

10-9-2023

Instrumentalising Dignity to Govern Digital Ethics: Towards a Unified View

Nick Lüthi

Universität Bern, Switzerland, nicola.luethi@unibe.ch

Christian Matt

Universität Bern, Switzerland, christian.matt@unibe.ch

Follow this and additional works at: <https://aisel.aisnet.org/wi2023>

Recommended Citation

Lüthi, Nick and Matt, Christian, "Instrumentalising Dignity to Govern Digital Ethics: Towards a Unified View" (2023). *Wirtschaftsinformatik 2023 Proceedings*. 58.

<https://aisel.aisnet.org/wi2023/58>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISeL). It has been accepted for inclusion in Wirtschaftsinformatik 2023 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Instrumentalising Dignity to Govern Digital Ethics: Towards a Unified View

Research in Progress

Nick Lüthi¹ and Christian Matt¹

¹ Universität Bern, Institut für Wirtschaftsinformatik, Information Management, Bern, Schweiz
{nicola.luethi, christian.matt}@unibe.ch

Abstract. The potential ethical issues of their digital products and services for consumers represent a challenge for many companies. Above obeying the laws, stakeholders expect companies to consider any ethical ramifications of their products and services. Companies have responded by establishing different digital ethics governance approaches, such as ethics boards, new specialist roles, or extended responsibilities for compliance officers. However, criticism holds that all approaches may only be “ethics washing”, lacking the effectiveness to influence product and service design effectively. An informed response requires a unified view to judge the effectiveness of digital ethics governance approaches on digital product and service design. We hold that dignity is the missing link to connect ethical governance approaches with their effects on the operational level. We, therefore, provide the CAGE Framework (Claims, Affronts, Governance, Effects) for dignity, which details the organisational lifecycle for effective digital ethics governance.

Keywords: dignity, digital ethics governance, CARE Theory, CAGE Framework

1 Introduction

While digital services and products (DSP) have profited tremendously from AI and other technological advancements, these advancements also pose severe ethical issues for consumers, i.e. situations in which the behaviour of organisations adversely affects consumers (Mirbabaie et al. 2022, Seppälä et al. 2021, Mittelstadt et al. 2016). For instance, companies use consumer data to push sales in moments when the latter feel vulnerable (Stahl et al. 2023), DNA tests, being taken by a small percentage of consumers, allow providers to make far-reaching deductions about the whole population (Marabelli et al. 2021), or online loan decisions being declined to applicants with good standing, but without further elaboration (Marjanovic et al. 2021). These examples show that the current regulations and laws may not sufficiently prevent organisations

from developing DSP that entail such ethical harm. However, not even explicitly formulated organisational principles suffice to avoid them (Mittelstadt 2019) since organisational processes often fail to prevent ethical issues early when companies develop new DSP (Nussbaumer et al. 2019).

Organisations have started to respond to these issues with different digital ethics governance approaches (Stahl et al. 2022). Some have created boards for digital ethics that review DSP before they are introduced to markets. Others have extended the duties of compliance officers to regulate potential ethical issues of DSP supported by new standards (IEEE 2021, Spiekermann 2021). Last, some companies have created the role of Digital Ethicists, whose task is to oversee the company's DSP concerning their ethical issues for consumers. These different approaches can partly be explained by organisations choosing distinct forms of IT governance (Tiwana et al. 2013), depending on their specific need for digital ethics and the existing organisational structure.

While these approaches may be laudable, their effects remain unclear. The literature has highlighted the need for digital ethics governance approaches (Mayer et al. 2021) as well as appropriate mitigation strategies (Stahl et al. 2022), and it has confirmed that ethical principles are hard to translate into organisational practice (Floridi 2021, Morley et al. 2020, Haimerl et al. 2022, Seppälä et al. 2021). But assessing this connection is also tricky for IS research. Recently, Leidner & Tona (2021, p. 361) have conceptualised dignity as a principle that “sets the scene for future IS research to investigate digital technologies from a societal grand challenge perspective”. With their seminal work, the authors have provided CARE, an essential theory for bridging the gap from digital ethics governance approaches to effects on DSP. CARE allows us to study the effectiveness of digital ethics governance approaches. The theory proposes that organisations seek to lessen dignity-induced tensions. Thus, the more effective a governance approach is, the smaller will resulting tensions be. By reducing such tensions, organisations effectively reduce the ethical issues of their DSP, as treating all relevant consumers with dignity enables more ethical DSP for them. However, due to the grand theory nature of CARE, we still lack a concrete framework to obtain a unified view that can help explain how digital ethics governance approaches specifically affect DSP for consumers. Such a unified view should detail how and to what effect organisations use digital ethics governance approaches. Considering this current state, the following question guides our study:

— RQ: *How do digital ethics governance approaches affect digital products and services for consumers?*

Adopting the grand theory CARE by Leidner & Tona (2021) to digital ethics governance, we provide the CAGE framework (Claims, Affronts, Governance Approach, Effect). Dignity enables us to understand the interplay of ethics governance approaches and existing tensions on DSP. Tensions serve as a measure to understand the ethical level of DSP. Our CAGE framework provides an internal, iterative lifecycle model for firms and serves as a basis for future research on digital ethics governance. Our final study will expand the CAGE framework to a mid-level theory for which we use a qualitative research approach. We currently conduct interviews with stakeholders from midsize and large companies that use different digital ethics governance approaches.

2 Conceptual Background

2.1 Governance of Digital Ethics

Researchers identified ethical themes in many IS works (Kern et al. 2022), and Stahl (2012) provided early guidance to summarise the existing body of work into four levels of normativity (meta-ethics, ethical theory, explicit morality and moral intuition). Research on the governance of digital ethics has highlighted the difficulties organisations face when incorporating digital ethics (Mittelstadt 2019) and the possible levels of engagement when doing so (Stahl et al. 2022). Research has conceptually highlighted potential dangers such as ethics washing when translating principles into organisational practice (Floridi 2021, Morley et al. 2020), as digital ethics governance approaches are susceptible to the same drivers driving greenwashing (Delmas & Burbano 2011). Empirical work on this level has shown unintended ethical issues when introducing AI (Strich et al. 2021), the effects of principles on organisational identity (Haimerl et al. 2022), or ethics guidelines as a first step for further organisational governance of digital ethics (Mayer et al. 2021, Becker et al. 2022). As an emerging field, there is currently no unifying understanding of the *how*, i.e., the effects of digital governance approaches on DSP. Therefore, IS research and companies are left in the dark about the concrete impacts of choosing a particular ethical governance approach and how the different approaches compare in their effects.

2.2 Dignity in IS

Dignity is, first and foremost, a philosophical concept (Hill 1992), and it has seen minimal adoption in the social sciences. Kantian ethics treats dignity as its enabling principle; a person's inherent dignity is crucial to formulating the ethical theory (Hill 1992). In this sense, dignity serves as the enabling principle for ethics. In other words: if everyone is treated with dignity, everyone is treated morally just. Other ethical theories use different principles, but dignity seems to have widespread adoption and is a suitable means to reach a consensus among differing ethical ideologies (Stahl et al., 2023). New movements like humanistic management also treat dignity as their primary principle (Kostera & Pirson 2017). Leidner & Tona (2021, p. 365) provide one of the first adaptations of the concept for the IS discipline as a "novel way to understand the use and consequences of information systems on human life". They identified three meanings of dignity across a wide variety of literature: *behavioural dignity*, which is the notion of living a virtuous life; *meritocratic dignity*, as a certain status accompanied by a particular treatment per this status; *inherent dignity*, as having dignity by virtue of possessing rationality. For the scope of this paper, we will treat dignity as an umbrella term for those three meanings. The three meanings of dignity all make specific *claims* about the person possessing it, i.e. "treatment that enables behavioral dignity and/or treatment that recognizes the inherent or meritocratic dignity of the person" (Leidner & Tona 2021, p. 347) and are open to *affronts* to that meaning of dignity, i.e. "treatment that humiliates, insults, or injures the behavioral, inherent, or meritocratic dignity of the person" (Leidner & Tona 2021, p. 347). Furthermore,

different meanings of dignity can come into tension leading to dignity disequilibriums, i.e. the tensions between different meanings of dignity created by claims and affronts. Leidner & Tona (2021) argue that organisations respond on the macro- and micro-level to disequilibriums by striving for a dignity equilibrium which resolves these tensions. These considerations and arguments, extended with three theoretical assertions, form the CARE theory (Claims, Affronts, Responses, Equilibrium), a grand theory of dignity for the IS discipline. Drawing from CARE, our goal is to operationalise it for assessing the effects of ethical governance approaches on operational outcomes at the service and product level and to develop it further into a mid-level theory using qualitative empirical evidence.

3 The CAGE Framework

Due to CARE's nature as a grand theory, it highlights organisational and individual responses to dignity-induced tensions. Still, it neither provides a connection between governance approaches nor accounts for the effects of governance approaches on DSP. As such, it also misses a specification for more narrow contexts, such as consumer-facing DSP. We thus take CARE's main ground and derive an operationalised process for our context: CARE is based on the tacit assumption that DSP are not neutral either in a conceptualised or a realised state (Leidner & Tona 2021, Martin 2019). This non-neutrality expresses itself as claims and affronts to stakeholder dignity, which create tensions between meanings of dignity. To resolve these tensions, organisations seek to establish means such as digital ethics governance approaches to lessen these tensions. Ideally, these means will create an equilibrium between dignity claims and affronts.

The CAGE framework forms an internal organisational lifecycle that applies to companies that develop DSP targeted at consumers (B2C). CAGE can handle complex realities such as companies developing large DSP portfolios or multiple tensions arising from one product/service. Due to targeting consumers, though, CAGE does not make assertions about internal stakeholder tensions. While ethical issues can certainly also occur during implementing of DSP within companies (e.g., in performance measurement), we focus on developing commercial offerings for consumers, as this is where we see the most significant potential for ethical issues.

CAGE proposes that effective digital ethics governance reduces dignity-related tensions of the targeted DSP for consumers. The CAGE framework is iterative and consists of an internal organisational lifecycle with five main steps.

1. The first conceptualisation of DSP makes claims about the dignity of consumers but also affronts the dignity of consumers.
2. These claims and affronts collide and induce tensions between different meanings of dignity (between stakeholder groups and/or within the same group).
3. The tensions necessitate organisational responses of the organisation developing DSP. To reduce the tensions, the organisation will either establish or, if already established, apply a digital ethics governance approach.
4. The governance approach affects the next iteration of DSP. It leads to ethically more sound DSP, supposing that the chosen governance approach effectively releases tensions.

- The new iteration of DSP will make claims about the dignity of consumers but also affront the dignity of stakeholders. Ideally, this iteration will lead to less/minimised tensions. The new DSP of the developing organisation will be affected by the established digital ethics governance approach.

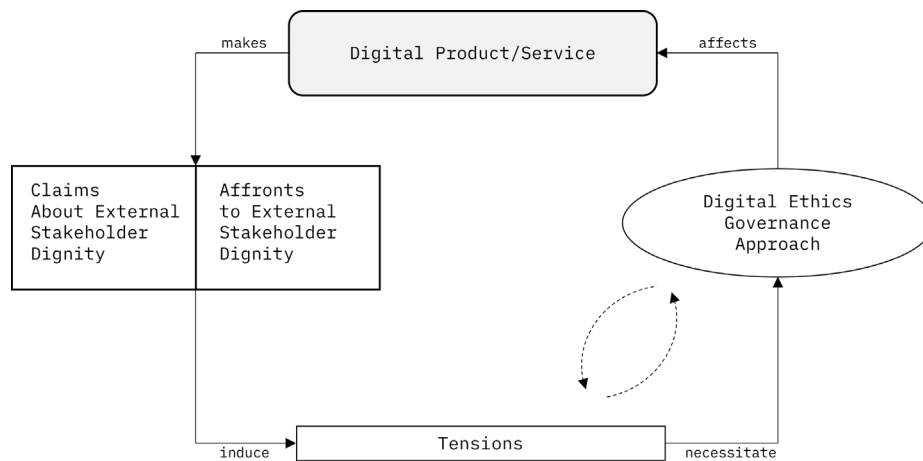


Figure 1. CAGE Framework

Appropriate digital ethics governance approaches should always reduce the number of existing tensions and minimise any future tensions. As such, CAGE is agnostic to the ethical theories/ principles used and thus able to be implemented for any form of digital ethics governance approach. To understand the effectiveness of a digital governance approach, CAGE compares dignity-induced tensions between iterations of DSP and assumes that fewer/smaller is better. To illustrate this: When shopping apps make recommendations to consumers, they create the dignity affront that consumers' purchasing decisions are better if they follow the app's suggestions. Such an affront targets the meritocratic dignity of consumers: because of their status (consumer), they warrant such and such treatment. This affront is in tension with the consumer's inherent dignity claim to wanting to make purchasing decisions autonomously. By targeting two different meanings of dignity (inherent vs meritocratic), this claim-affront pair creates a dignity tension experienced by the consumer. Effective digital ethics governance approaches would reduce such tensions by recognising the consumer's dignity claim or lessening the app's affront. For example, a company employing a Digital Ethics Specialist would enrol the specialist into the project team. After assessing the situation and recognising the consumer's claim, the specialist would suggest giving more autonomy to consumers by providing them with more information, such as a pop-up explaining why the app has recommended a particular item. The project team would then create a next iteration that provides consumers with additional explanations in an appropriate form for the app and thus reduce the tension between the app's treatment of users (affront) and their desire for autonomy (claim).

4 Method and Further Research

4.1 Method

Our further research will enrich this conceptual framework with a qualitative approach to provide empirical evidence to the research question. We collect data in the Netherlands, Germany and Switzerland. Interviews are conducted in German and English and based on a semi-structured guide consisting of two main blocks: processes/governance and dignity. In the first block, for example, we asked interviewees to describe their organisation's engagement with digital ethics, explain the processes/process steps explicitly catered to digital ethics or who was responsible and accountable for implementing those processes. Questions in the second block then focused on stakeholder treatment (internal & external), conflicts between stakeholders and how those are usually resolved. To attain rich insights, we chose a three-step sampling approach: We first interviewed consultants from Germany and Switzerland. After this first step, we reworked our interview guide and interviewed employees responsible for digital ethics governance (e.g. Digital Ethics Specialist/Ambassador/Leader/Director, Head of AI, Head of Compliance). The organisations of these experts range across industries (e.g. retail, technology, telecommunication, healthcare) and size. A first analysis of these interviews revealed three distinct digital ethics governance forms. Our third sampling step thus involved practitioners from three organisations (Data Scientists, Product Owners) responsible for designing and developing DSP, each representing one of the identified governance forms. In addition to the interviews, we collected material from most organisations, such as Codes of Conduct, Digital Ethics Guidelines, and professional ethics, to further our understanding of an organisation's governance for Digital Ethics. The principal researcher also participated in community events at which companies demonstrated their use cases and issues they face when implementing Digital Ethics Governance. Following Saldaña (2013), we use versus and values coding as first-cycle methods for analysis. These coding forms were chosen to understand better what conflicts arise and how these conflicts shape the responses of employees to digital ethics governance which in turn affect the DSP they are working on. Finally, we will use focused coding as second-cycle method.

4.2 Concluding Note

This research in progress starts with a conceptual approach to respond to the research question: Effective digital ethics governance approaches reduce the dignity-induced tensions for consumers and affect future iterations of DSP. The conceptualisation of dignity from the CARE theory of dignity (Leidner & Tona 2021) serves as a basis to assess the effects of a digital ethics governance approach on the design of DSP by investigating dignity-induced tensions. Our adaptive framework CAGE (Claims, Affronts, Governance, Effect) shows that to resolve dignity-induced tensions; organisations must choose appropriate governance approaches for the claims and affronts to dignity their DSP make for consumers in their current state.

References

- Becker, S. J., Nemat, A. T., Lucas, S., Heinitz, R. M., Klevesath, M. & Charton, J. E. (2022), 'A Code of Digital Ethics: Laying the foundation for digital ethics in a science and technology company', *AI & Society*. <https://doi.org/10.1007/s00146-021-01376-w>
- Delmas, M. A. & Burbano, V. C. (2011), 'The Drivers of Greenwashing', *California Management Review* **54**(1), 64–87.
- Floridi, L. (2021), Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical, in L. Floridi, ed., 'Ethics, Governance, and Policies in Artificial Intelligence', Philosophical Studies Series, Springer International Publishing, Cham, pp. 81–90.
- Haimerl, A., Mayer, A.-S. & Fiedler, M. (2022), 'Managing Organisational Identity Challenges Caused By AI Implementation: The Role Of AI Principles' in *ICIS 2022 Proceedings* 10.
- Hill, T. E. (1992), *Dignity and Practical Reason in Kant's Moral Theory*, Cornell University Press.
- IEEE (2021), 'IEEE Standard Model Process for Addressing Ethical Concerns during System Design', *IEEE Std 7000-2021* pp. 1–82.
- Kern, C. J., Noeltner, M. & Kroenung, J. (2022), 'The Case of Digital Ethics in IS Research – A Literature Review' in *ICIS 2022 Proceedings* 10.
- Kostera, M. & Pirson, M. (2017), *Dignity and the Organization*, Springer.
- Leidner, D. E. & Tona, O. (2021), 'The Care Theory of Dignity Amid Personal Data Digitalization', *MIS Quarterly* **45**(1), 343–370.
- Marabelli, M., Newell, S. & Handunge, V. (2021), 'The lifecycle of algorithmic decisionmaking systems: Organisational choices and ethical challenges', *The Journal of Strategic Information Systems* **30**(3), 101683.
- Marjanovic, O., Cecez-Kecmanovic, D. & Vidgen, R. (2021), 'Algorithmic pollution: Making the invisible visible', *Journal of Information Technology* **36**(4), 391–408.
- Martin, K. (2019), 'Ethical Implications and Accountability of Algorithms', *Journal of Business Ethics* **160**(4), 835–850.
- Mayer, A.-S., Haimerl, A., Strich, F. & Fiedler, M. (2021), 'How corporations encourage the implementation of AI ethics' in *ECIS 2021 Research Papers* 27.
- Mirbabaie, M., Brendel, A., & Hofeditz, L. (2022), 'Ethics and AI in Information Systems Research', *Communications of the Association for Information Systems* **50**(1), 726–753.
- Mittelstadt, B. (2019), 'Principles alone cannot guarantee ethical AI', *Nature Machine Intelligence* **1**(11), 501–507.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016), 'The ethics of algorithms: Mapping the debate', *Big Data & Society* **3**(2), 1–21.
- Morley, J., Floridi, L., Kinsey, L. & Elhalal, A. (2020), 'From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices', *Science and Engineering Ethics* **26**(4), 2141–2168.
- Nussbaumer, A., Pope, A. & Neville, K. (2019), 'A framework for applying ethics-by design to decision support systems for emergency management', *Information Systems Journal* **33**(1), 34–55.
- Saldaña, J. (2013), *The coding manual for qualitative researchers*. SAGE.
- Seppälä, A., Birkstedt, T. & Mäntymäki, M. (2021), 'From Ethical AI Principles to Governed AI' in *ICIS 2021 Proceedings* 10.
- Spiekermann, S. (2021), From value-lists to value-based engineering with IEEE 7000™, in '2021 IEEE International Symposium on Technology and Society (ISTAS)', pp. 1–6.
- Stahl, B. C. (2012), 'Morality, Ethics, and Reflection: A Categorisation of Normative IS Research', *Journal of the Association for Information Systems* **13**(8), 636–656.

- Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K. & Jiya, T. (2022), 'Organisational responses to the ethical issues of artificial intelligence', *AI & SOCIETY* **37**(1), 23–37.
- Stahl, B. C., Schroeder, D. & Rodrigues, R. (2023), *Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges*, Springer Nature.
- Strich, F., Mayer, A.-S. & Fiedler, M. (2021), 'What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity', *Journal of the Association for Information Systems* **22**(2), 304–324.
- Tiwana, A., Konsynski, B. & Venkatraman, N. (2013), 'Special Issue: Information Technology and Organisational Governance: The IT Governance Cube', *Journal of Management Information Systems* **30**(3), 7–12.