can lead to a loss of autonomy and control, to stress, or even to an addiction. IS, then, can have negative impacts on well-being.

Against this background, designing for well-being has received increasing attention in research in the last decade. Beyond accessibility, usability, and UX, well-being oriented design has established itself as an important criterion of a "good design" (Calvo and Peters 2014) in the Human–Computer Interaction (HCI) field. Following the positive psychology paradigm, research streams such as "positive technology" or "positive computing" have encouraged the investigation of technology designs for well-being. In parallel, the commercial market of well-being technology devices in different forms (apps, wearables, etc.) is growing rapidly. Well-being features–e.g., managing time spent, notification blockers–are increasingly added as core capabilities of IS used in the workplace and at home.

Designing IS for well-being can follow two complementary strategies: First, well-being can be increased through behavior changes of users by means of digital intervention designs. Self-tracking can help in understanding current behavior and the corresponding well-being states. On this basis, positive psychology interventions that have proved themselves able to positively influence well-being (Bolier et al. 2013) can be realized in the form of digital interventions. Second, IS can adapt to prevent negative outcomes on well-being during use. User-adaptive IS are a class of IS where the interaction with users is based on monitoring, analyzing, and responding to user activity in real-time and over longer periods of time. The underlying idea is that huge amounts of data about the users themselves, their tasks and contexts, are collected using different types of sensor technology. User activity is captured by sensors, e.g., in the form of electrocardiography (ECG) signals which are collected through wearable technology or eye-movement signals captured by eye-tracking technology. The collected data is then processed using machine learning techniques in order to automatically detect the affective-cognitive states of users; individualized user-centered IS adaptations can be designed on this basis. One example is intelligent notification management through dynamic notification adaptations, which may be triggered based on the analysis of user, task and context data collected by sensors. In the recently completed research project "Kern", funded by the German Ministry for Work and Social Affairs, we investigated the design of flow-adaptive notification systems for the workplace. In a first step, the flow was predicted based on ECG signals in combination with self-reported subjective data using supervised machine learning. Subsequently, the flow classifier was leveraged to design a flow-adaptive notification system to protect employees from incoming messages during flow states in real time. The field experiment with 30 employees using the system in a (home-)office environment has delivered promising results (see Rissler et al. 2020).

To conclude, it is important to emphasize that when designing IS for well-being, it is far from a straightforward task to define the actual design goal and measuring specific well-being outcomes. In light of this, it is first of all important to clearly conceptualize and break down the broad well-being concept into more specific constructs in order to clarify the nomological network. In addition, one has to be clear about whether the goal is to change user behavior or to adapt the IS to the existing behavior. Finally, in order to successfully design IS for well-being, it is necessary to involve all relevant stakeholders, ranging from users, designers and developers, to companies that provide and/or use technology, as well as governance actors in the society. With users' well-being a central priority, the existing business models of digital service providers need to be challenged and new legal boundaries enforcing specific designs should be considered. Moreover, since the design of user-adaptive IS requires access to privacy-sensitive data that may conflict with other human values, designing for health and well-being needs to become the subject of a broader public debate on societal values and their prioritization. The journey towards designing IS for well-being in work and private spheres has just started–and we still have a long way to go.

3 Value-based Engineering for Human Well-being

Sarah Spiekermann

An important way to work towards human and social well-being in system design is to construct systems in a more ethical way. Ethical system design can draw its inspiration from the Aristotelian approach to ethics. This classic perspective emphasizes the importance of human values and virtues worth striving for in order to reach "eudaimonia", which might be described as a state of self-actualization or well-being (see the contribution of Alexander Maedche, "Designing Information Systems for Well-being," above). In his Nicomachean Ethics, (Aristotle 2000) focused on human virtues he deemed important, such as courage, kindness, justice, and many others–all values of human conduct that are undermined by current IS. Value-based Engineering aims to avoid these adverse effects on virtues. It is about anticipating, assessing, and formulating system requirements that go beyond efficiency, profit and speed, as well as those non-functional value requirements that have already earned their place in traditional system design, such as usability, dependability or security.

In the past five years, values and virtues have been put forward in a myriad of listings by companies and global institutions (Jobin et al. 2019), as well as by legislators. An example is the ALTAI list of the EU Commission's High Level Expert Group on artificial intelligence (HLEG of the EU Commission 2020). Values called for in such listings include transparency, fairness, non-maleficence, responsibility, privacy, human autonomy, trustworthiness, sustainability, dignity, and solidarity. However, using such preconfigured value listings to build an ethical system is not sufficient. In fact, a lot of valid criticism has been voiced concerning the straightforward application of these lists in practice. This is because ethics is essentially contextual, and there is a risk of applying the logic of the list to problems that don't fit these lists. More importantly, value listings do not tell engineers how to effectively embed and respect values in the technical system design. "The truly difficult part of ethics—actually translating normative theories, concepts and values into good practices …is kicked down the road like the proverbial can. Developers are left to translate principles and specify essentially contested concepts as they see fit, without a clear roadmap for unified implementation" (Mittelstadt 2019 p. 503).

Some scholars in the field called "machine ethics" (Anderson and Anderson 2011) have taken up this challenge and made attempts to bring ethics closer to system-level design by developing ethical algorithms. These algorithms typically follow a simple weighing of harmful and beneficial decision consequences (an approach called Utilitarianism), or they follow a duty ethical approach where specific human principles are optimized (e.g., fairness). The work on ethical algorithms culminated in MIT's "Moral Machine Experiment" to inform the evasive actions of autonomous cars (Awad et al. 2018) with the help of "trolley economics." A shortfall of Machine Ethics (including the Moral Machine Experiment) is that the vast majority of its proposed algorithms is based only on utilitarianism or on duty ethics (Tolmeijer et al. 2020). In contrast, Virtue Ethics, which is one of the most timely and influential streams of moral philosophy, seems to be completely ignored when ethical algorithms are conceived (Tolmeijer et al. 2020). This is a pity considering its recognized importance for technology design (Vallor 2016). Virtue ethics aims to foster the value of human conduct. Its goal is to strengthen humans. Instead of aspiring to maximum algorithmic autonomy, virtue ethical algorithms would probably follow a different design paradigm, one that relies more on human interaction and that strives to improve the human decision maker instead of taking decision autonomy away from him or her. For this reason, it is regrettable that so little research is attributed to this form of potential Machine Ethics.

Machine Ethics and the intense public debate of MIT's Moral Machine Experiment has also taken attention away from what I would argue are much more relevant challenges for a more ethical IS world. These challenges include, among others, system-of-system control issues, data quality issues, sustainability issues, human control issues, as well as the ignorance of a system's long-term 2nd order value effects on stakeholders. Some of these grander challenges of ethical system design are anticipated by scholars working in value-sensitive design (Friedman and Kahn 2003) or participatory design (Frauenberger et al. 2015); however, the problem is that these works often get bogged down in the identification of very specific problems for which its authors find very specific technical solutions, but lack a generally applicable methodology to address value challenges across contexts.

Here, I believe, an important research opportunity opens up for the IS community, which has been historically strong in method design and modeling. One might say that a proper system development life cycle (SDLC) model is missing for ethical and value-based engineering. The only rigorous approach currently available to fill this gap is the IEEE 7000^{TM} standard (IEEE 2021). IEEE 7000^{TM}, which is at the heart of what has been called Value-based Engineering. The standard provides engineers with a clear system design and development framework; or in other words, an ethical SDLC (Spiekermann 2021). It uses various ethical theories to elicit relevant values, and subsequently prioritizes these with the help of corporate or industry value listings. It then derives a new artifact called "ethical value requirement" (EVR) that is translated into system requirements. System requirements are derived with the help of risk assessment.

Whether Value-based Engineering with IEEE 7000^{TM} will be taken up on a large scale remains to be seen. Early trials, however, show that if companies really want to build and operate their IS in an ethical way they will need to consider their "value proposition," which means not only changing the technology they build but also their business models (see the contribution of Alexander Maedche on "Designing Information Systems for Well-being in Private and Work Life" above). True value creation is not a matter of technology design alone but also of strategy, corporate culture, and companies' willingness to forgo some profit for the sake of community, integrity, and accountability.

4 Selected Values of Outstanding Importance for IS Research

4.1 Health and Well-being

Henner Gimpel and Alexander Benlian

Health and well-being are intrinsically and instrumentally valuable (Frankena 1973; Ryan et al. 2008) and are closely intertwined. The World Health Organization suggests that "health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO 1948, preamble). Philosophers have criticized this definition for being too all-encompassing (e.g., Callahan 1973). Nevertheless, health is not only statistical normality but also a normative ideal (Nordenfelt 1993). It is a prerequisite for flourishing and living a fulfilling life. For this reason, it is no surprise that "good health and well-being" is one of the United Nations' Sustainable Development Goals.

There is ample evidence for IS both promoting and weakening health and well-being. Let us consider the dark side first: a side effect of digitalization is the impairment of psychological and physical health (Gimpel and Schmied 2019). Interruptions by information and communication technologies (ICTs), techno-overload, blurred boundaries between the workplace and the private domain and other digital stressors often result in exhaustion, cognitive and emotional irritation, and physical illness (Chen and Karahanna 2018; Benlian 2020; Califf et al. 2020). Pirkkalainen and Salo (2016) reviewed two decades of research on this dark side of ICT use. Among the four phenomena they identified, three impaired health and well-being: technostress, IT addiction, and IT anxiety. These phenomena of ICT use may have detrimental influences on individuals, for example, in the form of loneliness (Matook et al. 2015), burnout (Srivastava et al. 2015), or diseases of the musculoskeletal or cardiovascular system (Gimpel et al. 2019).

On the bright side, ICTs also seem to promote certain aspects of health and well-being. Healthcare is a shining example of how digitalization can achieve higher efficiency and effectiveness. Examples at the individual level are the support of patient self-management by m-health apps (Gimpel et al. 2021) and health education and disease prevention (Kirchhof et al. 2018). The interaction of patients and providers via patient portals improves health outcomes (Bao et al. 2020). At the organizational level, effective use of ICT affords improved efficiency and effectiveness in healthcare processes (Burton-Jones and Volkoff 2017; Gimpel and Schröder 2021). At the societal level, ICT supports public health as, for example, witnessed in the COVID-19-pandemic, where ICT aided the containment of infections via physical distancing, working from home, and contact tracing (Adam et al. 2020; Trang

et al. 2020), as well as analysis, modeling, prediction of the pandemic, and managing vaccination campaigns (Klein et al. 2021). Chen et al. (2019) conducted a bibliometric study of health IS research from 1990 to 2017. They identified major research themes, such as "Clinical Health IS," "Administrative Health IS," and "Consumer Health IS," that are covered in many research papers. The premise remains beyond the realm of Health IS that individual assistance systems and other ICTs can support users' eudaimonic well-being by helping them in their pursuit of virtues and excellences (e.g., via provision of product information and context information for ethical consumer decisions), by abetting continuous reflection on goals and actions (e.g., via self-tracking of behavior and goal achievement), by encouraging self-affirming attitudes and self-knowledge (e.g., via online self-help communities for patients with rare diseases), and by promoting exercise of reason and free will (e.g., via provision of health information to allow for a more informed and balanced discussion with healthcare professionals). However, for each of these potential positive effects, there are contrarian examples. Thus, to what extent this claim is true certainly deserves more research attention (see also the contribution of Annika Baumann, Irina Heimbach and Hanna Krasnova on "Digitization of the Individual" below).

While we have many case examples of the beneficial effects of ICT on health and well-being in specific contexts, we lack a unifying and overarching theoretical perspective on these effects. Thus, we should continue behavioral and design-oriented work on situated observations or instantiations and substantive theories. Simultaneously, we should work towards more abstract mid-range or potentially even grand theories of how ICT may promote health and well-being. Regarding the dark side of digitalization, more research is needed to identify and conceptualize the risks and side effects of digitalization. Furthermore, we should leverage our competencies in design-oriented work to envision preventive measures that might mitigate or nullify these adverse effects (see the contributions of Alexander Maedche and Sarah Spiekermann above).

4.2 Trust in Automation

Annika Baumann and Björn Niehaves

In recent decades, our lives have undergone a tremendous transformation, with automation increasingly permeating professional and private contexts. At the heart of automation are algorithms that represent "a sequence of unambiguous instructions for solving a problem, that is, for obtaining a required output for any legitimate input in a finite amount of time" (Levitin 2003, p. 3). Algorithms provide the basis for machine learning and artificial

intelligence, which use the underlying instructions either learned via input data or explicitly programmed. Algorithms work across multiple areas of our lives and range from viewing personalized feeds on social media (Lazer 2015) to potentially riding in autonomous cars in the near future (Choi and Ji 2015). With users increasingly relying on automation in private and professional settings, trust constitutes a critical component (Glikson and Woolley 2020) as it is one of the primary drivers to adopt the technology and for an individual to autonomously follow suggested actions (Benbasat and Wang 2005; McKnight et al. 2011; Freude et al. 2019).

Two conceptualizations of trust are currently prevalent in the context of user interaction with technological artifacts. The first conceptualization aligns trust with the more human-like trust dimensions such as integrity, competence, and benevolence (Benbasat and Wang 2005). A second perspective incorporates technological particularities using more system-like dimensions such as reliability, functionality, and helpfulness (McKnight et al. 2011). Importantly, how trust shapes the boundaries of human-automation-interaction seems to depend on several factors, including human character, the underlying automation itself, and the surrounding environment where the interaction takes place (Schaefer et al. 2016). Thus, the socially constructed meaning of terms associated with automation influences individuals' expectations of technological characteristics, potentially resulting in cognitive biases and erroneous assumptions regarding the system (Felmingham et al. 2021). Consequently, vital pre-conditions for a successful collaboration between humans and technology, like trust, are already shaped before an interaction occurs. Nevertheless, since trust has a dynamic element (McKnight et al. 1998), it changes with the experiences made upon interacting with automation. Overall, trust between humans and technology appears to be a multi-faceted, time-sensitive phenomenon that needs further investigation, with specific consideration of the nature of its initial development and its course over time.

State-of-the-art research discusses both negative and positive implications of automation. On the bright side, research discusses the economic capabilities and associated success chances of automation (Pasquale 2015). For example, it has been shown that automating algorithms can provide more accurate predictions than humans in various contexts (Cheng et al. 2016; Kleinberg et al. 2017). Thus, automation can offer a fertile ground for economic gains across industries. Furthermore, the algorithm-enabled large-scale analysis of data seems to support the tackling of global challenges such as climate change (Rolnick et al. 2019). At the same time, the dark side of automation and algorithmic decision-making has been increasingly in the spotlight of scholarly attention (O'Neil 2016; Eubanks

2018). For example, automation has been shown to create biases towards specific entities (e.g., Lambrecht and Tucker 2019; see also the contribution of Irina Heimbach, Oliver Hinz and Marten Risius on "Bias, Fairness, and Transparency" below), and to facilitate extremists' views through the algorithm-induced creation of echo chambers on social media platforms (e.g., Kitchens et al. 2020; see also the contribution of Antonia Köster and Marten Risius on "Fake News and Online Extremism" below).

While research into how individuals, organizations, and society interact with automation is gaining traction, several research gaps remain. As algorithmic automation increasingly establishes itself as a new norm, future studies need to shed more light on the underlying mechanisms that are at play when users are interacting with it. As user perceptions play out between the poles of algorithm aversion (Dietvorst et al. 2015; Jussupow et al. 2020) and algorithm appreciation (Logg et al. 2019), obtaining a more in-depth understanding of the factors influencing user attitudes towards algorithms appears especially critical. For example, just like their human counterparts, algorithms are imperfect; that is, they may and do err, as no system reaches a level of complete perfection (Martin 2019). These mistakes, however, may severely diminish trust towards automation, leading to changes in individual attitudes and perceptions in the short and long term (e.g., Dietvorst et al. 2015; Prahl and Van Swol 2017). Hence, further investigation into how trust can be repaired after such instances of failure constitutes another promising avenue for future research.

4.3 Algorithmic Bias, Fairness and Transparency

Irina Heimbach, Oliver Hinz and Marten Risius

Against the background that artificial intelligence-based predictions are said to be often faster, cheaper, more reliable, and better scalable than predictions made by humans (Mei et al. 2020), artificial intelligence technologies have found their way into businesses in virtually all industries (McAfee et al. 2012), influencing and transforming many of the societal decisions that we make today (Cowgill 2018). However, there is also the risk that decision-making supported or automated by algorithms may unintentionally and unexpectedly shape societal outcomes for the worse (see Rahwan et al. (2019) for a discussion). The issues of bias, fairness, and transparency relate to the core of IS research.

Such biases can be caused by four problems: First, the data for training can be biased. Second, the model of the algorithm itself may be a possible cause for discrimination. Third, the presentation form of the information given by the algorithms can lead to unfair decisions. Finally, the user trying to use the system can come up with a biased or

misinformed decision. Policymakers try to address these potential problems by prescribing high degrees of transparency and explainability.

Researchers and practitioners point to an increasing amount of evidence that indicates how the broad use of algorithms can lead to an inferior treatment of already disadvantaged parts of society, thereby contributing to even more societal tensions, a phenomenon frequently referred to as algorithmic discrimination (Sweeney 2013; Ensign et al. 2017; Lambrecht and Tucker 2019; Obermeyer et al. 2019). Reported examples are autonomous recruitment systems with a gender bias (Mann and O'Neil 2016) or jurisdictional decision support systems suffering from a racial bias (Polonski 2018). Biased or discriminatory decision-making resulting from defective algorithms or data is a prototypical example for research following an imperative technical approach (Sarker et al. 2019). This line of research considers technology as the major antecedent to social outcomes and human decision-making. At the same time, IS researchers should acknowledge that biased data is also the result of real-world discrimination. It reflects how humans design organizational processes. Biases in algorithms may (unknowingly) be introduced through the developers' background and upbringing. This view conceptualizes bias and fairness issues as a result of the interplay between socio-technical components and, hence, is prototypical for IS Research (Sarker et al. 2019).

Regulators and researchers have identified transparency as a key to avoid bias and ensure fair algorithmic decision-making. However, even if we were able to openly obtain access to relevant algorithms and data, there would still be natural barriers to transparency that need to be overcome. First, there is an issue of how to even assess the degree to which algorithm-based decisions are biased. Relating to this issue is the question of what corrective actions to undertake (e.g., which observations to ex-/include) to rectify the biased data. And lastly, we need to find ways to disentangle these black-box algorithms and make them explainable or at least interpretable (Kim and Routledge 2018). By overcoming these transparency issues, IS researchers can contribute to a better society and solve issues of biases and discrimination.

The interplay-oriented perspective between socio-technical components should also consider the societal implications of the increased exposure to algorithms (Sarker et al. 2019). As algorithms become increasingly ubiquitous, research needs to consider the organizational implications of personally distorted attitudes towards algorithms, such as automation bias, algorithm aversion, and the fear of technology paternalism. By addressing these issues, IS scholars can offer a substantial contribution to the betterment of society (Majchrzak and Markus 2012).

The current state of research on algorithmic transparency, fairness, and bias could, in general, be characterized by two streams of work. The first stream embraces discussion papers of a prescriptive and conceptual nature (e.g., Burrell 2016; Carlson 2017; Hosseini et al. 2018; Felzmann et al. 2019) with a special focus on developing fair, transparent, and explainable/interpretable algorithms (Rudin 2019; Rai 2020). The second stream consists of empirical studies that aim to go beyond the anecdotal evidence of algorithmic bias and discrimination (Kleinberg et al. 2017; Lambrecht and Tucker 2019) and investigate the general role of algorithms and data characteristics on trust building and the individual's attitudes towards algorithmic management (Kizilcec 2016; Lee 2018; see also the contribution by Annika Baumann and Björn Niehaves on "Trust in Automation" above). A challenge is that previous research is scattered across various disciplines and tends to focus on specific aspects of the problem while neglecting a more holistic IS view that algorithms are part of a socio-technical system, which connect tasks, humans, technology, and various levels of decision-making contexts.

IS research as a cross-sectional discipline with a long tradition of looking at IT as a sociotechnical system has a great opportunity–and the capability–to make substantial contributions to future research. First, IS theorists paired with researchers from other disciplines can elaborate on a unified and concise understanding and measurement of the concepts of algorithmic transparency and fairness. Second, IS engineers can develop system and data requirements as well as validation tests for fair and transparent algorithms. Third, behavioral IS researchers can empirically test how algorithmic characteristics (perceived transparency and fairness) affect decision-making behavior, or how they reveal human and organization-related rather than technology-centric issues that lead to potentially undesired outcomes like bias and discrimination.

5 Selected Challenges Addressable by IS Research

5.1 Digital Work, Digital Labor Markets, and Gig Economy

Alexander Benlian and Henner Gimpel

Digital, platform-mediated labor markets (e.g., Uber, Airbnb, Amazon Mechanical Turk) have permeated many economic sectors by now, provoking debate about the implications of this form of "gig" work organization. Most accounts emphasize the problematic effects on gig workers and ask questions about algorithmically controlled labor processes and the increasing precarity in such digital labor markets.

Are digital labor markets akin to *digital cages*? Scholars following such a starkly dystopian perspective ominously question what happens when the boss is an algorithm, which uses anopticon powers to continuously monitor and sanction workers (Curchod et al. 2020; Möhlmann et al. 2021). Algorithms encode managerial decisions and workplace rules into the digital tools that workers must use to complete their tasks. In this way, workers' autonomy to resist, elude, or challenge the rules that platform providers establish as conditions of participation is severely constrained. In addition, platforms individualize and alienate their labor force, depriving workers of interpersonal contact spaces that have traditionally made it possible for workers to challenge managerial authority (Kellogg et al. 2020).

Are digital labor markets *catalysts of precarity*? According to this view, platforms are a manifestation of a much broader trend that has enabled firms to externalize risks which they had previously been compelled to shoulder. The effect is to bereave the worker of long-standing social protections such as a minimum wage, safety and health regulation, retirement income, health insurance, and worker compensation (van Doorn 2017). The issue, in this view, is thus a broad socioeconomic shift that dismantles many of the labor market shelters which workers had previously enjoyed, leaving them in an increasingly vulnerable position (Schor et al. 2020).

While previous research has looked into several critical aspects of platform labor markets affecting gig workers, such as legitimacy, fairness, privacy, and marginalization (e.g., Deng et al. 2016; Wiener et al. 2020; Möhlmann et al. 2021), we believe that there are several opportunities for further research:

First, it would be worthwhile to hone in on the values and ethics inscribed into algorithms that select, match, guide, and control workers in digital labor markets (Saunders et al. 2020; see also the contribution of Irina Heimbach, Oliver Hinz, and Marten Risius on "Bias, Fairness, and Transparency" above). The encroaching influence of machine learning algorithms–which can embed and reproduce inherent biases and threaten to entrench the past's societal problems rather than redress them (Rosenblatt 2018)–is particularly evident in dynamic pricing and matchmaking between customers and workers (algorithmic matching), as well as in screening workers and guiding their behavior (algorithmic control) (Möhlmann et al. 2021; Wiener et al. 2022). The values of privacy, accountability, fairness, and freedom of access are increasingly coming to the fore of discussions around digital labor markets (Deng et al. 2016) and big digital platforms more generally (van der Aalst et al. 2019).

Second, there is an abundance of research on platform operators and service providers, yet a dearth of research on

the developers who create the matching and control algorithms at the core of the platform's operations and scalability (Vallas and Schor 2020). Developers, who are often independent contractors themselves, are exposed to severe tensions between the platform operator's goals and the gig workers' interests, and may revolt when fundamental labor rights are violated. How do developers relate to algorithmic design's potentially manipulative and invasive consequences for the workers' livelihood and cope with value conflicts on a daily basis? On a broader note, we know very little about the process by which algorithms come into being, are negotiated between different parties and updated over time. What purposes and values drive the design and operation of digital labor platforms?

Third, from the perspective of gig workers, an interesting avenue for future research is an inquiry into practices of and prospects for collective action: The various forms of resistance and "algoactivistic practices" to circumvent or subvert algorithms are particularly prevalent in digital labor markets, yet still largely under-investigated (Kellogg et al. 2020). How and why do workers comply with or deviate from algorithmic management on platforms? Can workers join forces with the customers they serve, altering the "geometry of power" (Rahman and Valentine 2021) in this triadic relationship between platform providers, customers, and workers?

5.2 Personal Data Markets and Surveillance Capitalism

Manuel Trenz

With personal data dubbed the oil of the digital economy and a key to competitive advantage, it is no surprise that there is a market for individuals' data. In fact, there has always been one, with credit reporting agencies and consumer data brokers collecting and selling data on individuals for decades. However, the scope of available, collected, and aggregated data has expanded significantly through the rise of digital platforms that now track every action individuals conduct online and even combine offline and online data sources.

As a consequence, a large number of firms have emerged that collect, aggregate, analyze, package, and sell data about individuals. This, in turn, has led to more refined targeting options with, for instance, advertisers on Facebook being able to select their target audiences based on demographics, education, financial details, life events, parental and relational status, interests, specific behaviors, etc. (Facebook, Inc. 2021). While Facebook and Google are the most visible examples of such companies, many others operate in the shadows and beyond public attention (Schneier 2015; Melendez and Pasternack 2019). For example, Acxiom Corporation offers data on more than 700 million individuals worldwide by merging data

elements from hundreds of sources (Acxiom 2018). These data include demographics, political views, economic situation, health, relationship status, activities, interests, consumption preferences, as well as psychometric characteristics. While firms benefit from improved risk prediction, targeting, or innovation opportunities, these personal data markets come with significant problems for individuals, social systems, politics, and economics (Spiekermann et al. 2015b). The most obvious issue is the question of information privacy, as individuals lose control over their data. Beyond that, detailed profiles give rise to discrimination based on race, gender, or income. Moreover, they may also simply result in wrong inferences, as these profiles can be erroneous, drawn from merged, incomplete or faulty datasets (see also the contribution of Irina Heimbach, Oliver Hinz, and Marten Risius on "Bias, Fairness, and Transparency" above). This can lead to situations where individuals are rejected from loan applications, jobs, memberships, or even denied bail without having access to the database against which they are judged and left with little options to influence or delete the data and contest the inferences collected about them. As the data in today's personal data markets is most of the time collected, aggregated, analyzed, and sold without individuals' knowledge or even without a truly informed consent, those markets have aroused the interest of regulators. Moving beyond the individual level and considering the economy as a whole, regulators are worried about the consolidation and aggregation of market power towards a few large platforms (Parra-Arnau 2018) that can exercise manipulative powers. Considering the key role of personal data in today's economy, exclusive access to these data may lead to excessive market dominance and hamper competition.

Touching upon topics such as market design and digital platforms (e.g., Bimpikis et al. 2019), (inter-organizational) data-driven innovation (e.g., Kastl et al. 2018; van den Broek and van Veenstra 2018), and information privacy (e.g., Karwatzki et al. 2017), personal data markets are a phenomenon at the center of interest of IS research. Because personal data markets are highly intrusive into the intimate lives of individuals, research on this topic requires a perspective that extends well beyond technological and economic issues.

Prior studies on personal data markets can be structured along three major research streams. The first stream has investigated the development and functioning of existing personal data markets. This includes studies that uncover and classify personal data markets and their business models (Agogo 2020; Fruhwirth et al. 2020). We also have first insights into the role of technological implementations to collect data across platforms (Krämer et al. 2019) and into strategic choices made by the data market providers (Zhang et al. 2019). A second stream of research is concerned with the valuation of personal data (Gkatzelis et al. 2015; Spiekermann and Korunovska 2017) and approaches aimed at allowing people to participate in the economic value of their information (Wessels et al. 2019). Prior studies investigating digital self-disclosure have often employed a privacy calculus perspective, which suggests that users weigh the perceived benefits against the perceived risks of sharing data as a basis for their decision-making (Dinev et al. 2015; Abramova et al. 2017). However, the rationale of benefit or value in this context is usually limited to the value that individual users gain from their consumption or participation but ignores that the economic value derived from personal data extends far beyond this. While users provide or generate the data that enables personal data markets to create value, they often play no role in determining how these data are used nor participate financially. If individuals were to actively participate in those markets, they appear to have preferences for data markets that preserve their anonymity (Schomakers et al. 2020). Such participatory personal data markets could then make use of developed mechanisms through which individuals may decide on which data to conceal at what price (Parra-Arnau 2018). The third stream of research pertains to studies on the ethical, legal, and societal impacts of personal data markets, which have mostly centered around the phenomenon of privacy itself (Spiekermann et al. 2015a). From a regulatory perspective, studies have investigated the implications of existing policies such as GDPR on the design of IS (Jakobi et al. 2020) and formulated the need for different policy interventions to protect, for instance, the weakest groups in our society (Montgomery 2015).

Given the significant economic and societal impact of personal data markets and the attention they received from regulatory bodies, media, and companies participating in the digital economy, research on personal data markets is comparably scarce. Beyond an expansion of the research streams described above, future research should investigate alternative approaches to personal data markets with the goal of making them less intrusive. From an economic perspective, this includes considering competitive strategies and business models for participatory, responsible, user-centered personal data markets to make them a sustainable alternative to current models. From a technological and regulatory perspective, we still lack effective solutions that empower individuals to take control of what data traces they leave behind, what data about them is being stored, what inferences are drawn from it, and how others use it. From a societal and ethical perspective, the implications of existing personal data markets seem to be predominantly negative. However, there also seems to be a significant social value in personal data for research, crisis management, health management, and innovation that

could be unlocked by advancing approaches to how behavioral, perceptual, or medical data can be shared ethically and responsibly.

The unique combination of technological and economic expertise should allow IS researchers to be at the forefront of guiding and monitoring the development of ethical personal data markets, informing regulatory bodies, and facilitating an informed, consent-based release and use of personal data for the social good.

5.3 Online Misinformation and Extremism

Antonia Köster and Marten Risius

Social media platforms such as Facebook, Twitter, and YouTube have transformed how information is produced, consumed, and disseminated. While empowering users with the opportunity to participate and with access to knowledge, news and opinions of others, this transformation has also been accompanied by a rise in misinformation campaigns (Lazer et al. 2018), which are frequently exploited by extremists to further their malicious agenda (Winter et al. 2020). Indeed, as any user is potentially a content creator, social media platforms have developed into a breeding ground for misinformation (Kim and Dennis 2019).

Over the past few years, the spread of misinformation has led to considerably negative individual, economic and societal implications. For example, the sharing of fake news on the COVID-19 pandemic has escalated and caused misinformation on public health matters (Laato et al. 2020), directly impacting individual well-being (Brennen and Nielsen 2020; Apuke and Omar 2021). Furthermore, fake news in combination with social media bots and micro-targeted political advertisements played a decisive role in the outcome of political events, such as the UK referendum on EU membership and the US presidential election in 2016 (Allcott and Gentzkow 2017; Liberini et al. 2020). Beyond politics, fake news can have an impact on the economy. Fake stories may attract the attention of financial market investors and thereby lead to stock market reactions (Vosoughi et al. 2018; Clarke et al. 2020). Hence, misinformation that is created and disseminated with the help of digital technologies has grave implications in the modern age.

Despite the pervasiveness of online misinformation and, in particular, fake news, we currently lack an understanding of the enabling characteristics of technology and its unique role in these processes. Some research points out that not only users generate fake news but also technology can be used to do so (Calvillo et al. 2021; Bringula et al. 2021). For instance, artificial intelligence can be used to create comments on news articles or even generate the articles themselves (Zellers et al. 2019). An emerging

technological development that is gaining attention among researchers studying misinformation are "deepfakes" (Westerlund 2019; Liv and Greenbaum 2020). Deepfake is a portmanteau of "deep learning" and "fake" and describes hyper-realistic video manipulation based on neural networks (Westerlund 2019). These deep learning algorithms enable facial mapping (i.e., swapping an individual's face in a video with another), and they have been found to have a powerful effect on creating false memories (Liv and Greenbaum 2020). At the same time, technology is not only used to create misinformation but also to detect it. Tech companies rely on machine learning or artificial intelligence to automatically detect fake news online (Woodford 2018; Newman 2020). However, users respond differently to these fact-checking services. While some perceive such services as useful and respond mindfully to identified fake news, other users do not trust these detection algorithms (Brandtzaeg et al. 2018). To further complicate the detection issue, research points towards an "implied truth effect". This describes the phenomenon that flagging some articles as fake news makes users automatically assume that other non-flagged articles are truthful–even if they have not yet been fact-checked (Pennycook et al. 2020). In this context, further research is needed to address the challenges of technologically enabled misinformation detection and creation (e.g., deepfake videos) (Shu et al. 2020).

The adverse effects of online misinformation have prompted researchers to investigate the interaction between humans and technology regarding what may explain higher susceptibility to fake news (e.g., Bryanov and Vziatysheva 2021; Sindermann et al. 2020). When summarizing the findings of scholarly articles on the topic, Bryanov and Vziatysheva (2021) identify three broad categories of determinants; namely, message characteristics, individual factors, and accuracy-promoting interventions. Several researchers have examined the importance of belief consistency and confirmation bias (Kim and Dennis 2019; Sindermann et al. 2020; Calvillo et al. 2021; Bringula et al. 2021), referring to the tendency of people to be more susceptible to fake news that aligns with pre-existing values, beliefs, or political views. Second, individual factors, including cognitive modes, predispositions, and news and information literacy differences may determine individual susceptibility to fake news. For example, lower trust in science, media and government (Roozenbeek et al. 2020), specific personality traits (e.g., lower levels of agreeableness, conscientiousness, open-mindedness, and higher levels of extraversion), as well as certain media consumption characteristics (e.g., amount of Instagram visits and more hours of news consumption) have been linked to increased susceptibility to misinformation (Calvillo et al. 2021; Bringula et al. 2021). Additionally, emotional

factors, such as higher levels of emotionality, have been linked to susceptibility to fake news (Martel et al. 2020). Finally, accuracy-promoting interventions, such as specific warnings or nudges that make individuals reflect the truthfulness of information, may influence the credibility of fake news. The problem of misinformation is further exacerbated by the social media platforms' algorithmic filtering that exposes users to news and content based on their interests and past behaviors, thereby facilitating repeated exposure to more misinformation (Kitchens et al. 2020). Further research that explores the interaction between the human or social factors and the technological aspects of fake news will help to better understand the individual's susceptibility to online misinformation.

Beyond being harmful by its very nature, online misinformation also supports online radicalization and extremism, as prominently evidenced by the recent attacks on the US capitol (Kanno-Youngs and Sanger 2021). Online extremism has become a pressing issue on social media platforms as highlighted, for example, by FBI Director Christopher Wray stating that "*social media has become, in many ways, the key amplifier to domestic violent extremism*" (Volz and Levy 2021, p.1). Digital technologies have enabled this new form of extremism that presents various unique challenges; these include the rapidly changing technological landscape (Fisher et al. 2019; Winter et al. 2020) as well as the extremists' abilities to leverage these new technologies for their malicious purposes (Conway 2017) and to respond to counter-extremist measures (e.g., platform migration) (Conway and Macdonald 2019; Nuraniyah 2019).

Currently, platform providers and third parties (e.g., government authorities, NGOs) struggle to develop and implement effective measures to combat misinformation and online extremism (e.g., Sharma et al. 2019). This is partly a result of the unique technological implications that are insufficiently understood. For example, extremism is in essence a strong deviation from something that is considered "normal" or "ordinary" (Winter et al. 2020). Online services that operate globally face region-specific understandings of humanist values and societal norms, which lead to a different understanding of what is locally understood as extreme. When proposed countermeasures to online extremism such as content moderation or account tracing and removal lack the region-specific awareness, they threaten to violate civic liberties such as the freedom of speech and personal privacy (Monar 2007; Nouri et al. 2019). Against this background, the field of IS, with its sociotechnical perspective on the interaction between social elements (individual and group norms) and the technical artifact (e.g., encrypted services, global platforms), is in a favorable position to support tech companies and regulators by comprehensively considering the interactions between technological and social components. In this way, research can help to assess and alleviate growing concerns that the increasing ability to interact online may not only lead to undetected disinformation but also contribute to more polarized societies as individuals adopt more extreme views (Kitchens et al. 2020; Qureshi et al. 2020). In this context, IS research should address this comparatively open field by shedding light on the relationship between on- and offline radicalization, how online technologies (e.g., different social media platforms, content stores, blockchain technologies) attract and support online extremist activities, and what strategies online extremists pursue to counter regulatory measures (e.g., migrating to fringe platforms, adopting peer-to-peer encrypted technologies).

5.4 Digitization of the Individual

Annika Baumann, Irina Heimbach and Hanna Krasnova

The use of digital technologies for private purposes is steadily increasing. For example, the number of smartphone users reached 3.6 billion in 2020 and is projected to grow even further (Statista 2021a). In addition, the average time spent on social media a day amounts to more than 2 h daily worldwide (Statista 2021b). The market of fitness and activity trackers that allow users to monitor their health-related behaviors (e.g., daily steps, heart rate, sleep) is booming, with "end-user spending on wearable devices" worldwide expected to reach US$81.5 billion in 2021 (Gartner 2021). With social media, smartphones, smartwatches, and other digital technologies rapidly becoming an integral part of life for consumers across the world, a growing number of stakeholders voice the need to better understand the implications of this ongoing transformation. Within this development, the paradigm of the "*digitization of the individual*" has become a central issue for IS research (Vodanovich et al. 2010; Vaghefi et al. 2017; Turel et al. 2020). At its core, it implies that digital technologies heavily influence user perceptions, cognitions, emotional reactions, and behavior (Vanden Abeele 2020), and can thereby contribute to individual and societal outcomes. However, scientific evidence on the direction and strength of the effects remains contradictory.

On the one hand, the rise of the use of digital technologies has been met with optimism. Inventions such as the use of a mobile app and wearable device have been linked to weight loss, for example (Kim et al. 2019). In the context of vulnerable groups, the growing use of smartphones has been shown to support communication, contribute to user safety, enable political and social participation (AbuJarour and Krasnova 2017), and lead to user empowerment (AbuJarour et al. 2021). Similarly, social media platforms were initially hailed for their

potential to facilitate social interaction, promote feelings of social connectedness (Koroleva et al. 2011), and enhance social capital for millions of users worldwide (Ellison et al. 2007). On the other hand, the use of digital technologies has also brought a lot of disillusionment regarding the unintended negative impact of the growing digitalization of individuals above and beyond what was expected. A journalistic investigation revealed that sensitive data provided by users during their app use (e.g., details on users' diet, exercise activities, ovulation cycle) was shared and reused for commercial purposes (Schechner and Secada 2019). Furthermore, smartphone use has been associated with a multitude of adverse effects, ranging from worsened sleep (Demirci et al. 2015; Huang et al. 2020) and deteriorated relational cohesion (Krasnova et al. 2016) to poor academic performance (Lepp et al. 2014), anxiety, and depression (Demirci et al. 2015). In a similar vein, participation in social media has been shown to be addictive (Hou et al. 2019) and has been linked to exhaustion and fatigue (Bright et al. 2015), increasingly bad mood, lower life satisfaction (Kross et al. 2013), symptoms of depression (Cunningham et al. 2021), and body dissatisfaction (Tiggemann and Zaccardo 2015). For comprehensive meta-analyses we hereby refer exemplary to the works of Appel et al. (2020), Huang (2017) and Liu et al. (2019).

ICT-enabled changes in perception at the micro-level may also collectively contribute to the emergence and proliferation of issues affecting society at large. For example, the time spent on social media has been linked to lower perceptions of inequality, which may skew redistribution preferences and affect corresponding voting behavior (Baum et al. 2020). In a similar fashion, social media use has been shown to influence users' political views, giving rise to echo chambers and contributing to polarization (Barberá et al. 2015). Furthermore, hostile expressions common on social media platforms (Crockett 2017) can potentially have an invidious effect on users, interfering with such socially relevant behaviors as free expression and participation in political processes and social life. Considering the far-reaching potential of these technologies to affect individuals and society at large, IS research has an opportunity to make a substantial contribution in the following directions:

First, the understanding of the "digitized individual" paradigm should be unified. For example, Turel et al. (2020) define a digitized individual as someone who uses at least one digital technology. In contrast, Kilger (1994) refers solely to virtual identity, while Clarke (1994) describes a "digital persona" to be a model of an individual based upon the data collected and analyzed about this person. Better alignment of terminology used in the scientific discourse and across multiple disciplines can promote more targeted exploration into this phenomenon.

Second, while individual outcomes and, as a consequence, societal outcomes of digital use can be far-reaching, the mechanisms behind them are still little understood. For example, concerns about the way social media platforms and content creators influence and bias our perceptions of reality become increasingly pressing. How, and in which specific ways, does the use of digital platforms and applications change our perception of ourselves, others, and the world around us? How do changes at the individual level translate into societal consequences? And what can be done to mitigate those detrimental developments?

Third, whereas past research has mainly focused on interpersonal differences when exploring the link between the use of digital technologies and individual outcomes, a new generation of studies advocates a stronger focus on longitudinal approaches that allow the exploration of the role of within-person differences (Beyens et al. 2020; Kross et al. 2021; Valkenburg et al. 2021b). For example, in a recent study by Valkenburg et al. (2021a, p. 56), 88% of adolescents "experienced no or very small effects" from social media usage (captured as an aggregate measure of self-reported time on WhatsApp, Instagram, and Snapchat) on self-esteem. At the same time, 4% of adolescents experienced positive effects, while 8% of adolescents experienced negative effects. Therefore, a more in-depth investigation into the within-person processes is needed. Furthermore, since a large share of studies into the individual outcomes of digital use are correlational, experimental approaches should be pursued with greater enthusiasm, as they allow causal inferences to be made about the relationships at play (e.g., Allcott et al. 2020; Brailovskaia et al. 2020; große Deters and Mehl 2013).

Fourth, methodological issues regarding the measurement of media use have been raised. Specifically, a large share of previous studies relied on retrospective self-reports to measure digital technology use by participants (e.g., in the form of constructs measuring "use," or self-reporting of time spent). However, a recently published meta-analysis raises concerns about the validity and accuracy of this approach as there seems to be only a moderate correlation between self-reported and logged metrics, concluding that the users either under- or over-report their digital media use (Parry et al. 2021). Future research should capture objective measures of platform use whenever possible, as well as strive for better operationalization of different aspects of digital media usage (Faelens et al. 2021). Importantly, in light of this, findings based on self-reported measures should be received with caution and verified for robustness with direct measures of actual behavior.

Fifth, whereas fitness and activity trackers and other mobile apps hold significant potential to improve users' health and well-being, their use may inherently conflict with such fundamental values as the individual right to

privacy, self-determination, and autonomy. Indeed, the data traces users leave behind can also be misused as part of scoring systems, or to make predictions about users' future performance at work or about future health outcomes. Hence, a more profound discussion of which values should be prioritized and how those tensions can be resolved might be necessary.

Finally, when it comes to exploring the detrimental outcomes of digital use, future research should focus on proposing and testing the effectiveness of corrective actions to mitigate the adverse effects of digital technology use for individuals (e.g., lower well-being, fatigue, technostress, overspending). At the time of writing, interventions involving digital detox are already providing encouraging evidence on the reversibility of harmful influences (e.g., Allcott et al. 2020; Brailovskaia et al. 2020).

References

Abramova O, Wagner A, Krasnova H, Buxmann P (2017) Understanding self-disclosure on social networking sites - a literature review. In: 22nd Americas conference on information systems. Boston, pp 1–10

AbuJarour S, Krasnova H (2017) Understanding the role of ICTs in promoting social inclusion: the case of Syrian refugees in Germany. In: Proceedings of the 25th European conference on information systems. Guimarães, pp 1792–1806

AbuJarour S, Köster A, Krasnova H, Wiesche M (2021) Technology as a source of power: exploring how ICT use contributes to the social inclusion of refugees in Germany. In: Proceedings of the 54th Hawaii international conference on system sciences. A virtual AIS conference, pp 2637–2646

Acxiom (2018) Annual Report 2018. In: Annu. Rep. https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_ACXM_2018.pdf. Accessed 19 Nov 2021

Adam M, Werner D, Wendt C, Benlian A (2020) Containing COVID-19 through physical distancing: the impact of real-time crowding information. Eur J Inf Syst 29:595–607. https://doi.org/10.1080/0960085X.2020.1814681

Agogo D (2020) Invisible market for online personal data: an examination. Electron Mark. https://doi.org/10.1007/s12525-020-00437-0

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31:211–236

Allcott H, Braghieri L, Eichmeyer S, Gentzkow M (2020) The welfare effects of social media. Am Econ Rev 110:629–676. https://doi.org/10.1257/aer.20190658

Aristotle (2000) Nichomachean Ethics. Cambridge University Press, Cambridge

Anderson M, Anderson SL (2011) Machine ethics. Cambridge University Press, New York

Appel M, Marker C, Gnambs T (2020) Are social media ruining our lives? A review of meta-analytic evidence. Rev Gen Psychol 24:60–74

Apuke OD, Omar B (2021) Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. Telemat Inform 56:101475

Awad E, Dsouza S, Kim R et al (2018) The Moral Machine experiment. Nature 563:59–64. https://doi.org/10.1038/s41586-018-0637-6

Bao C, Bardhan IR, Signh H et al (2020) Patient-provider engagement and its impact on health outcomes: a longitudinal study of patient portal use. MIS Q 44:699–723

Barberá P, Jost JT, Nagler J et al (2015) Tweeting from left to right: is online political communication more than an echo chamber? Psychol Sci 26:1531–1542. https://doi.org/10.1177/0956797615594620

Baum K, Köster A, Krasnova H, Tarafdar M (2020) Living in a world of plenty? How social network sites use distorts perceptions of wealth inequality. In: Proceedings of the 28th European Conference on Information Systems. A virtual AIS conference, pp 1–16

Benbasat I, Wang W (2005) Trust in and adoption of online recommendation agents. J Assoc Inf Syst 6:72–101

Benlian A (2020) A daily field investigation of technology-driven stress spillovers from work to home. MIS Q 44:1259–1300. https://doi.org/10.25300/MISQ/2020/14911

Beyens I, Pouwels J, van Driel II et al (2020) Social media use and adolescents' well-being: developing a typology of person-specific effect patterns. Commun Res. https://doi.org/10.1177/00936502211038196

Bimpikis K, Crapis D, Tahbaz-Salehi A (2019) Information sale and competition. Manag Sci 65:2646–2664. https://doi.org/10.1287/mnsc.2018.3068

Bolier L, Haverman M, Westerhof GJ et al (2013) Positive psychology interventions: a meta-analysis of randomized controlled studies. BMC Public Health 13:1–20. https://doi.org/10.1186/1471-2458-13-119

Brailovskaia J, Ströse F, Schillack H, Margraf J (2020) Less Facebook use – more well-being and a healthier lifestyle? An experimental intervention study. Comput Hum Behav 108:106332. https://doi.org/10.1016/j.chb.2020.106332

Brandtzaeg PB, Følstad A, Chaparro Domínguez MÁ (2018) How journalists and social media users perceive online fact-checking and verification services. J Pract 12:1109–1129. https://doi.org/10.1080/17512786.2017.1363657

Brennen JS, Nielsen RK (2020) COVID–19 has intensified concerns about misinformation. Here's what our past research says about these issues. In: Reuters Inst. https://reutersinstitute.politics.ox.ac.uk/risj-review/covid-19-has-intensified-concerns-about-misinformation-heres-what-our-past-research. Accessed 19 Nov 2021

Bright LF, Kleiser SB, Grau SL (2015) Too much Facebook? An exploratory examination of social media fatigue. Comput Hum Behav 44:148–155. https://doi.org/10.1016/j.chb.2014.11.048

Bringula RP, Catacutan AE, Garcia MB et al (2021) "Who is gullible to political disinformation?" Predicting susceptibility of university students to fake news. J Inf Technol Polit. https://doi.org/10.1080/19331681.2021.1945988

Bryanov K, Vziatysheva V (2021) Determinants of individuals' belief in fake news: a scoping review determinants of belief in fake news. PLoS ONE 16:e0253717. https://doi.org/10.1371/journal.pone.0253717

Burrell J (2016) How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data Soc 3:1–12. https://doi.org/10.1177/2053951715622512

Burton-Jones A, Volkoff O (2017) How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. Inf Syst Res 28:468–489

Califf CB, Sarker S, Sarker S (2020) The bright and dark sides of technostress: a mixed-methods study involving healthcare IT. MIS Q 44:809–856. https://doi.org/10.25300/MISQ/2020/14818

Callahan D (1973) The WHO definition of 'health.' Hastings Cent Stud 1:77–87

Calvillo DP, Garcia RJB, Bertrand K, Mayers TA (2021) Personality factors and self-reported political news consumption predict susceptibility to political fake news. Personal Individ Differ 174:110666. https://doi.org/10.1016/j.paid.2021.110666

Calvo RA, Peters D (2014) Positive computing: technology for well-being and human potential. MIT Press, Cambridge

Carlson A (2017) The need for transparency in the age of predictive sentencing algorithms. Iowa Law Rev 103:303–329

Chen A, Karahanna E (2018) Life interrupted: the effects of technology-mediated work interruptions on work and nonwork outcomes. MIS Q 42:1023–1042. https://doi.org/10.25300/MISQ/2018/13631

Chen L, Baird A, Straub DW (2019) An analysis of the evolving intellectual structure of health information systems research in the information systems discipline. J Assoc Inf Syst 20:1023–1074

Cheng J-Z, Ni D, Chou Y-H et al (2016) Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci Rep 6:24454. https://doi.org/10.1038/srep24454

Choi JK, Ji YG (2015) Investigating the importance of trust on adopting an autonomous vehicle. Int J Hum-Comput Interact 31:692–702. https://doi.org/10.1080/10447318.2015.1070549

Clarke R (1994) The digital persona and its application to data surveillance. Inf Soc 10:77–92

Clarke J, Chen H, Du D, Hu YJ (2020) Fake news, investor attention, and market reaction. Inf Syst Res 32(1):35–52

Conway M (2017) Determining the role of the internet in violent extremism and terrorism: six suggestions for progressing research. Stud Confl Terror 40:77–98. https://doi.org/10.1080/1057610X.2016.1157408

Conway M, Macdonald S (2019) Introduction to the special issue: Islamic state's online activity and responses, 2014–2017. Stud Confl Terror 42:1–4

Cowgill B (2018) The impact of algorithms on judicial discretion: evidence from regression discontinuities. Working Paper

Crockett MJ (2017) Moral outrage in the digital age. Nat Hum Behav 1:769–771

Cunningham S, Hudson CC, Harkness K (2021) Social media and depression symptoms: a meta-analysis. Res Child Adolesc Psychopathol 49:241–253

Curchod C, Patriotta G, Cohen L, Neysen N (2020) Working for an algorithm: power asymmetries and agency in online work settings. Adm Sci Q 65:644–676

Davenport T, Beck J (2001) The attention economy: understanding the new currency of business. Harvard Business Review Press, Boston

Demirci K, Akgönül M, Akpinar A (2015) Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. J Behav Addict 4:85–92. https://doi.org/10.1556/2006.4.2015.010

Deng X, Joshi KD, Galliers RD (2016) The duality of empowerment and marginalization in microtask crowdsourcing: giving voice to the less powerful through value sensitive design. MIS Q 40:279–302

Diener E (1984) Subjective well-being. Psychol Bull 95:542–575. https://doi.org/10.1037/0033-2909.95.3.542

Diener E, Suh EM, Lucas RE, Smith HL (1999) Subjective well-being: three decades of progress. Psychol Bull 125:276–302. https://doi.org/10.1037/0033-2909.125.2.276

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: people erroneously avoid algorithms after seeing them err. J Exp Psychol Gen 144:114–126. https://doi.org/10.1037/xge0000033

Dinev T, McConnell AR, Smith HJ (2015) Research commentary – informing privacy research through information systems, psychology, and behavioral economics: thinking outside the "APCO" box. Inf Syst Res 26:639–655

Ellison NB, Steinfield C, Lampe C (2007) The benefits of Facebook "friends:" social capital and college students' use of online social network sites. J Comput-Mediat Commun 12:1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Ensign D, Friedler S, Neville S, et al (2017) Runaway feedback loops in predictive policing. ArXiv Prepr. ArXiv170609847

Eubanks V (2018) Automating inequality: how high-tech tools profile, police, and punish the poor. St. Martin's, New York

Facebook (2021) Facebook ad center, detailed targeting options. In: Facebook. https://www.facebook.com/ad_center/create/pagead/?entry_point=fb4b_create_ad_cta&page_id=864331173712397. Accessed 21 Aug 2021

Faelens L, Hoorelbeke K, Soenens B et al (2021) Social media use and well-being: a prospective experience-sampling study. Comput Hum Behav 114:106510

Felmingham CM, Adler NR, Ge Z et al (2021) The importance of incorporating human factors in the design and implementation of artificial intelligence for skin cancer diagnosis in the real world. Am J Clin Dermatol 22:233–242. https://doi.org/10.1007/s40257-020-00574-4

Felzmann H, Villaronga E, Lutz C, Tamò-Larrieux A (2019) Transparency you can trust: transparency requirements for artificial intelligence between legal norms and contextual concerns. Big Data Soc 6:1–14. https://doi.org/10.1177/2053951719860542

Fisher A, Prucha N, Winterbotham E (2019) Mapping the Jihadist information ecosystem: towards the next generation of disruption capability. Royal United Services Institute for Defence and Security Studies, London

Frankena WK (1973) Ethics, 2nd edn. Prentice Hall, Englewood Cliffs

Frauenberger C, Good J, Fitzpatrick G, Iversen OS (2015) In pursuit of rigour and accountability in participatory design. Int J Hum-Comput Stud 74:93–106. https://doi.org/10.1016/j.ijhcs.2014.09.004

Freude H, Heger O, Niehaves B (2019) Unveiling emotions: attitudes towards affective technology. In: Proceedings of the 40th International conference on information systems. Munich, pp 1–18

Friedman B, Kahn P (2003) Human values, ethics, and design. In: Jacko J, Sears A (eds) The Human-computer interaction handbook. Lawrence Erlbaum, Mahwah

Fruhwirth M, Rachinger M, Prlja E (2020) Discovering business models of data marketplaces. In: Proceedings of the 53rd Hawaii international conference on system sciences. Hawaii, pp 5736–5747

Gimpel H, Schröder J (eds) (2021) Hospital 4.0: Schlanke, digital-unterstützte Logistikprozesse in Krankenhäusern. Springer, Wiesbaden

Gimpel H, Lanzl J, Regal C et al (2019) Gesund digital arbeiten?! Eine Studie zu digitalem Stress in Deutschland. Projektgruppe Wirtschaftsinformatik des Fraunhofer FIT, Augsburg

Gimpel H, Manner-Romberg T, Schmied F, Winkler TJ (2021) Understanding the evaluation of mHealth app features based on a cross-country Kano analysis. Electron Mark Online Ahead Print. https://doi.org/10.1007/s12525-020-00455-y

Gimpel H, Schmied F (2019) Risks and side effects of digitalization: a multi-level taxonomy of the adverse effects of using digital technologies and media. In: Proceedings of the 27th European conference on information systems. Stockholm, pp 1–15

Gkatzelis V, Aperjis C, Huberman BA (2015) Pricing Private Data Electron Mark 25:109–123. https://doi.org/10.1007/s12525-015-0188-8

Glikson E, Woolley AW (2020) Human trust in artificial intelligence: review of empirical research. Acad Manag Ann 14:627–660. https://doi.org/10.5465/annals.2018.0057

große Deters F, Mehl MR, (2013) Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. Soc Psychol Personal Sci 4:579–586. https://doi.org/10.1177/1948550612469233

HLEG of the EU Commission (2020) Assessment list for trustworthy AI (ALTAI). Brussels

Hosseini M, Shahri A, Phalp K, Ali R (2018) Four reference models for transparency requirements in information systems. Requir Eng 23:251–275

Hou Y, Xiong D, Jiang T et al (2019) Social media addiction: its impact, mediation, and intervention. Cyberpsychol J Psychosoc Res Cyberspace 13(1):4. https://doi.org/10.5817/CP2019-1-4

Huang C (2017) Time spent on social network sites and psychological well-being: a meta-analysis. Cyberpsychol Behav Soc Netw 20:346–354

Huang Q, Li Y, Huang S et al (2020) Smartphone use and sleep quality in Chinese college students: a preliminary study. Front Psychiatry 11:352. https://doi.org/10.3389/fpsyt.2020.00352

IEEE. (2021). IEEE 7000 - Model process for addressing ethical concerns during system design. In. Piscataway: IEEE computer society. https://engagestandards.ieee.org/ieee-7000-2021-for-systems-design-ethical-concerns.html. Accessed 19 Nov 2021

Jakobi T, von Grafenstein M, Legner C et al (2020) The role of IS in the conflicting interests regarding GDPR. Bus Inf Syst Eng 62:261–272. https://doi.org/10.1007/s12599-020-00633-4

Jobin A, Ienca M, Vayena E (2019) The global landscape for AI ethics guidelines. Nat Mach Intell 1:389–399

Jussupow E, Benbasat I, Heinzl A (2020) Why are we averse towards Algorithms? A comprehensive literature review on algorithm aversion. In: Proceedings of the 28th European Conference on Information Systems. A virtual AIS conference, pp 1–16

Kahneman D, Diener E, Schwarz N (eds) (1999) Well-being: foundations of hedonic psychology. Russell Sage

Kanno-Youngs Z, Sanger DE (2021) Extremists emboldened by Capitol attack pose rising threat, Homeland Security says. N. Y. Times. https://www.nytimes.com/2021/01/27/us/politics/homeland-security-threat.html. Accessed 19 Nov 2021

Karwatzki S, Trenz M, Tuunainen VK, Veit D (2017) Adverse consequences of access to individuals' information: an analysis of perceptions and the scope of organisational influence. Eur J Inf Syst 26:688–715. https://doi.org/10.1057/s41303-017-0064-z

Kastl J, Pagnozzi M, Piccolo S (2018) Selling information to competitive firms. RAND J Econ 49:254–282. https://doi.org/10.1111/1756-2171.12226

Kellogg KC, Valentine MA, Christin A (2020) Algorithms at work: the new contested terrain of control. Acad Manag Ann 14:366–410

Keyes CLM (1998) Social well-being. Soc Psychol Q 61:121–140. https://doi.org/10.2307/2787065

Kilger M (1994) The digital individual. Inf Soc 10:93–99

Kim A, Dennis AR (2019) Says who? The effects of presentation format and source rating on fake news in social media. MIS Q 43:1025–1039

Kim JW, Ryu B, Cho S et al (2019) Impact of personal health records and wearables on health outcomes and patient response: three-arm randomized controlled trial. JMIR MHealth UHealth 7:e12070

Kim TW, Routledge BR (2018) Informational privacy, a right to explanation, and interpretable AI. In: 2018 IEEE symposium on privacy-aware computing. pp 64–74

Kirchhof G, Lindner JF, Achenbach S et al (2018) Stratified prevention: opportunities and limitations. Clin Res Cardiol 107:193–200

Kitchens B, Johnson SL, Gray P (2020) Understanding echo chambers and filter bubbles: the impact of social media on diversification and partisan shifts in news consumption. MIS Q 44:1–32

Kizilcec RF (2016) How much information? Effects of transparency on trust in an algorithmic interface. In: Proceedings of the 2016 CHI conference on human factors in computing systems. pp 2390–2395

Klein AZ, Magge A, O'Connor K et al (2021) Toward using Twitter for tracking COVID-19: a natural language processing pipeline and exploratory data set. J Med Internet Res 23:e25314

Kleinberg J, Lakkaraju H, Leskovec J et al (2017) Human decisions and machine predictions. Q J Econ 133:237–293

Koroleva K, Krasnova H, Veltri NF, Günther O (2011) It's all about networking! Empirical investigation of social capital formation on social network sites. In: International conference on information systems. Shanghai, pp 1–20

Krämer J, Schnurr D, Wohlfarth M (2019) Winners, losers, and Facebook: the role of social logins in the online advertising ecosystem. Manag Sci 65:1678–1699. https://doi.org/10.1287/mnsc.2017.3012

Krasnova H, Abramova O, Baumann A, Notter I (2016) Why phubbing is toxic for your relationship: understanding the role of smartphone jealousy among "Generation Y" users. In: European conference on information systems. İstanbul, pp 1–20

Kross E, Verduyn P, Demiralp E et al (2013) Facebook use predicts declines in subjective well-being in young adults. PLoS ONE 8:e69841. https://doi.org/10.1371/journal.pone.0069841

Kross E, Verduyn P, Sheppes G et al (2021) Social media and well-being: pitfalls, progress, and next steps. Trends Cogn Sci 25:55–66

Laato S, Islam AN, Islam MN, Whelan E (2020) What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? Eur J Inf Syst 29:288–305

Lambrecht A, Tucker C (2019) Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of stem career ads. Manag Sci 65:2966–2981

Lazer D (2015) The rise of the social algorithm. Science 348:1090–1091. https://doi.org/10.1126/science.aab1422

Lazer DM, Baum MA, Benkler Y et al (2018) The science of fake news. Science 359:1094–1096

Lee MK (2018) Understanding perception of algorithmic decisions: fairness, trust, and emotion in response to algorithmic

management. Big Data Soc 5:1–16. https://doi.org/10.1177/2053951718756684

Lepp A, Barkley JE, Karpinski AC (2014) The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. Comput Hum Behav 31:343–350. https://doi.org/10.1016/j.chb.2013.10.049

Levitin A (2003) Introduction to the design & analysis of algorithms. Addison-Wesley

Liberini F, Russo A, Cuevas Á, Cuevas R (2020) Politics in the Facebook era - evidence from the 2016 US presidential elections. Center for Economic Studies and ifo Institute, Munich

Liu D, Baumeister RF, Yang C (2019) Digital communication media use and psychological well-being: a meta-analysis. J Comput-Mediat Commun 24:259–274

Liv N, Greenbaum D (2020) Deep fakes and memory malleability: false memories in the service of fake news. AJOB Neurosci 11:96–104. https://doi.org/10.1080/21507740.2020.1740351

Logg JM, Minson JA, Moore DA (2019) Algorithm appreciation: people prefer algorithmic to human judgment. Organ Behav Hum Decis Process 151:90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Majchrzak A, Markus ML (2012) Technology affordances and constraints in management information systems (MIS). In: Kessler E (ed). Encyclopedia of management theory. Sage, Forthcoming, USA

Mann G, O'Neil C (2016) Hiring algorithms are not neutral, https://hbr.org/2016/12/hiring-algorithms-are-not-neutral. Accessed 21 Aug 2021

Martel C, Pennycook G, Rand DG (2020) Reliance on emotion promotes belief in fake news. Cogn Res Princ Implic 5:47. https://doi.org/10.1186/s41235-020-00252-3

Martin K (2019) Ethical implications and accountability of algorithms. J Bus Ethics 160:835–850. https://doi.org/10.1007/s10551-018-3921-3

Matook S, Cummings J, Bala H (2015) Are you feeling lonely? The impact of relationship characteristics and online social network features on loneliness. J Manag Inf Syst 31:278–310

McAfee A, Brynjolfsson E, Davenport TH et al (2012) Big data: the management revolution. Harv Bus Rev 90:60–68

McKnight DH, Cummings LL, Chervany NL (1998) Initial trust formation in new organizational relationships. Acad Manage Rev 23:473–490

McKnight DH, Carter M, Thatcher J, Clay P (2011) Trust in a specific technology. ACM Trans Manag Inf Syst TMIS 2:1–25. https://doi.org/10.1145/1985347.1985353

Mei X, Lee H, Diao K (2020) Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat Med 26:1224–1228

Melendez S, Pasternack A (2019) Here are the data brokers quietly buying and selling your personal information. In: Fast Co. https://www.fastcompany.com/90310803/here-are-the-data-brokers-quietly-buying-and-selling-your-personal-information. Accessed 20 Jan 2020

Mittelstadt B (2019) Principles alone cannot guarantee ethical AI. Nat Mach Intell 1:501–507

Möhlmann M, Zalmanzon L, Henfridsson O, Gregory RW (2021) Algorithmic management of work on online labor platforms: when matching meets control. MIS Q, Forthcoming

Monar J (2007) Common threat and common response? The EU's counter-terrorism strategy and its problems. Gov Oppos 42:292–313

Montgomery KC (2015) Youth and surveillance in the Facebook era: policy interventions and social implications. Telecommun Policy 39:771–786. https://doi.org/10.1016/j.telpol.2014.12.006

Newman J (2020) This AI fact-checking startup is doing what Facebook and Twitter won't. In: Fast Co. https://www.fastcompany.com/90535520/this-ai-fact-checking-startup-is-doing-what-facebook-and-twitter-wont. Accessed 27 Aug 2021

Nordenfelt L (1993) Quality of life, health and happiness. Averbury, Aldershot

Nouri L, Lorenzo-Dus N, Watkin A-L (2019) Following the whack-a-mole: Britain First's visual strategy from Facebook to Gab. Royal United Services Institute for Defence and Security Studies, London

Nuraniyah N (2019) The evolution of online violent extremism in Indonesia and the Philippines. Royal United Services Institite for Defence and Security Studies, London

O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy. Crown, New York

Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Sci 366:447–453

Parra-Arnau J (2018) Optimized, direct sale of privacy in personal data marketplaces. Inf Sci 424:354–384. https://doi.org/10.1016/j.ins.2017.10.009

Parry DA, Davidson BI, Sewall CJ et al (2021) A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. Nat Hum Behav. https://doi.org/10.1038/s41562-021-01117-5

Pasquale F (2015) The black box society: the secret algorithms that control money and information. Harvard University Press, London

Pennycook G, Bear A, Collins ET, Rand DG (2020) The implied truth effect: attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. Manag Sci 66:4944–4957. https://doi.org/10.1287/mnsc.2019.3478

Pirkkalainen H, Salo M (2016) Two decades of the dark side in the information systems basket: suggesting five areas for future research. In: European conference on information systems. Istanbul, pp 1–16

Polonski V (2018) AI is convicting criminals and determining jail time, but is it fair? https://www.weforum.org/agenda/2018/11/algorithms-court-criminals-jail-time-fair. Accessed 21 Aug 2021

Prahl A, Van Swol L (2017) Understanding algorithm aversion: when is advice from automation discounted? J Forecast 36:691–702. https://doi.org/10.1002/for.2464

Qureshi I, Bhatt B, Gupta S, Tiwari AA (2020) Call for papers: Causes, symptoms and consequences of social media induced polarization (SMIP). Inf Syst J 1–11

Rahman HA, Valentine MA (2021) How managers maintain control through collaborative repair: evidence from platform-mediated "Gigs." Organ Sci 32:1149–1390. https://doi.org/10.1287/orsc.2021.1428

Rahwan I, Cebrian M, Obradovich N et al (2019) Machine behaviour. Nature 568:477–486

Rai A (2020) Explainable AI: from black box to glass box. J Acad Mark Sci 48:137–141

Rimol M (2021) Gartner forecasts global spending on wearable devices to total $81.5 billion in 2021. Gartner, Stamford

Rissler R, Nadj M, Li MX et al (2020) To be or not to be in flow at work: physiological classification of flow using machine learning. IEEE Trans Affect Comput. https://doi.org/10.1109/TAFFC.2020.3045269

Rolnick D, Donti PL, Kaack LH, et al (2019) Tackling climate change with machine learning. ArXiv190605433 Cs Stat

Roozenbeek J, Schneider CR, Dryhurst S et al (2020) Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 7:201199. https://doi.org/10.1098/rsos.201199

Rosenblatt A (2018) Uberland: how algorithms are rewriting the rules of work. University of California Press, Oakland

Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1:206–215

Ryan RM, Deci EL (2001) On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. Annu Rev Psychol 52:141–166. https://doi.org/10.1146/annurev.psych.52.1.141

Ryan RM, Huta V, Deci EL (2008) Living well: a self-determination theory perspective on eudaimonia. J Happiness Stud 9:139–170

Ryff CD, Keyes CLM (1995) The structure of psychological well-being revisited. J Pers Soc Psychol 69:719–727. https://doi.org/10.1037/0022-3514.69.4.719

Sarker S, Chatterjee S, Xiao X, Elbanna A (2019) The sociotechnical axis of cohesion for the is discipline: its historical legacy and its continued relevance. MIS Q 43:695–720

Saunders C, Benlian A, Henfridsson O, Wiener M (2020) IS control and governance. MIS Q Res Curations 1–14

Schaefer KE, Chen JY, Szalma JL, Hancock PA (2016) A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. Hum Factors 58:377–400

Schechner S, Secada M (2019) You give apps sensitive personal information. Then they tell Facebook. In: Wall Str. J. https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636?mod=e2tw. Accessed 19 Nov 2021

Schneier B (2015) Data and Goliath: the hidden battles to collect your data and control your world, reprint. Norton, New York

Schomakers E-M, Lidynia C, Ziefle M (2020) All of me? Users' preferences for privacy-preserving data markets and the importance of anonymity. Electron Mark 30:649–665. https://doi.org/10.1007/s12525-020-00404-9

Schor JB, Attwood-Charles W, Cansoy M et al (2020) Dependence and precarity in the platform economy. Theory Soc 49:833–861

Sharma K, Qian F, Jiang H et al (2019) Combating fake news: a survey on identification and mitigation techniques. ACM Trans Intell Syst Technol 10:1–41

Shu K, Bhattacharjee A, Alatawi F et al (2020) Combating disinformation in a social media age. Wiley Interdiscip Rev Data Min Knowl Discov 10:1–39

Sindermann C, Cooper A, Montag C (2020) A short review on susceptibility to falling for fake political news. Curr Opin Psychol 36:44–48. https://doi.org/10.1016/j.copsyc.2020.03.014

Spiekermann S (2021) Value-based Engineering: Prinzipien und Motivation für bessere IT Systeme. Inform Spektrum 44:247–256

Spiekermann S, Korunovska J (2017) Towards a value theory for personal data. J Inf Technol 32:62–84. https://doi.org/10.1057/jit.2016.4

Spiekermann S, Acquisti A, Böhme R, Hui K-L (2015a) The challenges of personal data markets and privacy. Electron Mark 25:161–167. https://doi.org/10.1007/s12525-015-0191-0

Spiekermann S, Böhme R, Acquisti A, Hui K-L (2015b) Personal data markets. Electron Mark 25:91–93. https://doi.org/10.1007/s12525-015-0190-1

Spiekermann-Hoff S, Krasnova H, Hinz O (2021) 05/2023 – Technology for humanity. In: Bus Inf Syst Eng. https://www.bise-journal.com/?p=1940

Srivastava SC, Chandra S, Shirish A (2015) Technostress creators and job outcomes: theorising the moderating influence of personality traits. Inf Syst J 25:355–401

Statista (2021a) Number of smartphone users worldwide from 2016 to 2023. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. Accessed 19 Nov 2021

Statista (2021b) Daily time spent on social networking by internet users worldwide from 2012 to 2020. https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/. Accessed 19 Nov 2021

Sweeney L (2013) Discrimination in online ad delivery. Queue 11:10–29

Tiggemann M, Zaccardo M (2015) Exercise to be fit, not skinny": the effect of fitspiration imagery on women's body image. Body Image 15:61–67. https://doi.org/10.1016/j.bodyim.2015.06.003

Tolmeijer S, Kneer M, Sarasua C et al (2020) Implementations in machine ethics: a survey. ACM Comput Surv 53:6. https://doi.org/10.1145/3419633

Trang S, Trenz M, Weiger WH et al (2020) One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. Eur J Inf Syst 29:415–428

Turel O, Matt C, Trenz M, Cheung CMK (2020) An intertwined perspective on technology and digitised individuals: linkages, needs and outcomes. Inf Syst J 30:929–939

Vaghefi I, Lapointe L, Boudreau-Pinsonneault C (2017) A typology of user liability to IT addiction. Inf Syst J 27:125–169

Valkenburg PM, Beyens I, van Driel II et al (2021a) Social media use and adolescents' self-esteem: heading for a person-specific media effects paradigm. J Commun 71:56–78. https://doi.org/10.1093/joc/jqaa039

Valkenburg PM, van Driel II, Beyens I (2021) The associations of active and passive social media use with well-being: a critical scoping review. PsyArXiv Prepr. https://doi.org/10.31234/osf.io/j6xqz

Vallas S, Schor JB (2020) What do platforms do? Understanding the gig economy. Annu Rev Sociol 46:273–294

Vallor S (2016) Technology and the virtues – a philosophical guide to a future worth wanting. Oxford University Press, New York

van den Broek T, van Veenstra AF (2018) Governance of big data collaborations: how to balance regulatory compliance and disruptive innovation. Technol Forecast Soc Change 129:330–338. https://doi.org/10.1016/j.techfore.2017.09.040

van der Aalst W, Hinz O, Weinhardt C (2019) Big digital platforms. Bus Inf Syst Eng 61:645–648

van Doorn N (2017) Platform labor: on the gendered and racialized exploitation of low-income service work in the 'on-demand' economy. Inf Commun Soc 20:898–914

Vanden Abeele MMP (2020) Digital wellbeing as a dynamic construct. Commun Theory 31(4):932–955. https://doi.org/10.1093/ct/qtaa024

Vodanovich S, Sundaram D, Myers M (2010) Digital natives and ubiquitous information systems. Inf Syst Res 21:711–723

Volz D, Levy R (2021) Social media plays key role for domestic extremism, FBI director says. In: Wall Str. J. https://www.wsj.com/articles/social-media-is-key-amplifier-of-domestic-violent-extremism-wray-says-11618434413. Accessed 15 Oct 2021

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. Science 359:1146–1151. https://doi.org/10.1126/science.aap9559

Wessels N, Gerlach J, Wagner A (2019) To sell or not to sell – antecedents of individuals' willingness-to-sell personal information on data-selling platforms. In: Proceedings of the 40th international conference on information systems. Munich, pp 1–17

Westerlund M (2019) The emergence of deepfake technology: a review. Technol Innov Manag Rev 9:39–52. https://doi.org/10.22215/timreview/1282

WHO (1948) Constitution of the World Health Organization. World Health Organization, Geneva

Wiener M, Cram WA, Benlian A (2020) Technology-mediated control legitimacy in the gig economy: conceptualization and nomological network. In: Hirschheim R et al (eds) Information systems outsourcing. Progress in IS Springer, Cham

Wiener M, Cram A, Benlian A (2022) Algorithmic control and gig workers: a legitimacy perspective of Uber drivers. Eur J Inf Syst Forthcom. https://doi.org/10.1080/0960085X.2021.1977729

Winter C, Neumann P, Meleagrou-Hitchens A et al (2020) Online extremism: research trends in internet activism, radicalization, and counter-strategies. Int J Confl Violence IJCV 14:1–20

Woodford A (2018) Expanding fact-checking to photos and videos – about Facebook. In: Facebook. Accessed 27 Aug 2021

Zellers R, Holtzman A, Rashkin H, et al (2019) Defending against neural fake news. 1–21. arXiv:190512616

Zhang X, Zhang R, Yue WT, Yu Y (2019) What is your data strategy? The strategic interactions in data-driven advertising. In: Proceedings of the 40th International conference on information systems. Munich, pp 1–9