Opinion Paper

What is it about humanity that we can't give away to intelligent machines? A European perspective

Crispin Coombs [a,*], Patrick Stacey [a], Peter Kawalek [a], Boyka Simeonova [a], Joerg Becker [b], Katrin Bergener [b], João Álvaro Carvalho [c], Marcelo Fantinato [d], Niels F. Garmann-Johnsen [e], Christian Grimme [b], Armin Stein [b], Heike Trautmann [b]

[a] Centre for Information Management, School of Business and Economics, Loughborough University, UK
[b] European Research Center for Information Systems, University of Münster, Germany
[c] Department of Information Systems, University of Minho, Portugal
[d] Center for Artificial Intelligence, University of São Paulo, Brazil
[e] Dept. for Information Systems, University of Agder, Norway

ARTICLE INFO

ABSTRACT

One of the most significant recent technological developments concerns the development and implementation of 'intelligent machines' that draw on recent advances in artificial intelligence (AI) and robotics. However, there are growing tensions between human freedoms and machine controls. This article reports the findings of a workshop that investigated the application of the principles of human freedom throughout intelligent machine development and use. Forty IS researchers from ten different countries discussed four contemporary AI and humanity issues and the most relevant IS domain challenges. This article summarizes their experiences and opinions regarding four AI and humanity themes: Crime & conflict, Jobs, Attention, and Wellbeing. The outcomes of the workshop discussions identify three attributes of humanity that need preservation: a critique of the design and application of AI, and the intelligent machines it can create; human involvement in the loop of intelligent machine decision-making processes; and the ability to interpret and explain intelligent machine decision-making processes. The article provides an agenda for future AI and humanity research.

1. Introduction

One of the most significant recent technological developments concerns developing and implementing intelligent, interactive, and highly networked machines within organizations and society. These "intelligent machines" are characterized by autonomy, the ability to learn, and the ability to interact with other systems and humans. They draw on new advances in technologies such as artificial intelligence (AI) and robotics, enabling them to undertake tasks that human workers could previously only complete (Coombs, Hislop, Taneva, & Barnard, 2020). Referring to what some have called the second machine age (Brynjolfsson & McAfee, 2016), analysts and commentators have highlighted the growing tensions between human freedoms and machine controls. For example, the more advanced intelligent machines become (e.g., more human-like

androids), the more blurred the physical, psychological, and social boundaries between machines and humans. Should machines be "looking" after clinical patients, educating students, and making complex financial or security decisions? While the experienced and anticipated benefits of these technologies for individuals, organizations and societies are apparent (e.g., Calo, Hunt-Bull, Lewis, & Metzler, 2011; Luxton, 2014) rapid technological developments in this area may also posit some severe risks. For example, using simulations for patients with delusional or psychotic psychopathologies in the absence of careful monitoring may put these patients' health at significant risk (Luxton, 2014). Torras (2015) warns about potential negative impacts of robot nannies on children's psychological development. For instance, how could a robot achieve a balance between protecting a child from danger and restricting his/her freedom (hence, affecting the child's

development to become mature and autonomous)? Such advancing interactions between machines and humans are psychologically complex and evoke critical ethical questions (Coombs et al., 2020).

Against this background, a workshop on the application of the principles of human freedom throughout intelligent machine development and use was conducted at the annual European Research Center for Information Systems (ERCIS) meeting held at Loughborough University, United Kingdom in September 2019. Forty IS researchers from ten different countries discussed AI and humanity issues and the challenges most relevant for the IS domain. The IS researchers represented a wide range of different IS perspectives. This article summarizes their experiences and opinions and combines them with the academic literature on AI and humanity. The workshop participants contributed to the debate by detailing their thoughts and ideas regarding four AI and humanity themes: *Crime & conflict*, *Jobs*, *Attention*, and *Wellbeing*. These four themes were selected because they represent common debates in the literature and more widely in public discourse. Some examples of how the themes apply include: In *Crime and conflict*, how can the AI arms race be addressed through humanistic practices and logics? In *Jobs*, how much displacement can we tolerate? In *Attention*, how much time can we spend online, how much control do we exhibit when online, and who or what mediates where attention is given? In *Wellbeing*, can responsibility for our mental and physical wellbeing be shared with AI? These themes serve as a framework for a debate of what might be lost and what might be retained, and the identification of common themes within.

This article contributes to the ongoing discussion in the International Journal of Information Management regarding issues and challenges raised by AI. Recent studies have considered a range of different topics including the strategic use of AI (Borges, Laurindo, Spínola, Gonçalves, & Mattos, 2020), AI and future of work (Coombs, 2020), consumers acceptance of AI devices (Gursoy, Chi, Lu, & Nunkoo, 2019), the impact of AI on decision making (Duan, Edwards, & Dwivedi, 2019), AI's influence on human cognition (Hu, Lu, Pan, Gong, & Yang, 2021) and using AI tools to predict human behaviour (Abubakar, Behravesh, Rezapouraghdam, & Yildiz, 2019). Research has also considered AI issues in a range of contexts including disaster management (Fan, Zhang, Yahja, & Mostafavi, 2021), sustainability (Nishant, Kennedy, & Corbett, 2020), digital and social media marketing (Dwivedi et al., 2020), and for responding to the COVID-19 pandemic (Sipior, 2020). Many of these studies touch on how the boundary between AI controls and human freedoms should be managed. This article extends this debate by identifying three attributes of humanity that need preservation: a critique of the design and application of AI, and the intelligent machines it can create; human involvement in the loop of intelligent machine decision-making processes; and the ability to interpret and explain intelligent machine decision-making processes.

The remainder of the article is structured as follows. Section two provides the theoretical framing of the debate regarding AI and humanity. Section three explains the workshop setting and provides a critical reflection of the workshop discussion of each of the four themes, including potential research problems and critical questions to be explored by IS researchers. Section four then synthesizes the key insights from the workshop to provide an agenda for future research. Section five concludes the article and acknowledges the limitations.

2. Theoretical background on human freedom and machines – Patrick Stacey

Computational technologies are being misused to threaten, take away, or even abuse our basic human rights (Schippers, 2018). For example, our human rights to life, privacy, freedom of expression, and work. Consider these selected four human rights in light of the following examples:

- The AI arms race; the fear that LAWS (lethal autonomous weapons systems) will put autonomous robotic systems in charge of life and death decisions

- Google's participation in Project Maven, a military program that uses machine learning to analyze drone surveillance footage.
- The Edward Snowden revelations that revealed gross invasions of privacy by the NSA.
- Cambridge Analytica's scraping of personal data, being used in the manipulation and interference in democratic elections, damaging the right to freedom of expression.
- Automated screening of CVs by machine learning algorithms.

There are attempts to protect human users from such technological abuses, such as:

- UN Framing and Guiding Principles on Business and Human Rights
- Magna Carta for the digital age (Tim Berners Lee)
- Future of Life Institute's 23 Asilomar AI Principles
- The Electronic Frontier Foundation (EFF)
- The EU-funded Humane AI project (www.humane-ai.eu)

But such initiatives lag accelerating computational change and the invention of new applications. This motivated MIT's Work of the Future task force to call for greater institutional agility to protect our basic freedoms. However, in our view, solely relying on these initiatives to ensure freedom is inadequate. What is required is a fundamental understanding and application of human freedom principles throughout systems use and development (Stacey & Tether, 2015).

To this end, we must naturally define human freedom. This is an age-old quest, but we must do it and keep doing it. Drawing on established early modern-to-modern philosophy, to be a free human is essentially being able to critique the world to make informed choices (c.f. Kant). According to Hegel (2018), critique must also deal with the history of phenomena globally and apply mediation, synthesis and negation to any contradictions arising. Humans earn their freedom through a mindful critique of phenomena in the world. But is such critique unique to humans? AI is already being used to write news stories and financial reports:

"Companies in this business aim to relieve humans from the burden of the writing process by using algorithms and natural language generators to create written content. Feed their platforms some data — financial earnings statistics, let's say — and poof! In seconds, out comes a narrative that tells whatever story needs to be told." (Marr, 2019).

Arguably, this is neither Kantian nor Hegelian critique; rather, it is a narrative being generated from datasets. However, for Heidegger, this is the beginning of a harmful development. Heidegger posited that technology presents the ultimate danger to humanity, mainly when humans are no longer the source of meaning-making (Heidegger, 1954[1]). This is especially so if AI develops the ability to write independent critical essays. Critique then could be a crucial concept that humanity hangs onto as a means of liberation.

For Nietzsche, the act of critique liberates us from any limiting conceptual constraints and conventions to enable us to be ultimately creative and ingenious (Kellner & Lewis, 2007). Critique is also about creativity and using it to overcome constraint (Molnar, Nandhakumar, & Stacey, 2017; Stacey & Nandhakumar, 2009). Nietzsche and Foucault advocated aesthetic creativity as a form of human resistance to a mass culture of conformity. We see various creativity studies within computer science, including design fiction (e.g., Coulton & Lindley, 2019). Yet, computers also putatively indulge in aesthetic creativity. For example, Microsoft produced an 'original Rembrandt' using machine learning. To do this, algorithms read the signs of Rembrandt's style and (re)created an original image, which was 3D-printed into material form (Baraniuk,

---

[1] ***The Question Concerning Technology*** (German: *Die Frage nach der Technik*) is a work by Martin Heidegger, in which the author discusses the essence of technology. Heidegger originally published the text in 1954, in *Vorträge und Aufsätze*.

2016). Mario Klingemann, a German artist who uses AI in his work, has radical views on creativity. "Humans are not original," he says; "We only reinvent, make connections between things we have seen." While humans can only build on what we have learned and what others have done before us, "machines can create from scratch" (Ibid). Computers are already impinging on a core means of Nietzschean human liberation. Humans often define themselves by how creative they are - a question of identity (Adarves-Yorno, Postmes, & Haslam, 2007). But this is somewhat paranoid - machines creating art does not abuse nor prevent human beings from making art. However, if machines were *curating, deciding* on art submissions from human artists, then this is a different issue.

Yet a similar phenomenon is already upon us on the battlefield, and the road - LAWS involves robots making autonomous life/death decisions, and cars can now decide how to drive themselves and their human passengers; we are already under the thumb of technology, abdicating specific critical responsibilities. We have moved from the computer as a decision support system (e.g., Sprague & Watson, 1993) to an autonomous decision system (McDermid, 2019). Again, Heidegger's thesis of the ultimate danger of technology surfaces - humans are no longer the sole source of meaningful, critical, responsible action. And yet, AI is described as having no understanding, despite being as smart as an eighth-grade student (Metz, 2019). This makes the autonomous decision system sound even more dangerous. To subvert such technological dangers and retain human freedoms, creative critique must be directed at several mediating contextual forces; for example, capitalistic computerization movements and advocacy (e.g., Elliott & Kraemar, 2008).

The pace of computational technological change is accelerating rapidly, unlike in any other era. This sounds facile perhaps, but it is recognized in a recent study by the UN entitled "The impact of the technological revolution on labour markets and income distribution." The UN study states the pace of breakthroughs in several clusters, including gene editing and machine learning, and that these signify that a new technological revolution is at hand, and every industry will be affected. The study even discusses how AI could replace medical doctors. Computational technologies threaten to do for cognitive ability what factory machines did for muscle power (Dwivedi et al., 2019). At this pace of change, there is little or no time for the user or business to weigh things up, to reflect. This is perhaps what some businesses hope for, of course – a temporal trap in which the switching costs are too high because nobody has the time for critique anymore. Complementing the pace of computational change is the capitalistic legitimacy of technology (Elliott & Kraemar, 2008; Markard, Wirth, & Truffer, 2016). This has created a socio-cultural context in which Tech Giants are assumed to be the 'good guys' despite the ethical quagmires that embroil Apple, Facebook and Google[2]. It is akin to Nietzsche on the invention of good and bad – the strong inventing the term 'good' to make themselves *feel* 'good' about their actions. This capitalistic legitimacy is reinforced through Advocacy. Advocates of computerization movements, such as cloud computing, spread their message through public discourse in various segments of society such as media, academics, visionaries, and professional societies. This discourse sediments technological frames (Orlikowski & Gash, 1994) which are composite understandings about how computer technology works and could be used. Advocacy and frames are sedimented into mass media, scientific journals, TED talks, and trade journals (Iacono & Kling, 2001). Also, User eXperience Design theory (UX), for example, nudge theory and gamification (Liu, Santhanam, & Webster, 2017), attempts to smooth the flow of technological capital from invention to user acceptance and consumption.

In conclusion, humans need to preserve the well-established ability of critique for themselves. Limits should be placed on machines *per se*

but also the mediating, hegemonic discourses, and advocacies. One could critique this critique (i.e., meta critique) and revert to a Bauhausian view of human-machine unity. Currently, this is the dominant ontology in IS via sociomateriality (Orlikowski & Scott, 2008). However, the human being has de-levelled for too long in IS; we need a re-turn to humanism (Stacey, 2019).

Drawing on the above ideas of human liberty through critique, we discuss four themes throughout the remainder of this article:

1 Crime & Conflict
2 Jobs
3 Attention
4 Wellbeing

3. The workshop setting

We conducted a workshop in September 2019 with 40 IS researchers from nine European countries: Finland, Germany, Italy, Lithuania, Norway, Poland, Portugal, Spain and the United Kingdom, and Brazil. The participating scholars have diverse IS-related backgrounds, ranging from technical to managerial. As a consequence, not only do their research interests vary, but also the methods they apply. At the outset of the workshop, an inspiration session was presented to introduce the concept of new humanism and machines and the four themes selected for further investigation in the workshop. An initial open plenary discussion revealed four cross-cutting issues relevant to each theme: *bias, ethics, responsibility,* and *control*. It was agreed that each theme discussion should use these cross-cutting issues to guide conversations to ensure a coherent article could be produced.

To facilitate the discussion between workshop participants and capture key discussion points, a World Café format was followed (World Café Community Foundation, 2015). Small groups participated in four themed table discussions. Each table had a host, designated at the start of the discussion session that facilitated the meeting, captured key points on post-its, wrote up a summary of the discussion and remained on the table throughout the afternoon. Post-it notes were used to capture the ideas generated on each table. The table hosts are co-authors of this article. Their contributions reflect the table theme that they hosted during the workshop, the summary of the discussion and integration with the academic literature. Four table rotations ensured all experts contributed to all the research themes. After the rotations were complete, the experts were allowed to review and change their inputs through a plenary discussion.

4. Crime and Conflict - Niels F. Garmann-Johnsen, Marcelo Fantinato

4.1. Introduction

AI (Artificial Intelligence) is already heavily involved in Law and the wider procedures of Criminal Justice. AI is utilized in legal casework, to pick out who to scrutinize (through profiling) and direct police resources to patrol different geographical areas. There is the potential for AI to differentiate between offenders prosecuted for the same crime: an algorithm might predict which offender has the greatest chance of repeating a crime, meaning that sentencing periods will become based on automated behaviour predictions rather than a human assessment. This resembles the sci-fi-novel "Minority Report" and the issue of technological determinism raised therein. It becomes important to consider the chosen frame of reference behind such applications. These examples seem to manifest a view that there are technical design solutions to crime and conflict. Has such frame-of-reference obviated all alternative, non-technical approaches to crime prevention, e.g., the potential of social reform, the role of human development, childhood, socialization and so forth (Gabriel, 2015)?

---

[2] Google recently exploited child users of Youtube, being fined USD170m (Associated Press, 2019)

4.2. Bias

In the discussion, it was established that there should be no bias in AI when replacing a human institution or process such as a jury. In other words, the merits of the replacement (human by machine) must be a technical advantage, and there is no technical advantage worth having if there is bias. The ability to identify and diminish bias is, therefore, itself a technical skill. If we (sentient, emotional humans) fear machines' outputs because these machines can be biased and unfair, then why do we have them? From this, it follows that meeting this goal comes down to system design, *par excellence*. Machines need to be calibrated to make predictions of human behaviour by making a comparison with actual behaviour. This implies closed and open systems development and evaluation that mitigates against and overcomes problems that are already reported such as machines that might learn prejudice from people, the unseen bias in data models, or logically sound but socially deleterious conclusions (Silberg & Manyika, 2019). Arguably then, we approach a new phase of Information Systems Development (ISD), still more demanding than the discipline hitherto known, in which the real-world effects are continuously monitored, analyzed and improved. This is a deepening of the concerns of ISD so that it remains responsible for the impact of software in relation to bias throughout its use.

4.3. Ethics

The final decision must be human. This is a crucial constitutional element. The last word in decision-making should never be left to machines. We can let ourselves be influenced by machines' recommendations, but humans must retain a power balance that ensures control. This is a simple principle that is likely to be complex in practice. People will have to work to understand decisions and consequences from mixed AI/human worlds, seeking to ensure social acceptability and that the broad implications of decisions are understood. In this respect, the partnership with machines might save costs and generate greater accuracy but then will, also, develop higher-order complexity for social resolution. The justification of machines then hinges on the nature of this higher-order complexity. For machines to be valuable overall, they must support higher-level, more socially beneficial thinking and discussion by humans. In other words, by virtue of the assistance of machines, we humans can build social processes relating to crime and justice that better facilitate human society. If alternately, the utilization of machines leads to complex problems and harmful conditions that humans must manage and resolve, then that is a debit against the whole worth of the machines.

4.4. Control

Following the core condition that the final decision must be human, machines in criminal investigations or court systems could serve only as referees and not as decision-makers. The implication of this is additional complexity and difficulty in the human debate around the machines: since datasets could be manipulated and biased, how do we avoid that except by building the court's institutional responsibilities and the human stakeholders' education? It follows that the perceived benefits of algorithmic referees might be offset by greater complexity at other levels and new forms of uncertainty.

A recent topic in AI is "explainable AI" (Adadi & Berrada, 2018). This is concerned with explaining how AI decisions were made. The argument is that AI must be designed to be traceable through the algorithm's entire path before producing the result. Such a requirement is highly problematic for many AI techniques, such as some types of neural networks (e.g., convolutional nets), especially where deep learning is used, as such systems work with many mappings and mathematical transformations (filters and codings). Such layered complexity makes it almost impossible to explain how the result was achieved, i.e., to track-back.

4.5. Responsibility

Extending this issue of control opens a question of responsibility for machines' contribution: are they legal subjects? Is accountability expressed solely in relation to the machine, or is there additional human accountability? Again, an implication that the "final decision must be human" is a requirement that there should always be a human author or owner that could be held legally responsible for the damage. Once again, this implies that algorithmic machines' apparent cost savings and accuracy advantages will create costs and issues elsewhere in the system.

4.6. Future research

Many scenarios can be developed, many of which have frightening implications, e.g., should machines be afforded control of weaponry for civil behaviour or war? How does this work when, as we know from history, human progress is sometimes dependent on civil disobedience? The ethical questions are many, but we need ways of establishing them in popular discourse.

It follows that an open debate is needed leading to public awareness. In our discussions, we saw potential and possibilities for better decision-making in AI, as well as these challenges; the tone was not pessimistic. We saw the need for safety-mechanisms designed into systems, by law and engineering but, once in place, a majority of scholars saw advantages in the application of machine learning and decision making, or, perhaps, *decision-aiding*.

Looking at the problem from a specific focus on Information Systems allows us to perceive machines' role in decision-processes. This might be an area of important insight: to enlarge the scope from the decision to how data is gathered, utilized and how consequences manifest among stakeholders. Extending this, it is comprehensible that we could ask machines to justify decisions or advice, aiding scrutiny. Following the same logic of looking at an enlarged decision-making process, it can be understood that essential information and knowledge-asymmetries are likely to develop, between machines and human minds. Such asymmetries will be so great in favour of machines that it might be impossible for humans to scrutinize decisions themselves. There will be a shift of Information Systems to becoming a discipline of mapping consequences and making adjustments to optimize machines to society's needs. A principle of restraint might be needed to ensure this.

Discussions turned to the book "Coping With The Future," and its chapter 10, wherein Sampanikou and Johnsen investigate a posthuman, machine-dominated future with transhumanism, the merger of machines and human biology or thinking, as a "grey-zone" (Johnsen, Holtskog, Ennals, & John, 2018). They challenge the assumption that post- and trans-humanism posed by new technologies are the *only* threats to humanism but concede worrying signs that emphasize the need for the ongoing debate on the challenges posed by such developments.

5. Jobs - João Álvaro Carvalho

5.1. Introduction

Jobs are so ingrained to human life that they become part of one's identity: What is your name? What do you do for a living?

What one "does for a living" is a crucial aspect of personal characterization. It is a pivot and a fulcrum to modern identity. This is not just in adulthood. It starts early in someone's life - what do you want to be when you are grown up? It accompanies the young person almost always through education: *what are the career prospects of this programme?*

Jobs are central to human life. They direct education choices; they provide financial rewards and act as a conditioning mechanism through social recognition and status. Furthermore, jobs confer a sense of self-fulfilment and realization. They are a means by which a person can feel that s/he plays a role in society, that this person contributes to the

creation of wealth and the social order.

Thus, jobs are germane to wellbeing and humanity's social organization. Ideally, they constitute a mechanism that, in a unified/integrated way, can provide for needs across the whole range of human experience (McLeod, 2007).

In our workshop discussion, we acknowledged that the future for jobs is not bright. There are plenty of texts that warn of difficulties ahead of employment and job-markets (e.g. Acemoglu & Restrepo, 2020). The future is likely to be a mixed prospect. For some, work opportunities might be plentiful, and the jobs they take might be close to the ideal. For many others, work might be challenging, and their position might fail to completely account for some of their human needs. This needs to be understood as a crisis working across Maslow's hierarchy. It is not just that higher-order needs might be sacrificed but also that we already know that it is not rare to hear reports of jobs that fail to provide basic needs such as proper nourishment and shelter.

5.2. Bias

Bias concerning jobs might mean many things, but a pertinent one is to consider algorithms' role in employee selection and promotion. Specific individuals with certain experiences and backgrounds might be increasingly rewarded by algorithms that cannot be persuaded by idiosyncratic factors (e.g., overcoming social circumstances and coping with ill health). As a black box, it can be argued that algorithmic selection and promotion is fair in that a consistent set of rules is applied. Still, in the more complex world of society, any formula might have the potential to be reductive and unfair.

5.3. Ethics

In our discussion, it was presented that the relationship between jobs and employment used to be more-or-less straightforward. Within our lifetimes, arrangements have been reasonably predictable. Competencies, their character, requirements, and their market value were all reasonably well understood. Relations between employers and employees were also predictable. It was often reasonable to expect a long-term relationship between employer and employee.

Machine transformations affect such state-of-affairs. A new technological society with its fast-evolving capabilities generates circumstances of adaptation and change among human competences. This potentially changes ethical relations between employees and employers, and between employees (or economic actors more generally) and the market abstraction. What is the trade-off for people? What do they gain from adapting their skills, and what do they lose? New roles for machines through AI, robotics and other developments encourage the commoditization of competences. Work experience seems to be increasingly expendable. Control is exerted by economic interests beyond the person, enabled by machines and tools of permanent surveillance. New employment schemas potentially disrupt the known social order, aggravating inequalities and provoking social fissures. Suppose the role of jobs in the social order becomes fragmented. In that case, new reward systems will be sought potentially including solutions such as guaranteed basic income, new forms of money, or taxing robots and other machines in the way that society once would tax humans.

5.4. Control

Jobs have an obvious *work* dimension. This is work in the sense of effort, application, energy in the execution of tasks, physical or intellectual. Saving this work effort is a clear driver for technology. Historically, this has been predominantly in the case of dull, routine work. The market's logic is not so confined; however, it applies machines to reduce costs, irrespective of whether work is dull or not. We have to question existing assumptions about the substitution of human labour by machines. Historically, humans welcomed technology to support or automate work. Overall, the belief remains that automation will eliminate excessive effort, costs, or the dull part of work - thus freeing humanity from inhumanity: less work effort, more time for leisure, and other activities that contribute to wellbeing. Might such beneficial relations by machines be able to do what we find interesting, as well as what we find dull?

The picture has always been complex. Problems accompany benefits. Working time freed using technology is filled with more work-duties. Reduced working hours have been a mirage for most people. Furthermore, humans should have learned that technology is not just an enhancer of human capabilities; technology also enables new capabilities that are not in humans' reach, either individually or in cooperation with each other. Technology, starting by being a facilitator of human work, soon becomes a transformer of human labour and then of the social order previously established.

History shows how the technological augmentation of human physical and movement capability led to a revolution of social arrangements and order through industry. The new revolution, currently underway, is related to the technological augmentation of human cognitive and communicative capabilities – information technology (IT), computation and networks. At the dawn of this new revolution, we can try to figure out what the prospect is for humankind.

An observable consequence of the widespread use of advanced algorithmic machines is an increase in the rhythm of life. Since the mass interconnection of computers through the Internet, perceptively space and time have shrunk. New forms of cooperation and coordination are possible or necessary. Humans themselves have been expected to fit the new technological milieu. We deal with business at a new pace, often at the cost of extra working hours and reducing leisure and family time. In this sense, there is a question of the control of time, and the usurpation of the human clock by network time. This network time is based upon the interconnection of machines and human actors spread across geographies and time zones. In this picture, work is saved, but there is no obvious priority to automating dull work or serving human values beyond the acquisition of material goods. The logic of the network is cost-efficiency so-as-to bring prosperity through markets. Hence, prosperities are traded: the ability to fit labour within human needs is traded for further efficiency and, behind that, potential material gains.

Work is automated according to what technology can do. This is a significant point of control. There is a lot that new machines can do, including making decisions. Today, decisions are made by the logic of the network. Soon, the machines themselves will additionally make decisions.

5.5. Responsibility

Central to human society, the nature and abundance of jobs have significant consequences across society's formation and running. The modern prospect of the automation of intellectual capabilities brings severe challenges to social arrangements. It is likely to be much more just the disappearance of some (a lot of) jobs and the emergence of some new ones. There are likely to be emergent effects across society and its power structures.

Data shows that already the distribution of wealth is increasingly affected, raising ethical issues in society. The contemporary benefits of the material consumerism are now characterized by increasing inequalities in their distribution (Brynjolfsson & McAfee, 2015). Whilst this inequality problem is already understood as a dilemma for society, an additional and accompanying dilemma might be a problem of responsibility. Who or what is responsible for decisions in human society? If machines are increasingly in charge of decisions, how does society vet them and whose interests do they serve? Who owns the machines? Problems of inequality and issues of responsibility will seriously weaken social cohesion, trust, and accountability.

The logic we know to date is that eventually, what can be automated will be automated. However, it is not clear what is going to happen to

responsibility. Are we reversing a democratic tradition of trying to share responsibility by automating decisions on some algorithmic basis, and thereby concentrating responsibility? Responsibility used to be assigned to humans. What will happen when machines make decisions? Who will understand what is in the algorithm?

5.6. Future research

Such a fundamental place is given to jobs in society that the changes now occurring merit profound, multi-disciplinary research. Concerning the Information Systems community, the study of machines' performance in their social setting has been a long tradition. Potentially the work done to date is only a harbinger of more necessary and greater studies ahead. Moreover, topics such as business process are potentially relevant because they allow researchers to catalogue and critique changing machine/employee relations on a process-by-process basis. From such a base, Information Systems research can deploy its technical expertise to support social critique development. What kinds of machine, or deployment of a machine, serve wider societal ends and which do not? Our scholarship might then increasingly benefit human growth as much as it currently serves firm efficiency and economic development. To this end, Maslow, as cited earlier, might become an essential framework for many of our studies (McLeod, 2007).

6. Attention - Christian Grimme, Heike Trautmann

6.1. Introduction

Following Simon, (1971), the ability to consume information in the "information-rich world" is scarce. In other words, information richness implies attention scarcity. Mechanistically, attention is consumed by information. Therefore, the further growth of information requires increasingly complicated decision-making over which information to attend to and which information to ignore. Algorithms intervene in this process, seeking to direct attention to advertisers (e.g., Facebook, Google advertising), or towards the needs of algorithms themselves (e.g., "Update Now," "Your computer needs to close," "Log-on for full service.")

Understanding attention as a scarce resource opens it to economic vocabulary (Davenport & Beck, 2001): *paying* attention to information reduces the amount of overall individually available attention that each can *spend*. We *deal* with attention by recognizing one piece of information and ignoring another piece of information. In the data-centric world, where we are confronted with a massive amount of raw and extracted information ("information overload" see, e.g., Ashton, 1974; Roetzel, 2019), the assignment of attention to the *information market* is essential to receive the maximum *revenue* in terms of the ability to fulfil our tasks in society and work. Thus, the challenge is to identify and consume the most *profitable* information to be up-to-date and *produce* new information.

To separate information into useful (to consume) and useless (to ignore) information, people have started to develop and to accept automated decision support systems such as search engines (e.g., Google), product-related recommender systems (like that implemented in Amazon's platform), media/news aggregators (like Google News), or social media and collaboration platforms that (ideally) function as an individualized filtering mechanism. Click-stream advertising functions as an attention prompt or controller. These systems try to infer from our behaviour, data consumption and data production, which information we need, and which information can be omitted (Bozdag, 2013). This development leads to two major streams of concern in the larger societal context that were discussed during the ERCIS Annual Workshop:

1 The individualised decision-making mechanism is not truly individualized but instead based on and biased by a global classification of information and personal profiles. It can also be used for

surveillance of people and the steering of population sub-groups or individuals (e.g., by governments, companies, or other actors making use of artificial intelligence techniques in a sophisticated manner). A central question is: is it necessary to meet ethical standards in steering and to control the focus of society (i.e., attention) to enable plausible societal outcomes, such as productivity in this new information-centric environment?

2 Strongly connected to (1) but on the individual level, the overload of information can lead to loss of focus and a reduction of personal intellectual capability or discretion, thus reducing individual productivity and wellbeing. This is potentially problematic both at an individual level and at the level of society as a whole. This implies that decision-support systems may be existentially important for managing personal wellbeing and maintaining personal status. The question is: do we want to accept information control and individual information filtering, or do we need a renaissance of human-centric (humanistic) ideals and capabilities?

Fig. 1 addresses the interaction of these two streams of discussion. Global classification and filtering are enabled and supported by the usage of services to reduce information overload. The widespread use of these mechanisms allows central "authorities" (service providers) to build up profiles and classify users. This helps services to improve their quality (in a seemingly individualized way) for users and augments (or is it biases?) the personal view by seemingly important (and sometimes expected) information. This self-reinforcing cycle can make the information consumer vulnerable to external control and might have an unnoticed dependence on filtering mechanisms. Simultaneously, the current information overload results in the tendency to lose capabilities to concentrate and critically reflect (Pennington & Tuttle, 2007; Schick, Gordon, & Haka, 1990), which prevents the user from escaping such a "vicious circle".

6.2. Bias

Media and politics are arenas that deal with aspects of an attention economy in the classical sense of actors *being rewarded with readers' or voters' attention*. The daily struggle of media and politics can be considered a marketplace where consumers are confronted with an offering of information and opinions. They may pick what they want by paying attention. This passive (take what you need) perspective is traditionally complemented by an active perspective, in which media and politics praise their offerings like market barkers vying to gain attention (Webster, 2014). The increasing use of automation to actively promote certain content means that society has shifted towards actively managed attention and away from passively constructed means of allocating attention.

Global classification of information and profiles, which may result into surveillance and steering.

Individual reduction of information overload by decision-making and filtering mechanisms.
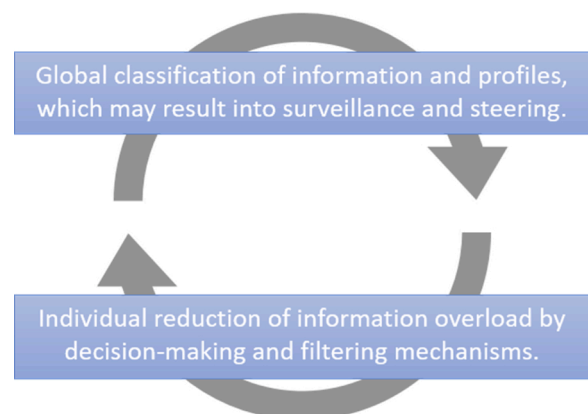
Fig. 1. Interaction of global and individual information filtering in the context of the attention economy.

Modern (social) networks and platforms, together with filter mechanisms and recommender systems, can lead to bias in information provision/consumption and manipulation of political views (Starbird, 2019). Effective individualization of social networks, news platforms, and virtual chat rooms creates closed groups. Filtering mechanisms – originally designed to facilitate network members' attention by bringing helpful and supporting information – lead to so-called "filter-bubbles" (Pariser, 2011). On the other hand, political actors can attack an individual, group, or public attention by artificially created trends or by flooding networks with misinformation. Vehicles like social bots (i.e., programs that support humans in massively spreading misinformation via social media channels) have become an essential technology in the struggle for attention (Grimme, Assenmacher, & Adam, 2018; Grimme, Preuss, Adam, & Trautmann, 2017) and provide for the possible misguidance of people.

The evaluation of the data, information and the associated producers has become of central interest to society. Rudimentary mechanisms have been developed to provide measurement and have become significant over social processes such as establishing popularity, approval, and influence. The more attention (e.g., likes on social media, views for posts and videos, downloads or citations for scientific work) is paid to content; the more critical the content producer becomes in public discussion and reception (Franck, 1999; Onnela & Reed-Tsochas, 2010; Shen & Barabási, 2014; Weng, Flammini, Vespignani, & Menczer, 2012). Thereby attention has become a central market value for humans themselves. It represents the personal importance, influence, and value of a person. It potentially makes some into influential and trustworthy multipliers – those with the highest' likes' and 'views', for example, become important communication hubs and outlets in a digital world. Returning then to the earlier analogy of an attention market, these multipliers become powerful warehouse workers that may significantly influence the offering and consumption of information.

6.3. Ethics

Attention is a precious commodity (Levitin, 2015; Zuboff, 2015). Machine AI can enhance or diminish the utilization and value of attention. The idea of a "human-centred" AI[3] is crucial to exploit AI techniques in a responsible, legal, and ethically sound manner (see CLAIRE project, Confederation of Laboratories for Artificial Intelligence Research in Europe[4]). At this point, machines alone are not yet sophisticated in automated reasoning and should not act autonomously. If individuals are kept in the loop and start to understand the potential conflicts over attention, and that machines might be a solution as well as a problem, then there is likely to be beneficial interest in AI and its use.

An example project that seeks human-centred AI is the recently initiated Humane AI project[5], funded by the European Union's Horizon 2020 research and innovation programme is a large international research consortium in cooperation with industry and political players. It aims at "designing and deploying AI systems that enhance human capabilities and empower both individuals and society as a whole to develop AI that extends rather than replaces human intelligence."

6.4. Control

The discussion provided insights into two major, seemingly contradicting, economic principles behind our attention's "datafication". The first and well-known observation is that global providers of digital services (e.g., multi-service providers like Google, communication and interaction providers like Facebook, commercial platforms like Amazon, and media networks and platforms like YouTube or Netflix) collect and intelligently analyze data to use them in targeted marketing. They systematically generate profiles for all kinds of users to provide commercial customers with these insights or to offer targeted advertisements and information provision via various contact channels. By tracking users' paid attention to ads, global information services can collect additional data and refine their information provision strategies. These marketing support activities generate revenue and should be considered as direct economic exploitation of user attention. Interestingly, data mining in the context of these exploitation activities can be understood as partly an *attention-mining* of users/customers of services.

Simultaneously, modern platforms offer paid service levels that allow the reduction of information overload that is only produced by marketing applications of the same services or by other information systems and technologies. YouTube offers a service level (YouTube Premium) that removes individualized advertisements. Effectively, the user then pays for services to retrieve attention consumed by promotions and information provided in return for platform services. The same holds for information technology itself: Despite collaboration support tools and advanced information filtering, the recovery of attention must be *bought*. An interesting example is a ReMarkable notepad[6] . It enables users to write on an electronic device like writing on paper (and to store and exchange documents via cloud technology) but explicitly abstains from providing communication and interaction interfaces (email, chat, collaboration platforms) to ensure undivided attention to the working task. Increasingly, with work-tools, collaborative apps, and the like, the deal is that users must pay to recover their attention.

Summarizing both aspects, the exploitation of attention as an economic good is twofold: (1) users essentially pay with attention, when using the (seemingly) free services offered online. (2) To retrieve personal attention, users must pay again, but this time financially. From an economic point of view, this enables an almost unlimited stream of (monetary) transactions.

6.5. Responsibility

About individuals' attention, unconditional trust of machines will neglect the need for one's own decisions and critique how one's time is spent. There are potential consequences of this for individuals themselves, and perhaps society more generally, as attention is co-opted by these machines (Zuboff, 2019). The support of advanced machines for decision-making should not come at the cost of compromising individual attention. An essential requirement is to hold responsibility with the individual. To achieve this, not only must people know how their attention is addressed and consumed, but also society must have access to sufficient explanation of AI. Machine-learning algorithms must find a compromise between gaining maximum outcome quality and ensuring a satisfactory level of explainability of the underlying mechanisms; thus necessitating increased research focus on explainable AI (e.g., Guidotti et al., 2018; Lundberg et al., 2019; Molnar, 2019).

6.6. Future research

Information overload is not a new topic. There was a discussion on countermeasures to it ten or twenty years ago (Savolainen, 2007; Whittaker & Sidner, 1997). Even then, a common issue was email overload. Researchers explored the behaviour of users and identified *filtering* as well as *withdrawal* strategies. The first strategy comprised individual cognitive prioritization followed by subsequent, manual selection of consumed information, while the latter strategy essentially consisted of unplugging a service (e.g., abstain from email).

In modern society, filtering strategies have outlived withdrawal

---

[3] The Age of Artificial Intelligence. Towards a European Strategy for Human-Centric Machines, EPSC Strategic Notes, European Political Strategy Centre, Issue 29, 27 March 2018.
[4] www.claire-ai.org
[5] https://www.humane-ai.eu

[6] https://remarkable.com

strategies and are now a standard part of information consumption and personal orientation. Technologies provide powerful support in searching, sharing, and interpreting information[7]. Further, information and attention have more obviously become economic resources underpinning the profits of major companies and supporting the careers of politicians. Garnering attention can be exchanged for societal inclusion, reputation, and power. Consequently, unplugging from this system has become very difficult if not impossible. This is even though this economic exploitation of information and attention increases information overload and the management of attention in a self-reinforcing spiral. As filtering leads to more precise targeting of information, the individual must further optimize that filter to continuously prevent it from becoming overloaded. There is a still increasing need for filtering and decision-making technologies. The real-time focus of society, which demands immediate attention and (re-) action makes these systems still more significant.

Under this setting and the virtual impossibility of unplugging from modern information technologies, some important research questions arise:

1. Is there no longer the option of giving time to reflection and reasoning, and will we cognitively degenerate? Alternatively, is it possible to adapt to information filtering challenges and keep at least partial independence from the machine (e.g., by becoming skilled in multi-tasking or through education)?
2. Is there any way to counteract the self-reinforcing spiral of technology support in attention economy or are machines *needed*? Can we escape the market of attention without losing personal or group reputation within a modern society?
3. If machines are needed (maybe more than today), how can we trust them and what fosters or hampers trust in their decision-making? Is transparency of methods (e.g., in AI) sufficient or just a vehicle to increase confidence and encourage carelessness in an increasingly information-abundant world?
4. Who is responsible for assessing the status quo, challenges, opportunities, and risks of misusing the distribution mechanisms of human attention objectively? Can frameworks be established or, on a meta-level, are we facing the threat of giving this away to machines as well? Is this in line with ethical standards we agree to in society?

From a humanistic (and thus human-centric) point of view, the discussion poses further challenges for humans' self-conception. If we give away personal responsibility of distributing our attention and the ability to reflect on decisions we make systematically, we also give up a central feature of humanism - namely individuality (Harari, 2015).

In the struggle for an attention management solution, we must find a trade-off between the dogma of individualism and the increasing importance of standards to protect our limited attention resources.

7. Wellbeing - Joerg Becker, Katrin Bergener, Armin Stein

7.1. Introduction

A definition of wellbeing is not obvious (Dodge, Daly, Huyton, & Sanders, 2012), but coalesces as the idea "[…] that wellbeing is a multi-dimensional construct." However, the authors conclude, that, "[i]n essence, stable wellbeing is when individuals have the psychological,

social and physical resources they need to meet a particular psychological, social and/or physical challenge" (Dodge et al., 2012). Increasing resources whilst keeping challenges stable results in an overall better state of wellbeing, whereas an increase in challenges without additional resources might reduce an individual's overall wellbeing as depicted in Fig. 2.

Also, wellbeing is associated with the idea that individuals might self-manage through self-reflection potentially further enabled by health apps and personal data (Topol, 2019). Against this backdrop, a question then develops of how technology affects the resources and the challenges of individuals, and whether this is desirable.

7.2. Bias

In the discussion, a very intuitive example was given that introduces a machine to a scenario, where it could empower the resources available to the agent to meet a specific challenge: This is as follows. During massive natural or human-made disasters like earthquakes, tsunamis, pandemics, or war, doctors and medical staff are often overwhelmed by patients and the chaotic environment. The emergency responders must decide whom to treat first, and who will be treated later under stressful situations. Typically, the first forty-eight hours are of uttermost importance (Schultz, Koenig, & Noji, 1996). An algorithm might feasibly provide support, helping medical workers prioritise their actions (who to help and who to leave for later) according to the patient's injury. This has an ethical implication of reducing the responder's burden of choice. This, in turn, might beneficially increase their mental resources to treat priority patients effectively. Logically, the scenario makes sense but under discussion limitations and complexities appear. An injured person, perhaps suffering from psychological and potentially lethal physical challenges might prefer to be judged by a human being (understood as a positive resource) rather than a machine (negative resource). There might be psychological reasons for this, such as the value of human interaction at a point of distress. There might also be fear of how the machine has been trained and whether it exhibits bias of some cause, e. g., skin colour or ethnicity (Yapo & Weiss, 2018). The ethical balance in such a situation is complicated. It is understood that humans are also biased in general, especially regarding race (see, e.g., Loiacono et al., 2013), yet this is understood as "natural" bias and more acceptable or at least identifiable and confrontable. In contrast, a machine is expected to be ultimately rational and removed from immediate scrutiny.

An alternative example of the intrusion of machines into wellbeing is the potential effect of the automated selection of news, e.g., social media. Often, online news consumed by individuals stems from single sources representing a unique ideological perspective (Flaxman, Goel, & Rao, 2016, p. 313). Whereas people expect quality papers to have a diverse and more heterogeneous choice of articles, the selection of (online) news by algorithms is usually based on the reader's identified interest. The information has usually been collected during past sessions, which might again enforce a singular perspective on topics. Especially when consuming negative news, this can lead to re-enforcing patterns leading to psychological challenges. This becomes especially apparent when using social media like Twitter (Flaxman, Goel, & Rao, 2013).

7.3. Ethics

The issue of bias immediately opens the question of ethics. Machines are omnipresent in health and care industries: Robots support surgeons during operations, making surgical treatments possible that were deemed too dangerous before (see, e.g., Kroh & Chalikonda, 2015). In geriatric care, artificial animals, and care robots (Wu, Fassert, & Rigaud, 2012) are being tested and already used to entertain and look out for older adults. Questions follow as to the ethics of these (e.g., are people entitled to human carers and "real" animals). Also, similar to issues of bias section, should machines substitute humans in the healthcare sector

---

[7] This certainly also holds for the authors' research on this topic by using online search engines (like Google), automated suggestions provided by publishers' online recommender services (e.g. Springer Link), hints in social networks (e.g. ResearchGate), and book recommendations by commercial platforms (e.g. Amazon). However, the authors claim to have intensively reflected on all of the literature statements in the context of the ERCIS Annual Workshop discussions.

Fig. 2. Definition of Wellbeing depicted as see-saw according to (Dodge et al., 2012).

to make decisions that have consequences about life and death. In the discussion, there was unanimous agreement that a machine should not decide when life-supporting systems can or should be switched off.

The discussants wondered whether it would be ethical to let robots treat people who are no longer able to distinguish themselves who or what it is that takes care of them (human/robot). It is worth mentioning that this discussion is already quite elaborate in medicine (see, e.g., Vandemeulebroucke, Dierckx de Casterlé, & Gastmans, 2018). Still, it has not dispersed to the field of Information Systems. Arguments against the use of care robots were that even though the patients might not recognize the difference between a human or machine anymore, the machine might lack empathy, interest in the person, and/or personal experience. Even if the patient does not know the difference, perhaps we, as a society, wish to represent ourselves through human contact rather than efficient machine mechanisms alone.

A counterexample further revealed the complexity of negotiating these boundaries. Discussants' agreed that they would consciously accept robots taking the lead during complex surgery. Interestingly, the reasons were partially the same as those used for rejecting care by machines in other circumstances: no distraction by emotions, no interest in the person, and no tiredness.

In light of Dodge et al. (2012) definition, we understand the resource view in these cases as everything that might help the patient recover, which should increase their wellbeing. Physical or psychological issues may negatively affect wellbeing, i.e., what brought the patients to their current state. We might argue that in the first case discussed (a decision about life and death) the same arguments should come into play as in the bias section: It would be unethical to give away the responsibility of such a decision to a non-human machine – the machine would therein become a negative resource affecting wellbeing. In a situation of care provision, machines may or may not be perceived as a positive resource influencing the patient's wellbeing, whereas, for surgery, they are perceived as positive.

7.4. Control

In our discussion, we extensively questioned if wellbeing includes "life satisfaction." If someone is happy with his/her current situation, even if it is not healthy according to common standard measures, should machines interfere with this? Examples we discussed were situations like sitting on the couch with a bag of potato crisps or going to a party and drinking too much beer. Both situations are not healthy according to a typical medical understanding and machines (like a fitness tracker) could point out unhealthy behaviour. Is this good for "life satisfaction"? This concept of "life satisfaction" is also described in the literature on subjective wellbeing as "evaluative wellbeing" (Steptoe, Deaton, & Stone, 2015), i.e., how satisfied people are with their lives. This does not necessarily correlate with healthy living. From this follows a utilitarian argument of whether people have the responsibility to stay healthy to avoid putting a strain on their country's healthcare system. Do we have to keep fit as service to society? Should machines actively point to our misbehaviour to facilitate this and take action if we do not act

accordingly? Do we then ultimately lose our freedom of choice to spend an irresponsibly unhealthy night with friends? What counts more: our freedom of choice or society as the greater good, i.e., "societal wellbeing"? If we let machines control our health wellbeing, we discussed that this might lead to technostress (Ayyagari, Grover, & Purvis, 2011), as people could feel obliged to behave as a controlling machine demands.

Apart from discussing the control of health issues, we discussed the development of interactions and communication in general, and the possible influence on our social exchanges from contact with AI tools like Alexa or Google Home in everyday life. How do we talk to machines? Do we have to be polite or not because it is "just a machine"? Do we get used to a reduced way of talking? Might it change how we talk to other humans? Will we be less polite or won't the increased interaction with AI machines make a difference?

The last topic we touched in our discussion concerning *wellbeing and control* were the implications of virtual reality, augmented reality, and sensors. If those technological advances become so good that we could not realize if we are at the beach or in VR, what would be the implications? Would people like to be "in the matrix"? Studies so far show that AR and VR can positively promote physical, social, and psychological wellbeing of adults (Lee, Kim, & Hwang, 2019). Does this mean that "the matrix" is to be encouraged?

7.5. Responsibility

Who is responsible for our wellbeing? This question underpins the relationship between machines and humans. The answer is us. We must keep responsibility for our wellbeing and ourselves. Besides, what influences our wellbeing is, of course, our surrounding. Our family, friends, colleagues, leaders, government, and society as a whole all influence us. In our community today, people like to define themselves through their job (Miscenko & Day, 2016; Walsh & Gordon, 2008); it also influences their wellbeing. It follows that responsibility shifts with changes in the job market where jobs are continuously delegated to machines and human work is continually replaced, not only in assembly-line work but also in other sectors like sales or analytics (where AI outperforms humans). Thus, in a situation where even knowledge workers are replaced by machines, there is a potential shift in responsibility. This brings the question of how people will feel about that. Will people enjoy the freedom of not having to work? Will they become depressed because they do not feel needed anymore? Will the transfer of jobs create a greater societal responsibility for the wellbeing of those affected?

7.6. Future research

Summing up our discussion on wellbeing, several topics might be worth researching. From an IS perspective, we need to clarify the differentiation of objective vs subjective wellbeing and if machines might support us in achieving one or the other or both.

An interesting topic to investigate is to have a closer look at how

younger generations feel about the issues that we discussed. It might be that the younger generation and digital natives would not expound the problems with wellbeing and machines in the same way. If they grow up with AI machines like Alexa, it will potentially feel normal for them to interact with AI in the same way they interact with humans.

8. Discussion and agenda for research

The previous sections have summarised the main themes that emerged from workshop discussions among IS experts regarding humanity and intelligent machines at the ERCIS annual general meeting in September 2019. The discussion focused on the need for new humanism perspectives to understand and explain our relations with intelligent machines. These relations were considered from four perspectives, crime and conflict, jobs, attention, and wellbeing. This section synthesizes the common threads that were evident across the workshop discussions through the lens of human freedoms. This discussion outlines the many strands of the recommended research agenda, and outlines the pressing research questions, summarised in Table 1.

The contribution by *Niels F. Garmann-Johnsen* and *Marcelo Fantinato* summarised the workshop discussion regarding humanity and intelligent machines in the context of crime and conflict. The discussants indicated that should intelligent machines be used for making legal decisions; then this change could only be acceptable if the machines were able to improve on human jury decision-making. To achieve this improvement, it would be critical that intelligent machine decision-

making is free from bias and does not lead to socially undesirable decisions. Keeping the human in the decision-making loop and ensuring that the ultimate decision-making control always remains with humans were identified as essential mechanisms to achieve this goal. Institutional rules were also identified as necessary to ensure continued responsible use of intelligent machines in the context of crime and conflict. The discussion revealed several important avenues for further research. For example, investigating how machines can be calibrated to make socially acceptable predictions of human behaviour (RQ1.1), avoiding unseen bias in data and logically sound but socially deleterious decisions (RQ1.2). It would also be valuable to investigate how humans can be retained in decision-making loops to build better social processes for justice systems (RQ1.3); how the institutional responsibilities of a court may be designed to account for the risk of dataset manipulation or bias (RQ1.4); and how human justice system stakeholders should be educated for interpreting machine recommendations (RQ1.5). Finally, further research would be valuable to determine how responsibility for machines' contribution to legal decision-making processes should be determined (RQ1.6).

João Álvaro Carvalho's contribution summarised the discussants' views regarding the impacts on intelligent machines on jobs and the implications for human freedoms. Intelligent machines are already influencing employee selection and promotion processes. It was acknowledged that these machines have the advantage of ensuring consistent rule application, but that in the more complex world of society, such algorithms have the potential to be reductive and unfair. The

Table 1
Summary of Research Questions for AI and Humanity Research.

AI and Humanity Themes Cross-cutting issues	1. Crime and Conflict	2. Jobs	3. Attention	4. Wellbeing
Bias	RQ1.1 How can machines be calibrated to make socially acceptable predictions of human behaviour by making a comparison with actual behaviour? RQ1.2 How can machines be designed to avoid repeating prejudice, the unseen bias in data, or logically sound but socially deleterious conclusions?	RQ2.1 How can job selection and promotion algorithms be designed to ensure fair and equitable decisions?	RQ3.1 How can we adapt to information filtering challenges and keep at least partial independence from the machine (e.g., by becoming skilled in multi-tasking or through education)?	RQ4.1 How can AI be designed and applied in safety-critical environments to support first responders' decision-making regarding people's wellbeing? RQ4.2 How can AI be designed to ensure unbiased, socially acceptable choices are made?
Ethics	RQ1.3 How should the human in the loop be retained to build better social processes relating to crime and justice?	RQ2.2 What do people gain from adapting their skills for AI and robotics and what do they lose? RQ2.3 How could new employment schemas from the widespread adoption of machines be designed to avoid increasing inequalities and provoking social fissures?	RQ3.2 Is there any way to counteract the self-reinforcing spiral of technology support in attention economy or are machines needed? RQ3.3 Can we escape the market of attention without losing personal or group reputation within a modern society?	RQ4.3 How can AI deployment conditions for health and social care (e.g., surgery vs social care) be determined? RQ4.4 How can dynamic mechanisms be designed to account for people's varying attitudes to AI deployment for health and social care?
Control	RQ1.4 How should the court's institutional responsibilities be designed to account for the risk of dataset manipulation or bias? RQ1.5 How should human justice system stakeholders be educated for interpreting machine recommendations?	RQ2.4 If machines can do what we find interesting and what we find dull, how should we determine which work machines should perform?	RQ3.4 If machines are needed (maybe more than today), how can we trust them and what fosters or hampers trust in their decision-making? RQ3.5 Is transparency of methods (e.g., in AI) sufficient or just a vehicle to increase trust and encourage carelessness in an increasingly information-abundant world?	RQ4.5 How can the conditions for AI deployment for social wellbeing be determined? How should machines respond to unhealthy behaviours?
Responsibility	RQ1.6 How should responsibility for the contribution of machines in legal decision-making processes be determined?	RQ2.5 If machines are increasingly in charge of decisions in the workplace, then how does society vet them and in whose interests do they work?	RQ 3.6 Who is responsible for assessing the status quo, challenges, opportunities, and risks of misusing the distribution mechanisms of human attention objectively? RQ3.7 Can frameworks be established or, on a meta-level, are we facing the risk of giving this away to machines as well? Is this in line with ethical standards we agree to in society?	RQ 4.6 How may technological unemployment from AI and machines influence people's sense of identity and wellbeing? RQ4.7 Will the transfer of jobs to machines create a greater societal responsibility for the wellbeing of those affected?

discussants also agreed that the future of work for some occupations looks bleak. Suppose high levels of technological unemployment are created from the implementation of intelligent machines. In that case, there may be a considerable disruption for the known social order, aggravating inequalities and provoking social fissures. The redistribution of wealth is already creating a dilemma and raising ethical issues for society. However, the discussants also highlighted that intelligent machines could enrich our working lives, enhancing human capabilities to previously unachieved levels. These observations emphasise several important directions for future research. For example, further research investigating how job selection and promotion algorithms may be designed for fair and equitable decisions would be valuable (RQ2.1). An exploration of the skills that people gain and lose from adapting to work with AI and robotics would help understand how jobs were evolving and changes necessary for education and training systems (RQ2.2). Some critical societal issues require further investigation such as how new employment schemas may be designed to avoid increasing inequalities and provoking social fissures (RQ2.3); how should organizations decide which work tasks should be performed by machines (RQ2.4); and if intelligent machines are increasingly in charge of decisions in the workplace, how should society vet them and in whose interests do they work (RQ2.5)?

Christian Grimme and Heike Trautmann provided a summary of the discussants' views regarding the intelligent machines and humanity from the perspective of attention. It was acknowledged that the market for our attention has become economically lucrative for third parties and led to a vast supply of media designed to compete for our attention. Consequently, we are becoming increasingly reliant on filter mechanisms and recommender systems that may lead to undesirable 'filter bubbles' that preference narrowly selected perspectives and information. The discussants also highlighted an important contradiction regarding global digital services providers' behaviours (e.g., Facebook, Google). On the one hand, these digital service providers actively seek to exploit user attention for economic gain from selling customer insight data or targeted advertisement opportunities. On the other hand, these same providers offer paid service levels that reduce information overload. The user is paying for services to retrieve attention consumed by advertisements and information that were initially provided in return for platform services. This cycle suggests the need for one's own decisions and critique of how one's time is spent, rather than placing unconditional trust in intelligent machines. Thus, people must know how their attention is addressed and consumed, but also society must have access to sufficient explanation of AI to make these choices. The debates identified in this section suggest several exciting and important avenues for further research. Research that investigates how we can respond to information filtering challenges and keep some independence from intelligent machines would be important to help policymakers and regulators respond to concerns about the power of digital service providers (RQ3.1). A further important research stream is needed to investigate whether the self-reinforcing spiral of technology support in the attention economy is inevitable (RQ3.2) and whether it would be possible to escape the market of attention without losing personal or group reputation (RQ3.3). It would also be necessary to revisit the concept of trust in technology that has often been the attention of previous IS research scholars. For the new generation of intelligent machines, how can trust mechanisms be designed into these machines (RQ3.4) and will these mechanisms (e.g., explainable AI) be sufficiently reliable to ensure society does not become complacent to the exploitation of attention (RQ3.5). At a societal and policy level, the discussion highlighted essential research questions regarding which bodies should be responsible for assessing the challenges, opportunities, and risks of misusing the distribution mechanisms of human attention (RQ 3.6) and whether frameworks can be established for this purpose without relying on intelligent machines for these activities as well (RQ3.7).

Joerg Becker, Katrin Bergener and *Armin Stein* summarised discussants views concerning the impacts of intelligent machines for humanity and

the implications for our wellbeing. They observe that stable wellbeing is dependent on achieving a suitable balance between psychological, social and physical resources and psychological, social and/or physical challenges. Discussants highlighted the challenges associated with decision-making for first responders in safety-critical environments. They suggest that intelligent machines could provide valuable, timely support and guidance but could also offer socially undesirable recommendations that are logical but not acceptable to the social actors involved. The discussants also recognized the risk of intelligent machine decision recommendations potentially reinforcing data biases. Discussants were unanimous in their agreement that a machine should not decide at what point in time life-supporting systems can or should be switched off. However, there was a contradiction in the discussants' views on the circumstances when the human could be taken out of the decision-making loop to deliver healthcare. Discussants agreed that they would consciously accept robots taking the lead during complex surgery but would be less willing to hand over control in social care settings, and not for life-or-death decisions. Interestingly, the reasons for accepting care were similar to those used for rejecting care in other settings: no distraction by emotions, no interest in the person, and no tiredness. Such attributes were considered advantageous for machines undertaking surgery but not for machines operating in social care environments. A further issue identified by discussants concerned individuals' responsibility for their wellbeing and whether intelligent machines should interfere with unhealthy behaviours. It was also recognized that a person's wellbeing and identity is often tightly linked to their job and place of work. Should intelligent machines remove some of these jobs, then this change may have a detrimental impact on people's wellbeing, requiring governmental policy responses. These issues highlight several important research areas warranting further investigation. For example, research that explores how AI can be designed and applied in safety-critical environments to provide unbiased, socially acceptable decision-making processes for first responders (RQ4.1 and 4.2) would be extremely valuable. It would be beneficial to investigate how to determine whether conditions are suitable for AI deployment in various settings (RQ4.3 and RQ4.5) and how the dynamic nature of people's varying attitudes to AI deployment in these settings can be accounted for (RQ4.4). It would also be necessary for research to explore how technological unemployment from intelligent machines may influence people's sense of identity and wellbeing (RQ 4.6) and whether there will be a greater societal responsibility for the wellbeing of those affected (RQ4.7).

9. Conclusions

This study presents a consolidated perspective on the new relations between intelligent machines and humanity from European IS academic experts. The collective insights stem from a workshop held at the annual European Research Center for Information Systems (ERCIS) meeting held at Loughborough University, United Kingdom in September 2019. The workshop discussions focused on four AI and humanity themes: Crime & conflict, Jobs, Attention, and Wellbeing. The critical issues raised in the workshop for each of these themes have been summarized and aligned with the academic literature. These insights have provided the foundation for outlining an exciting range of research questions to guide future IS research.

This article's title poses an important question regarding the essential aspects of humanity that should not be given up to intelligent machines. The outcomes of the workshop discussions suggest three attributes of humanity that need to be preserved.

First, humans must retain the role of meaningful, responsible critique of the design and application of AI, and the intelligent machines it can create. Critique is a vital concept that humanity can retain as a means to ensure liberation from intelligent machines. Suppose intelligent machines are used to help shape decision processes in life-changing situations, such as criminal court proceedings, or to aid emergency care first

responder in disaster situations. In that case, they should serve only as referees or expert guides and not as decision-makers. It is critical that such machine 'referees' or 'guides' should be subject to constant human critique.

Second, a human must be kept in the loop of intelligent machine decision-making processes. This involvement is vital to preserve our ability to systematically reflect on the decisions we make, which ultimately influence our individuality, a central feature of humanism. Suppose individuals are involved in the decision process and start to understand the potential conflicts over, for example, how our attention may be distributed. In that case, we will be in a stronger position to critique the risks and benefits of intelligent machines. Further, the balance of human involvement in intelligent machine decisions may be acceptable at one point in time, but not when circumstances change. Therefore, the challenge is not to find an optimum point for human inclusion in machine decision-making, but to determine how this point can be dynamically assigned as circumstances change.

Third, humans must retain the ability to interpret and explain intelligent machine decision-making processes. Intelligent machines must be designed to be traceable through the entire path taken by the algorithm before producing the final result. The ability to explain AI decisions is critical for supporting the ability of humans to critique AI decisions. While this requirement is highly problematic for many AI techniques, such as some types of neural networks and deep learning, it is critical to ensure that algorithmic machines' cost savings and accuracy advantages do not create costs and issues elsewhere in the system.

These three attributes are potentially critical to a new framing of human/machine relations. Taking the cue from *Stacey* (see Theoretical Background) that humanity must retain the ability of critique and seek a Bauhausian human-machine unity, these attributes then specify the terms of that critique. The new Bauhausian unity relies first on accepting this role of critique and then, secondly, deliberately enabling it through designing humans into the loop of intelligent machine decision-making and, thirdly, by enabling this through a traceable relationship – the unity requires that humans can interpret and explain intelligent machine decision-making progress. Through the acceptance of these terms, a new theoretical fork is reached, and the human requirements that have been "de-levelled" in a debate are elevated and prioritized in a re-turn to a newly expressed humanism (Stacey, 2019).

The AI humanity debate is rich and highly complex. This article has illustrated how examining the debate is critical, not only for the design and development of AI but also for how it should be applied within different social and cultural contexts. Inevitably, this article could only touch on these critical issues. However, consolidating IS expert perspectives provides an excellent foundation for reflecting on what might be lost and retained when forming new relations with intelligent machines.

CRediT authorship contribution statement

Crispin Coombs: Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision, Project administration. **Patrick Stacey:** Conceptualization, Writing - original draft, Writing - review & editing. **Peter Kawalek:** Conceptualization, Writing - original draft, Writing - review & editing. **Boyka Simeonova:** Conceptualization, Writing - original draft, Writing - review & editing. **Joerg Becker:** Conceptualization, Writing - original draft, Writing - review & editing. **Katrin Bergener:** Conceptualization, Writing - original draft, Writing - review & editing. **João Álvaro Carvalho:** Conceptualization, Writing - original draft, Writing - review & editing. **Marcelo Fantinato:** Conceptualization, Writing - original draft, Writing - review & editing. **Niels F. Garmann-Johnsen:** Conceptualization, Writing - original draft, Writing - review & editing. **Christian Grimme:** Conceptualization, Writing - original draft, Writing - review & editing. **Armin Stein:** Conceptualization, Writing - original draft, Writing - review & editing. **Heike Trautmann:** Conceptualization, Writing - original draft, Writing -

review & editing.

References

Abubakar, A. M., Behravesh, E., Rezapouraghdam, H., & Yildiz, S. B. (2019). Applying artificial intelligence technique to predict knowledge hiding behavior. *International Journal of Information Management, 49*, 45–57. https://doi.org/10.1016/J. IJINFOMGT.2019.02.006.

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *The Journal of Political Economy, 128*(6), 2188–2244.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052.

Adarves-Yorno, I., Postmes, T., & Haslam, S. A. (2007). Creative innovation or crazy irrelevance? The contribution of group norms and social identity to creative behavior. *Journal of Experimental Social Psychology, 43*(3), 410–416. https://doi.org/10.1016/J.JESP.2006.02.013.

Ashton, R. H. (1974). Behavioral implications of information overload in managerial accounting reports. *Cost and Management*, 37–40.

Associated Press. (2019). *YouTube fined $170m for collecting children's personal data*. The Guardian. https://www.theguardian.com/technology/2019/sep/04/youtube-kids-fine-personal-data-collection-children-.

Ayyagari, R., Grover, V., & Purvis, R. (2011). Technostress: Technological antecedents and implications. *MIS Quarterly: Management Information Systems, 35*(4), 831–858. https://doi.org/10.2307/41409963.

Baraniuk, C. (2016). Computer paints "new Rembrandt" after old works analysis. *BBC news*.

Borges, A. F. S., Laurindo, F. J. B., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2020). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management*, 102225. https://doi.org/10.1016/J.IJINFOMGT.2020.102225.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology, 15*(3), 209–227. https://doi.org/10.1007/s10676-013-9321-6.

Brynjolfsson, E., & McAfee, A. (2015). Will humans go the way of horses? *Foreign Affairs, 8*.

Brynjolfsson, E., & McAfee, A. (2016). The second machine age - work. *Progress, and prosperity in a time of brilliant technologies*. WW Norton & Co.

Calo, C. J., Hunt-Bull, N., Lewis, L., & Metzler, T. (2011). Ethical implications of using the Paro robot with a focus on dementia patient care. *Association for the advancement of artificial intelligence workshop* (pp. 20–24). WS-11-12 http://www.scopus.com/inward/record.url?eid=2-s2.0-80054913481&partnerID=40&md5=cf9ad1b74f81299aa3f48be4f8e9700f.

Coombs, C. (2020). Will COVID-19 be the tipping point for the Intelligent Automation of work? A review of the debate and implications for research. *International Journal of Information Management, 55*, 102182. https://doi.org/10.1016/J. IJINFOMGT.2020.102182.

Coombs, C., Hislop, D., Taneva, S., & Barnard, S. (2020). The changing nature of work in the age of intelligent machines. *Oxford handbook of digital technology and society* (p. 800). US: Oxford Univ Press. https://global.oup.com/academic/product/the-oxford-handbook-of-digital-technology-and-society-9780190932596?cc=eu&lang=en&#.

Coulton, P., & Lindley, J. G. (2019). More-than human centred design: Considering other things. *The Design Journal, 22*(4), 463–481. https://doi.org/10.1080/14606925.2019.1614320.

Davenport, T. H., & Beck, J. C. (2001). *The attention economy: Understanding the new currency of business*. Harvard Business School Press.

Dodge, R., Daly, A., Huyton, J., & Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing, 2*(3), 222–235. https://doi.org/10.5502/ijw.v2i3.4.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – Evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71. https://doi.org/10.1016/J.IJINFOMGT.2019.01.021.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., & Williams, M. D. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 101994. https://doi.org/10.1016/J. IJINFOMGT.2019.08.002.

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., & Wang, Y. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, 102168. https://doi.org/10.1016/J.IJINFOMGT.2020.102168.

Elliott, M. S., & Kraemar, K. L. (2008). *Computerization movements and technology diffusion: From mainframes to ubiquitous computing*. Information Today.

Fan, C., Zhang, C., Yahja, A., & Mostafavi, A. (2021). Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management, 56*, 102049. https://doi.org/10.1016/J.IJINFOMGT.2019.102049.

Flaxman, S., Goel, S., & Rao, J. M. (2013). Ideological segregation and the effects of social media on news consumption. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2363701.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly, 80*(Specialissue1), 298–320. https://doi.org/10.1093/poq/nfw006.

Franck, G. (1999). Scientific communication–a vanity fair? *Science, 286*(5437), 53–55.

Gabriel, J. (2015). *Young people's experiences of moving out of being 'not in education, employment or training' (NEET): An exploration of significant factors*. University of Birmingham. https://core.ac.uk/download/pdf/33528766.pdf.

Grimme, C., Assenmacher, D., & Adam, L. (2018). Changing perspectives: Is it sufficient to detect social bots? In G. Meiselwitz (Ed.), *Social computing and social media. User experience and behavior*. Springer.

Grimme, C., Preuss, M., Adam, L., & Trautmann, H. (2017). *Social bots: Human-like by means of human control?* ArXiv. http://arxiv.org/abs/1706.07624.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D., & Giannotti, F. (2018). *A survey of methods for explaining Black Box models*. http://arxiv.org/abs/1802.01933.

Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management, 49*, 157–169. https://doi.org/10.1016/J.IJINFOMGT.2019.03.008.

Harari, Y. N. (2015). *Homo deus: A brief history of tomorrow*. Penguin.

Hegel, G. W. F. (2018). *Hegel: The phenomenology of spirit*. Oxford University Press.

Heidegger, M. (1954). *Vorträge und Aufsätze*. Günther Neske.

Hu, Q., Lu, Y., Pan, Z., Gong, Y., & Yang, Z. (2021). Can AI artifacts influence human cognition? The effects of artificial autonomy in intelligent personal assistants. *International Journal of Information Management, 56*, 102250. https://doi.org/10.1016/J.IJINFOMGT.2020.102250.

Iacono, S., & Kling, R. (2001). Computerization movements: The rise of the internet and distant forms of work. *Information technology and organizational transformation: History, rhetoric, and practice*. London, England: SAGE PublicationsSage UK.

Johnsen, H. C. G., Holtskog, H., Ennals, J. R., & John, R. (2018). *Coping with the future: Rethinking assumptions for society, business and work*.

Kellner, D., & Lewis, T. E. (2007). Cultural critique. *The blackwell encyclopedia of sociology*. John Wiley & Sons, Ltd.. https://doi.org/10.1002/9781405165518.wbeosc172

Kroh, M., & Chalikonda, S. (Eds.). (2015). *Essentials of robotic surgery*. Springer International Publishing. https://doi.org/10.1007/978-3-319-09564-6.

Lee, L. N., Kim, M. J., & Hwang, W. J. (2019). Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. *Applied Sciences, 9*(17), 3556. https://doi.org/10.3390/app9173556.

Levitin, D. J. (2015). *Why it's so hard to pay attention, explained by science*. FastCompany.Com.

Liu, D., Santhanam, R., & Webster, J. (2017). Toward meaningful engagement: A framework for design and research of gamified information systems. *MIS Quarterly, 41*(4), 1011–1034. http://enterprise-gamification.com.

Loiacono, E., Urquhart, C., Beath, C., Craig, A., Thatcher, J. B., Vogel, D. R., & Zigurs, I. (2013). Thirty Years and Counting: Do We Still Need the ICIS Women's Breakfast? *Communications of the AIS, 33*(6), 81–96.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., & Lee, S.-I. (2019). *Explainable AI for trees: From local explanations to global understanding*. http://arxiv.org/abs/1905.04610.

Luxton, D. D. (2014). Recommendations for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence in Medicine, 62*(1), 1–10. https://doi.org/10.1016/j.artmed.2014.06.004.

Markard, J., Wirth, S., & Truffer, B. (2016). Institutional dynamics and technology legitimacy – A framework and a case study on biogas technology. *Research Policy, 45*(1), 330–344. https://doi.org/10.1016/J.RESPOL.2015.10.009.

Marr, B. (2019). *Artificial intelligence can now write amazing content – what does that mean for humans?* Forbes.

McDermid, J. (2019). *Autonomous decision making and robotics research*. University of York. https://www.york.ac.uk/research/themes/risk/autonomous-decision-making-robotics/.

McLeod, S. (2007). Maslow's hierarchy of needs. *Simply Psychology, 1*, 1–8.

Metz, C. (2019). *A breakthrough for A.I. technology: Passing an 8th-Grade science test*. New York Times.

Miscenko, D., & Day, D. V. (2016). Identity and identification at work. *Organizational Psychology Review, 6*(3), 215–247. https://doi.org/10.1177/2041386615584009.

Molnar, C. (2019). *Interpretable machine learning – A guide for making black Box models explainable*. Lulu.

Molnar, W., Nandhakumar, J., & Stacey, P. (2017). A paradox of progressive saturation: The changing nature of improvisation over time in a systems development project. *Journal of the Association for Information Systems, 18*(11), 814–836. https://doi.org/10.17705/1jais.00472.

Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management, 53*, 102104. https://doi.org/10.1016/J.IJINFOMGT.2020.102104.

Onnela, J.-P., & Reed-Tsochas, F. (2010). Spontaneous emergence of social influence in online systems. *Proceedings of the National Academy of Sciences of the United States of America, 107*(43), 18375–18380. https://doi.org/10.1073/pnas.0914572107.

Orlikowski, W. J., & Gash, D. C. (1994). Technological frames. *ACM Transactions on Information Systems (TOIS), 12*(2), 174–207. https://doi.org/10.1145/196734.196745.

Orlikowski, W. J., & Scott, S. V. (2008). Sociomateriality: Challenging the separation of technology, work and organization. *The Academy of Management Annals, 2*(1), 433–474. https://doi.org/10.5465/19416520802211644.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.

Pennington, R., & Tuttle, B. (2007). The effects of information overload on software project risk assessment. *Decision Sciences, 38*(3), 489–526. https://doi.org/10.1111/j.1540-5915.2007.00167.x.

Roetzel, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research, 12*(2), 479–522. https://doi.org/10.1007/s40685-018-0069-z.

Savolainen, R. (2007). Filtering and withdrawing: Strategies for coping with information overload in everyday contexts. *Journal of Information Science, 33*(5), 611–621. https://doi.org/10.1177/0165551506077418.

Schick, A. G., Gordon, L. A., & Haka, S. (1990). Information overload: A temporal approach. *Accounting Organizations and Society, 15*(3), 199–220. https://doi.org/10.1016/0361-3682(90)90005-F.

Schippers, B. (2018). *Why technology puts human rights at risk*. The Conversation. https://theconversation.com/why-technology-puts-human-rights-at-risk-92087.

Schultz, C. H., Koenig, K. L., & Noji, E. K. (1996). A medical disaster response to reduce immediate mortality after an earthquake. *The New England Journal of Medicine, 334*(7), 438–444. https://doi.org/10.1056/NEJM199602153340706.

Shen, H.-W., & Barabási, A.-L. (2014). Collective credit allocation in science. *Proceedings of the National Academy of Sciences of the United States of America, 111*(34), 12325–12330. https://doi.org/10.1073/pnas.1401992111.

Silberg, J., & Manyika, J. (2019). *Notes from the AI frontier: Tackling bias in artificial intelligence (and in humans)*. McKinsey Global Institute.

Simon, H. A. (1971). Designing Organizations for an Information-rich World. *Computers, communications, and the public interest*. Baltimore, MD: The Johns Hopkins Press.

Sipior, J. C. (2020). Considerations for development and use of AI in response to COVID-19. *International Journal of Information Management, 55*, 102170. https://doi.org/10.1016/J.IJINFOMGT.2020.102170.

Sprague, R. H., & Watson, H. J. (1993). *Decision support systems: Putting theory into practice*. Prentice Hall.

Stacey, P. (2019). *What is it that humanity should not give away to machines?* Bath Royal Literary and Scientific Institution. https://www.brlsi.org/events-proceedings/proceedings/what-it-humanity-should-not-give-away-machines.

Stacey, P., & Nandhakumar, J. (2009). A temporal perspective of the computer game development process. *Information Systems Journal, 19*(5), 479–497. https://doi.org/10.1111/j.1365-2575.2007.00273.x.

Stacey, P., & Tether, B. (2015). Designing emotion-centred Product Service Systems: The case of a cancer care facility. *Design Studies, 40*, 85–118. https://doi.org/10.1016/J.DESTUD.2015.06.001.

Starbird, K. (2019). Disinformation's spread: Bots, trolls and all of us. *Nature, 571*(7766), 449. https://doi.org/10.1038/d41586-019-02235-x.

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The Lancet*. https://doi.org/10.1016/S0140-6736(13)61489-0.

Topol, E. (2019). *The topol review: Preparing the healthcare workforce to deliver the digital future*. https://topol.hee.nhs.uk/.

Torras, C. (2015). Social robots: A meeting point between science and fiction. *MÈTODE Science Studies Journal, 5*, 111–115. https://doi.org/10.7203/metode.82.3546.

Vandemeulebroucke, T., Dierckx de Casterlé, B., & Gastmans, C. (2018). The use of care robots in aged care: A systematic review of argument-based ethics literature. *Archives of Gerontology and Geriatrics, 74*(September), 15–25. https://doi.org/10.1016/j.archger.2017.08.014.

Walsh, K., & Gordon, J. (2008). Creating an individual work identity. *Human Resource Management, 18*(1), 46–61. https://doi.org/10.1016/j.hrmr.2007.09.001.Final.

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. The MIT Press. https://doi.org/10.2307/j.ctt9qf9qj.

Weng, L., Flammini, A., Vespignani, A., & Menczer, F. (2012). Competition among memes in a world with limited attention. *Scientific Reports, 2*(1), 335. https://doi.org/10.1038/srep00335.

Whittaker, S., & Sidner, C. (1997). Email overload: Exploring personal information management of email. In S. Kiesler (Ed.), *Culture of the internet* (pp. 277–295). Lawrence Erlbaum.

World Café Community Foundation. (2015). *A quick reference guide for hosting world café*. http://www.theworldcafe.com/wp-content/uploads/2015/07/Cafe-To-Go-Revised.pdf.

Wu, Y. H., Fassert, C., & Rigaud, A. S. (2012). Designing robots for the elderly: Appearance issue and beyond. *Archives of Gerontology and Geriatrics, 54*(1), 121–126. https://doi.org/10.1016/j.archger.2011.02.003.

Yapo, A., & Weiss, J. (2018). Ethical implications of bias in machine learning. In *Proceedings of the 51st Hawaii International Conference on System Sciences, 9* pp. 5365–5372). https://doi.org/10.24251/hicss.2018.668.

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75–89. https://doi.org/10.1057/jit.2015.5.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for the future at the new frontier of power*. Profile Books.