Editorial to the Special Issue

Responsible and Human-Centered Artificial Intelligence

Standards, Processes and Behaviors

Stefan Güldenberg, Uta Wilkens, Tom Stoneham

With this special issue on Responsible and Human-Centered Artificial Intelligence (AI), we celebrate the start of the 80th anniversary volume of the *Swiss Journal of Business (*Established 1947 as *Die Unternehmung)*. For 80 years the *Swiss Journal of Business* has aimed to disseminate new findings in business research, to draw attention to important problems in society, research and business practice, to set research agendas, to present scientifically sound solutions and generally to promote the exchange between science and practice. With over 350 issues and more than 2,000 published research articles, we are proud that our transition to a new name and full open access is bearing fruit and that we have now reached an audience of close to 150,000 international online readers yearly. We are very much looking forward to further expanding our reach and impact by focusing on grand challenges and interdisciplinary topics that broaden the horizon of business research and practice. The topic of this first special issue of our 80th volume is an excellent example of this.

Many in media, business and politics talk about responsible and human-centered AI. But what exactly does it mean, and why has so little of it been implemented despite the pressing urgency, given the technological speed of AI development and the influence it has on human beings and their workplaces? This special issue is dedicated to the grand challenge of AI. It aims to stimulate the debate on responsible AI in corporate business ethics from an institutional and configurational perspective. The focus is on developing and implementing regulations, declarations and standards to ensure responsible AI design and deployment, and on integrating AI into business operations ethically. Emphasis is also placed on the human elements of AI ethics, including individual decision-making, reflective practices, AI literacy and leadership behaviors.

The application of AI in the workplace and the reflection on ethics related to AI-generated content, solutions and consequences for users and customers, is a matter for all

stakeholders – ethics is always an individual responsibility as well as a corporate one. Core research questions reflect on individual practices in AI-augmented work systems, responsible and reflective user behavior, requirements for maintaining ethical choices, applications, and behavior. This also includes leadership roles in ensuring transparency, fairness and equality in AI-assisted domains or algorithmic management.

The EU AI Act as well as the EU General Data Protection Regulation (GDPR) helps to systemize and understand risks and hidden ethical threats of corporate AI applications. This includes addressing the risk levels outlined in the EU AI Act specifying their criteria for responsible AI in corporate declarations and aligning these frameworks with corporate values and existing reporting systems such as ESG. The core research questions are related to the tension and harmonization between different standards, their impact on mitigating risks, and achieving sustainable development goals (SDGs). Another core area of research involves the visibility, acceptance and bargaining of declarations in the face of power differences.

From the perspective of corporate management and business ethics the responsible use of AI is more than just appealing, reflecting user behavior and non-negotiable fundamental values - it is a corporate responsibility for strategy, competitiveness and systems' resilience. For generating sustainable solutions, enhancing process resilience, and fostering organizational learning while using AI systems, corporate management must create structures, institutionalize governance mechanisms, and foster AI literacy for a responsible deployment of AI – not only on a technical basis for compliance regulation but as a social and cultural value. So far, little is known about organizational strategies, practices and outcomes with respect to responsible AI. Given the potential dislocation of accountability, core research questions address the location of responsibility, characteristics of responsibility-enhancing processes, the effectiveness of norms and practices, as well as consequences for SDGs.

In response to this challenging context, this special issue consists of seven articles, three full-length research articles and four shorter perspective contributions, all of which describe how companies reflect on AI ethics in their corporate strategy. All articles in this special issue share a common aim: to reflect and discuss the relevance and far-reaching implications of responsible and human-centered AI in practice.

The research articles provide conceptual frameworks, empirical evidence and case studies insights on how responsible and human-centered AI can be seen as a process and outcome, through the lens of ordonomics and against a socio-technical framework. They explain in great detail why responsible and human-centered AI should not only be seen as a technological challenge but a cultural and social one that requires smart organizational design and governance, how organizations construct collective interpretations of responsible AI along the dimensions of legitimacy, suitability, necessity, and proportionality, how corporations continuously adapt their practices while seeking for legitimization within a field of collective interpretation of responsible AI, and that even among companies with high obligations in public transparency only a minority systematically applies ethical AI standards on a high level where it is not legally required.

The perspective articles pinpoint selected discourses on AI ethics, broaden the view for future research, and address new topics for business ethics. In doing so, the contributing authors succeed in highlighting the necessity of responsible and human-centered AI as well as its practical implementation from various disciplinary and sector perspectives regarding

standards, processes and behaviors. It quickly becomes evident that the implementation of responsible and human-centered AI has implications for multiple disciplines, each characterized by distinct research traditions and practical relevance, necessitating their adequate representation.

The first research contribution "Responsible Use of Artificial Intelligence as Continuous Proportionalization" by *Simon Sturm, Florian Krause and Benjamin van Giffen* presents findings from original case study research in the fashion industry. The authors conceptualize responsible AI as both a process and outcome of social evaluation and propose a model ("continuous proportionalization") that explains how organizations construct collective interpretations of responsible AI along the dimensions of legitimacy, suitability, necessity, and proportionality. They argue that there is no established understanding of the responsible use of AI; instead corporations continuously adapt their practices while seeking for legitimization within a field of collective interpretation of responsible AI. The conceptual framework is illustrated with the case study of AI-based fashion image generation at OTTO, Germany's largest e-commerce company.

The second research contribution "Artificial Intelligence as a Socio-Economic Dilemma: Ordonomic Diagnosis–Reflection–Design for Education, Work and Governance" by *Patrick Hedfeld* analyzes AI through the lens of ordonomics, a normative-institutional approach that connects economic rationality with ethical reflection. While most discussions in AI ethics focus on principles such as fairness, transparency, and accountability, fewer studies address how these principles can be institutionalized through incentive-compatible rules. This research paper therefore conceptualizes AI not as a primarily technological challenge but as a social order problem that requires institutional design and governance. *Hedfeld* explicitly maps the classical ordonomic three-level schema - actor, institutional order, and market/discourse - onto an applied heuristic of Diagnosis–Reflection–Design, demonstrating how this triad operationalizes ordonomic reasoning for the AI context. Building on this foundation, the paper identifies and categorizes key AI-related social dilemmas (economic, epistemic, ethical, and educational). The analysis develops differentiated responsibilities across levels of coordination and proposes rule-based cooperation solutions that align individual incentives with collective welfare. By linking ordonomics to current frameworks such as Responsible AI, algorithmic accountability, and the EU AI Act, the paper positions ordonomics as a design-oriented ethics that bridges normative ideals and institutional economics. The result of the paper is a framework for diagnosing conflicts, reflecting responsibilities, and designing cooperative solutions that reconcile innovation with social responsibility.

In the third research contribution "Towards AI Governance in DAX40: A Typology of Organizational Guidelines for Self-Regulation" *Niklas Obermann, Daniel Lupp and Uta Wilkens* provide empirical evidence to organizational practices in coping with challenges of ethical AI. The authors systematically evaluate the self-regulating guidelines for AI ethics of the German DAX40 corporations against a socio-technical framework specifying different ways of how to demonstrate responsibility, whether it is directed towards the trustworthiness of the technology or may also include issues of organizational or personnel development. The outcome of their qualitative content analysis is "A typology of organizational guidelines for self-regulation", distinguishing DAX40 corporations with (1) non-codified self-regulation, (2) symbolic-technical self-regulation, and (3) comprehensive socio-technical self-regulation. Only the latter mirrors a broader discourse on AI ethics

including social dimensions of ethics in addition to technological characteristics of AI. The differences indicate that even among companies with high obligations in public transparency only a minority systematically applies ethical standards on a high level where it is not legally binding.

In the first of the four perspective contributions "Trust and Responsibility in AI: An Interdisciplinary Social-Sector Perspective" by *Nathan Chappell* the author states that the rapid adoption of artificial intelligence has intensified debates about responsibility, ethics, and trust. While regulatory frameworks and organizational ethics statements are proliferating, responsible AI is too often treated as compliance or reputation management rather than an organizing principle of practice. His perspective paper argues that social-sector organizations - including nonprofits, NGOs, and other mission-driven institutions - offer an instructive lens for rethinking responsible AI because they operate with structural vulnerability and high trust dependence. By drawing on business ethics, organizational theory, nonprofit and social-sector management, and human flourishing scholarship, *Chappell* proposes shifting from harm-avoidance toward trust-centered, flourishing-oriented AI integration.

The second perspective paper "Responsible und Human-Centered "AI" - Some Ethical Considerations" by *Peter G. Kirchschläger* provides an ethical as well as practical guidance towards responsible and human-centered AI. *Kirchschläger* challenges the term AI by referring to it as "data-based systems (DS)" and sees the necessity to identify ethical opportunities and risks of DS in order to promote the former and in order to avoid the latter – for the benefit of all people and the planet earth. He states in his perspective paper that companies can contribute to the realization of DS with ethics by, first, living up to the exclusive human responsibility for machines; second, while running innovation- and research-processes, by implementing always right from the start an interaction between ethics and technologies; third, by promoting global human rights-based regulation of DS as well as the establishment of an International Data-Based Systems Agency (IDA) at the UN enforcing this global regulation of DS.

The third perspective paper "Defining Knowledge in the Age of Society 5.0" by *Susanne Durst* takes a knowledge management perspective by introducing the concept of responsible knowledge management (rKM) and its usefulness for implementing ethically accepted AI solutions in organizations. Illustrative examples are presented to demonstrate the latter. *Durst* states in her paper that integrating the underlying principles of rKM into discussions related to responsible and human-centred AI in business ethics is expected to lead to the development and execution of more inclusive and responsible solutions to addressing the grand challenges at hand. She argues that this way of thinking can also contribute to achieving the United Nations Sustainable Development Goals, in particular Goal 5 "Gender Equality," Goal 10 "Reduced Inequalities," and finally Goal 17 "Partnerships", and that the use of AI raises the question of digital inequality, which is why a human-centred, inclusive, and collaborative approach such as rKM is more important than ever. The paper concludes with a series of research questions that serve as an outlook and inspiration for further reflection on rKM and ethically acceptable AI solutions in companies.

The fourth and final perspective paper "Intimate Machines, Disturbed Minds: Managing the Affective Cost of AI" by *Leona Chandra Kruse* and *Patrick Mikalef* focuses on the affective costs of AI on an individual as well as team level in organizations. The authors suggest preventive as well as corrective regulation and governance principles that

seek to limit problematic affective dynamics upstream while at the same time respond to relational and emotional disruptions once AI becomes embedded in work practices. At the end of their perspective paper *Chandra Kruse/Mikalef* propose a further research agenda for better managing the affective costs of AI in organizations.

This special issue provides a broad overview of the grand challenges, opportunities and ethical risks of human-centered AI in organizations and society at large. It offers both theoretically grounded and practice-oriented approaches and examples of how ethical governance standards, processes, behaviors, and regulations can be established. It also discusses potential limitations, shows existing contradictions, highlights research gaps as well as future research questions of responsible and human-centered AI. A central theme that emerges across all contributions is the critical importance of going beyond direct technology-related criteria of ethics and taking relational and social aspects into consideration, even if they are indirectly addressed, thereby exemplifying Kranzberg's First Law: "Technology is neither good nor bad; nor is it neutral" (1986, 545). AI ethics is conceptualized as an issue incorporated in all spheres of business, from governance and process design to people management and individual responsibility. The rapid adoption of AI technologies is rushing ahead of both hard and soft regulation, and of cultural and social norms which might guide business leaders, creating both risks and opportunities which need to be bounded by explicit engagement with ethics.

Given the current dynamic and disruptive technological developments we are pretty sure that this special issue is more the beginning than the end of an ethical discourse on responsible and human-centered AI in organizations and society. May it inspire future research and practical actions on ethical governance standards, processes, behaviors, and regulations regarding a more responsible and human-centered use of AI. We would like to thank all the authors involved in this special issue for their insightful contributions. We are especially grateful to our dedicated reviewers, who have made a significant contribution to ensuring the quality of this special issue.

References

Kranzberg, M. (1986). Technology and History: 'Kranzberg's Laws.'*Technology and Culture* 27(3): 544–560. https://doi.org/10.2307/3105385.

Stefan Güldenberg, Prof. Dr., is Managing Editor of the Swiss Journal of Business, President of the Swiss Society for Business and Management and Full Professor as well as Academic Director at the Graduate School of the EHL Hospitality Business School, Lausanne.

Address: EHL Hospitality Business School, HES-SO, University of Applied Sciences and Arts Western Switzerland, Route de Berne 301, 1000 Lausanne, 25, Switzerland,
E-Mail: stefan.guldenberg@ehl.ch, sjb@nomos-journals.de
ORCID: https://orcid.org/0000-0002-8698-4113

Uta Wilkens, Prof. Dr., is Full Professor of Work, Human Resources & Leadership at Ruhr University Bochum, Germany, member of acatech "Learning Systems", Chairperson of the Competence Center HUMAINE and the HUMAINE Network e.V., an Association for Human-Centered AI.

Address: Ruhr University Bochum, Institute of Work Science, Chair of Work, Human Resource & Leadership, Universitätsstr. 150, 44801 Bochum, Germany,
E-Mail: uta.wilkens@rub.de
ORCID: https://orcid.org/0000-0002-7485-4186

Tom Stoneham, Prof. Dr., is Full Professor of Philosophy at the University of York, UK. He is Ethics Lead for the UKRI Centre for Doctoral Training in Safe AI Systems and Convenor of the MA in Applied Ethics and Governance of Data Privacy, and President of the International Berkeley Society.

Address: Department of Philosophy, University of York, Heslington, York, YO10 5DD, UK, E-Mail: tom.stoneham@york.ac.uk
ORCID: https://orcid.org/0000-0001-5490-4927