AIS Transactions on Human-Computer Interaction

Volume 12
Issue 4 *Special Issue on AI Fairness, Trust, and Ethics*

Article 3

12-31-2020

# On Implementing Ethical Principles in Design Science Research

Ivo Benke
*Karlsruhe Institute of Technology*, ivo.benke@kit.edu

Jasper Feine
*Karlsruhe Institute of Technology*, jasper.feine@kit.edu

John R. Venable
*Curtin University*, j.venable@curtin.edu.au

Alexander Maedche
*Karlsruhe Institute of Technology*, alexander.maedche@kit.edu



## Recommended Citation

Benke, I., Feine, J., Venable, J. R., & Maedche, A. (2020). On Implementing Ethical Principles in Design Science Research. *AIS Transactions on Human-Computer Interaction, 12*(4), 206-227. https://doi.org/10.17705/1thci.00136

DOI: 10.17705/1thci.00136

Transactions on Human-Computer Interaction

Volume 12 Issue 4

12-2020

# On Implementing Ethical Principles in Design Science Research

Ivo Benke

*Institute of Information Systems and Marketing, Karlsruhe Institute of Technology, Karlsruhe, Germany, ivo.benke@kit.edu*

Jasper Feine

*Institute of Information Systems and Marketing, Karlsruhe Institute of Technology, Karlsruhe, Germany, jasper.feine@kit.edu*

John R. Venable

*School of Management, Curtin University, Perth, Australia, j.venable@curtin.edu.au*

Alexander Maedche

*Institute of Information Systems and Marketing, Karlsruhe Institute of Technology, Karlsruhe, Germany, alexander.maedche@kit.edu*



**Recommended Citation**

Benke, I., Feine, J., Venable, J. R., & Maedche, A. (2020). On Implementing Ethical Principles in Design Science Research. *AIS Transactions on Human-Computer Interaction*, *12*(4), pp. 206-227.

DOI: 10.17705/1thci.00136

Available at http://aisel.aisnet.org/thci/vol12/iss4/3

Transactions on Human-Computer Interaction

Research Paper DOI: 10.17705/1thci.00136 ISSN: 1944-3900

# On Implementing Ethical Principles in Design Science Research

**Ivo Benke**
Institute of Information Systems and Marketing,
Karlsruhe Institute of Technology, Karlsruhe, Germany

**Jasper Feine**
Institute of Information Systems and Marketing,
Karlsruhe Institute of Technology, Karlsruhe, Germany

**John R. Venable**
School of Management,
Curtin University, Perth, Australia

**Alexander Maedche**
Institute of Information Systems and Marketing,
Karlsruhe Institute of Technology, Karlsruhe, Germany

**Abstract:**

Technological innovations raise axiological questions such as what is right or wrong, good and bad, and so on (i.e., ethical considerations). These considerations have particular importance in design science research (DSR) projects since the developed artifacts often actively intervene into human affairs and, thus, cannot be free from value. To account for this fact, Myers and Venable (2014) proposed six ethical principles for DSR in order to support researchers to conduct ethical DSR. However, ethical principles per se—and the ethical DSR principles that Myers and Venable proposed—have an abstract nature so that they can apply to a broad range of contexts. As a consequence, they do not necessarily apply to specific research projects, which means researchers need to contextualize them for each specific DSR project. Because doing so involves much challenge, we explore how contemporary DSR publications have dealt with this contextualization task and how they have implemented the six ethical principles for DSR. Our results reveal that DSR publications have not discussed ethical principles in sufficient depth. To further promote ethical considerations in DSR, we argue that both DSR researchers and reviewers should be supported in implementing ethical principles. Therefore, we outline two pathways toward ethical DSR. First, we propose that researchers need to articulate the next generation of ethical principles for DSR using prescriptive knowledge structures from DSR. Second, we propose extending established DSR conceptualizations with an ethical dimension and specifically introduce the concept of ethical DSR process models. With this work, we contribute to the IS literature by reviewing ethical principles and their implementation in DSR, identifying potential challenges hindering efforts to implement ethics in DSR, and providing two pathways towards ethical DSR.

**Keywords:** Ethics, Principles, Design Science Research, Prescriptive Knowledge Structures, Ethical DSR.

Nigel Melville was the accepting senior editor for this paper.

# 1 Introduction

Innovative technologies, particularly technology based on artificial intelligence (AI), offer great opportunities for individuals and society. However, they also create high risks since they can exacerbate existing power asymmetries and biases (EU, 2018; Robert, Bansal, & Lütge, 2020; Shneiderman, 2020). Accordingly, such implications increasingly raise axiological questions such as what is right or wrong (i.e., ethical considerations) (Hassan et al., 2018). Such considerations have particular importance in designing AI-based technology (EU, 2019) because harmfully misusing AI-based technology might have a severe negative influence on the environment and society (EU, 2020). Examples in which actors have misused AI-based technology include how Cambridge Analytica manipulated public elections and how governments monitor minorities (Hern, 2018). Besides cautiously handling innovations through institutions or lawful and societal restrictions that affect how one applies them (Mittelstadt, 2019), engaging in ethical considerations before, during, and after one designs innovative artifacts ensures that their technological benefits outweigh their risks (EU, 2019).

However, engaging in ethical considerations when designing innovative technology can be complex. For some artifacts, one cannot easily determine whether more efficient and effective solutions would cause harm to individuals or society. Moreover, what if researchers cannot agree on what is good or harmful (Myers & Venable, 2014)? What if one could abuse artifacts in unintended ways? What implications do artifacts have for society if "humans using IT artifacts becomes IT using human artifacts" (Demetis & Lee, 2018, p. 1)? Faced with such questions, researchers and practitioners nowadays "feel that a more robust ethical compass is needed to guide development and ensure accountability" of innovative, often AI-enabled artifacts (EU, 2018, p. 7).

Particularly in the information systems (IS) discipline, researchers need to consider ethical issues when conducting research projects following the design science research (DSR) paradigm. DSR projects typically result in artifacts that actively intervene in human affairs, such as through the influence of AI-based technology (Baskerville & Myers, 2015; Shneiderman, 2020). Therefore, they can have positive and negative implications on individuals and society (Fischer, 2017). To illustrate, consider AI-based conversational agents for customer service or enterprise collaboration. From a customer service perspective, one can design conversational agents with social cues to unconsciously influence customers' behavior (Feine, Gnewuch, Morana, & Maedche, 2019a) or to discriminate against individuals (Feine, Gnewuch, Morana, & Maedche, 2020b). In turn, social cues can increase user satisfaction (Gnewuch, Morana, Adam, & Maedche, 2018; Rietz, Benke, & Maedche, 2019) and help conversational agents to adapt to specific users, contexts, and tasks (Feine, Morana, & Maedche, 2019b). From an enterprise perspective, conversational agents might be a threat to rationalize human workers (Feine, Morana, & Maedche, 2020c, 2020d) but can also provide valuable support in employees' work routines (Benke, Knierim, & Maedche, 2020; Feine, Adam, Benke, Maedche, & Benlian, 2020a). Ergo, conversational agent designers always have to consider both positive and negative design implications and must engage in ethical considerations and design trade-offs (André et al., 2019; Benke, 2020). As a consequence, "information systems as design science cannot be value-free" (Iivari 2007, p. 56).

Contemporary DSR methodologies, however, do not explicitly provide ethical guidance. When analyzing the most prominent DSR process models, it becomes apparent that not one explicitly mentions the term ethics. To address this omission, Myers and Venable (2014) proposed six ethical principles that focus on designing innovative artifacts following the DSR paradigm. These ethical DSR principles—like many other ethical principles in research and practice—provide information about what ethical outcomes researchers should obtain (e.g., ensure privacy). However, they do not tell DSR researchers how to achieve them (e.g., how to explicitly ensure privacy in a specific DSR project) because, for one, high-level ethical principles need to apply to a broad range of different contexts and, thus, do not necessarily apply to a specific research project (Mittelstadt, 2019). In turn, individual researchers should ideally translate high-level principles to their specific projects. This contextual translation, however, represents a complex endeavor that complicates efforts to introduce ethical principles in DSR. Due to this complexity, we investigate the extent to which DSR publications have dealt with these challenges and how they have implemented ethical principles in their DSR projects. Specifically, we address the following research question (RQ):

**RQ:** To what extent do contemporary design science research publications implement ethical principles?

To answer our research question, we first step back in time and summarize ethical considerations in IS, design science, and public and private initiatives from the past. Subsequently, we assess how contemporary

DSR publications have implemented Myers and Venable's (2014) six ethical principles for DSR. To do so, we conduct a systematic literature review (SLR) to analyze DSR publications in the top journals and top conferences for DSR. Based on the ethical considerations in recent DSR literature that we discovered, we critically examine important challenges in implementing ethical principles in DSR. Subsequently, we propose that 1) a structure for articulating more applicable ethical principles should guide the next step towards ethical DSR and that 2) researchers should enrich existing DSR conceptualizations with an ethical dimension. Specifically, we propose an ethical DSR (E-DSR) process model.

# 2 Related Work

## 2.1 Ethical Considerations in Information Systems

We define ethics as the "philosophical discipline interested in questions of right and wrong, good and bad, do's or don't" (Moore, 1993, p. 110). However, the key difficulty remains how to decide what such terms (e.g., right, wrong, good, bad) mean. Attempts to answer these questions belong to the philosophical branch of axiology (Hassan, Mingers, & Stahl, 2018). Axiology covers ethical theories that one can classify into three dominant traditions: virtue ethics (i.e., strongly associated with Aristotle and focuses on an agent's character), duty ethics (i.e., strongly associated with Immanuel Kant and focuses on an agent's motivation), and consequentialism (i.e., strongly associated with John Stuart Mill and Jeremy Bentham and focuses on an act's consequences) (Hassan et al., 2018). Early discussions to answer these questions in a business context date back to Adam Smith (Mingers & Walsham, 2010), the originator of modern market-economy with his seminal work *The Wealth of Nations*, though he has also published an extensive ethical discourse called *The Theory of Moral Sentiments* (Evensky, 2005).

Ethical considerations in IS also have a long history (Hassan et al., 2018). Several research streams have been interested in providing guidance "toward ethical information systems" (Mingers & Walsham, 2010, p. 2). With Weiner's (1950) and Moor's (1985) work, ethics in IS has its origins in computer ethics. In 1986, Mason was one of the first to name ethical issues of the information era; namely, privacy, accuracy, property, and accessibility (PAPA). In addition, Mingers and Walsham (2010) discussed the ethical development in IS in depth, reviewed Smith and Hasnas' (1999) IS and ethical business philosophies, and looked back on critical IS research closely interwoven with Habermasian philosophical perspectives (i.e., theory of communicative action, critical social theory, and discourse ethics). These perspectives see the force of the better argument as the primary legitimation for decision making in social discourse (Mingers & Walsham, 2010). As a consequence, philosophical considerations on ethics have found their way into institutional review boards, codes of conduct, and collections of resources on ethics (e.g., the Association for Information Systems' (AIS) information systems ethics) (Association for Information Systems, 2019).

## 2.2 Ethical Considerations in Design Science

Since designing socio-technical systems involves both technical and social subsystems, ethical considerations in design science have high relevance (Mingers & Walsham, 2010). Philosopher John Rawls was one of the first to introduce ethical discussions in the design process, which led to a "design turn in applied ethics" (van den Hoven, 2008, p. 59). He shifted the focus to the fundamental question: "How can we design the systems, institutions, infrastructures, and IT applications in the context of which users will be able to do what they ought to do" (van den Hoven, 2008, p. 59). More drastically, Mumford and Beekmann (1994, in Leitch & Warren, 2010, p. 2) argued that "if a technical system is created at the expense of a social system, the result obtained will be suboptimal". Driven by this idea, Enid Mumford conceptualized ethical design guidance in the effective technical and human implementation of computer-based systems (ETHICS) approach (Mumford & Ward, 1968). For the first time, ETHICS addressed design with regard to conflicting priorities between job satisfaction and computing-system adoption (Mumford & Ward, 1968). Carrying this argument into the IS discipline, Mason noted that the IS community needs to take "responsibility for the social contract that emerges from the systems we design and implement" (Mason, 1986, p. 11).

Value-sensitive design (VSD) constitutes one approach to address this ethical dilemma. VSD shares, adopts, and expands the traditional approach of participatory design to include all stakeholders in the design process (Friedman, Kahn, & Borning, 2008). VSD accounts for the values that users view as important and implements them in the design process (Friedman et al., 2008; van den Hoven, 2008). Moreover, Fischer (2017) argued that design trade-offs are central to achieve an ethical design. He describes quality of life as a target objective for design science, which involves certain design trade-offs (Fischer, 2017). As a

consequence, many scholars argued for a need for more ethical considerations in DSR (Baskerville & Myers, 2015; Mingers & Walsham, 2010; Prat, Comyn-Wattiau, & Akoka, 2015; Venable, 2009). For instance, Venable (2009) stated that "there remain important questions to be asked about the goals and ethical practice of DSR" (p. 94). In his work, he critically examined different stakeholders' problems and design goals. By applying critical systems heuristics (CSH) (Ulrich, 2002) to DSR, he identified different stakeholder roles in DSR (i.e., client, decision maker, professional, witness) and provided advice on how to address their needs in the design process. Following these thoughts, Myers and Venable (2014) explicitly stated six ethical principles that researchers should consider in DSR (see Table 1). They proposed that researchers should apply ethical considerations as early as possible in a DSR study in order to avoid negative implications for individuals and society (Myers & Venable, 2014). In addition, researchers should include ethical considerations when evaluating artifacts in order to account for risks to people, organizations, or the public (Venable et al., 2016). Likewise, Prat et al. (2015, p. 3) argued that "methods are needed to evaluate the long-term organizational impact of IS artifacts and their societal impact, including ethicality and side effects".

**Table 1. Ethical Principles for DSR (Myers & Venable, 2014)**

| Ethical principle | Description |
|---|---|
| Public interest | Critically consider what benefit or harm may result for all stakeholders. |
| Informed consent | Obtain informed consent from any person involved in the project. |
| Privacy | Ensure that adequate safeguards protect privacy. |
| Honesty and accuracy | Researchers should not plagiarize and should precisely outline their research findings. |
| Property | Ensure that stakeholders agree about who owns the IP. |
| Quality of the artifact | The artifact should be high quality to ensure safety in use. |

## 2.3 Ethical Considerations of Public and Private Initiatives

Beyond research institutions, political, societal, and commercial organizations have also addressed ethical concerns and established society-wide policies, standards, and principles (Jobin, Ienca, & Vayena, 2019; Mittelstadt, 2019). At its core, the United Nations (UN) Agenda for Sustainable Development 2030 aims to improve the world via ethical principles (United Nations, 2016). Likewise, the Geneva Declaration of Principles and Plan of Action propagates a clear ethical mandate in order to reduce negative social implications from IS developments (ITU, 2005). Moreover, 52 proficient experts from industry and academia have defined the European Union (EU) Ethics Guidelines for Trustworthy AI in order to maximize AI's benefits while minimizing its social risk (EU, 2019). In addition, several companies have identified the need to state and publish ethical principles. Particularly, we have seen much development in the AI-enabled artifact domain since the AI technology has the potential to cause a universal change in society for better or for worse (EU, 2019). A recent meta-analysis even found that ethics in AI has become a global topic (Mittelstadt, 2019) that drives the overall ethical debate. Over 80 public and private initiatives have proposed ethical principles to guide AI artifact development, deployment, and governance. We show some studies that have recently reviewed ethical AI principles in Table 2.

In their work, Jobin et al. (2019) analyzed ethical principles and guidelines from 84 documents issued by public and private institutions and revealed a convergence of those principles around five ethical principles: transparency, justice and fairness, non-maleficence, responsibility, and privacy. However, one cannot easily apply most of these principles, and they lack reinforcement mechanisms (Hagendorff, 2020). As a consequence, deviations from these principles have no implications and, thus, the principles often fail (Hagendorff, 2020). Thus, most ethical principles "mainly serve as a marketing strategy" (p. 10) since "economic incentives are easily overriding commitment to ethical principles" (p. 10). Critics argue that many companies engage in so-called "ethic washing" in order to avoid and postpone government regulations (Vincent, 2019), which undermines consumer trust (Peukert & Kloker, 2020). Nevertheless, untruthful attempts to define companywide ethical principles can also backfire as the failure of Google's first ethics board has shown (Jee, 2019).

**Table 2. Reviews Analyzing Existing Ethical AI Principles**

| Reference | Applied method | Number of investigated publications | Key results |
|---|---|---|---|
| Vakkuri & Abrahamsson (2018) | Mapping study | 83 (research papers) | Identified 83 academic papers that refer to AI and ethics since 2012. |
| Jobin et al. (2019) | Literature review | 84 (public and private institutions) | Most ethical principles refer to five principles (transparency, justice and fairness, non-maleficence, responsibility, and privacy) but lack implementation strategies. |
| Mittelstadt (2019) | Comparison study | 84 (public and private institutions) | Compared 84 ethical AI principles with principles of medical ethics and identifies four major shortcomings in existing ethical AI principles. |
| Hagendorff (2020) | Literature review | 21 (public and private institutions) | AI ethics often fails since it is necessary to build tangible bridges between abstract ethical principles and technical implementations. |

# 3 Methodology

To investigate how contemporary DSR publications have implemented ethical principles, we conducted a SLR following established guidelines (Kitchenham, 2004; Webster & Watson, 2002). Kitchenham (2004) divides the SLR process into two successive stages: planning the review (i.e., defining a search strategy) and conducting the review (i.e., executing the data-extraction strategy and analysis). We followed this procedure as we illustrate in Figure 1.

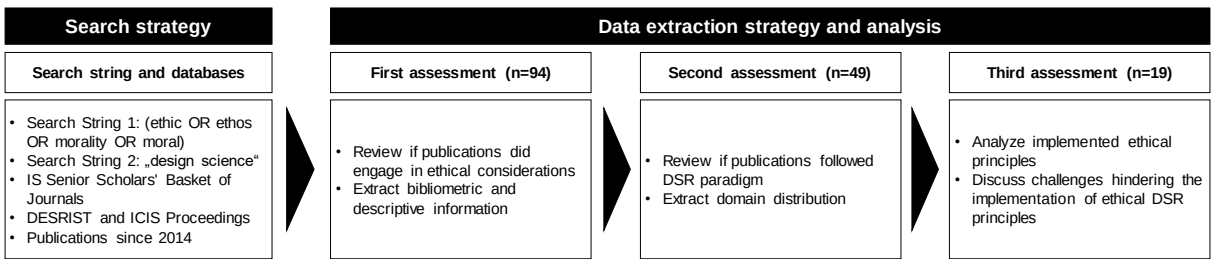


**Figure 1. Search and Data-extraction Strategy**

## 3.1 Search Strategy

To define the search strategy, we first developed two search strings. The first search string included synonyms of ethics (e.g., ethic, ethos, moral, morality) following a prominent online dictionary ("Ethics", n.d.). The second search string accounted for publications in the DSR context (i.e., design science). We conducted the literature review to identify all papers that meet both search strings. Because many papers might only mention terms from one string in their title, abstract, and keywords, we decided to search for both search strings separately and then to review the identified papers manually to identify papers that mentioned ethics or its synonyms and design science.

Subsequently, we determined the journals and databases to search. Because we focused on assessing state-of-the-art DSR research, we restricted our search to the IS Senior Scholars' basket of journals and to the proceedings of the leading DSR conferences (i.e., the International Conference on Design Science Research in Information Systems and Technology (DESRIST) and the International Conference on Information Systems (ICIS)). Moreover, we limited the timeframe of our literature search. We decided to focus on publications published after Myers and Venable (2014) proposed the six ethical principles for DSR in order to reveal how DSR publications have implemented them since its publication. Finally, we applied the search string to the respective databases to retrieve relevant publications: the AIS electronic library for *MIS Quarterly* (*MISQ*), *Journal of the Association for Information Systems* (*JAIS*), and ICIS proceedings; INFORMS database for *Information Systems Research* (*ISR*); Science Direct for *Journal of Strategic Information Systems* (*JSIS*); Springer for *Journal of Information Technology* (*JIT*) and DESRIST proceedings; Taylor & Francis for the *European Journal of Information Systems* (*EJIS*) and *Journal of*

*Management Information Systems* (*JMIS*); and Wiley for *Information Systems Journal* (*ISJ*). We conducted the search at the end of May, 2019; as such, it included all papers from January, 2014, to May, 2019, that met the conditions of one of the two search strings.

### 3.2 Data-extraction Strategy

In the next step, we defined our data-extraction strategy. We followed a three-step assessment in order to extract, collate, and summarize the selected publications (Kitchenham, 2004).

In the first assessment, we reviewed the retrieved publications for their semantic relatedness to ethics and its synonyms to exclude publications that used the terms in our search string but did not engage in an ethical discussion (e.g., citing a reference that includes the term). Subsequently, we analyzed these publications for bibliometric and descriptive information (i.e., author, year, journal, and publication categories adapted from *MISQ*) (MIS Quarterly, 2019).

In the second assessment, we screened the publications we identified from the first assessment and searched for publications that followed the DSR paradigm (Hevner, March, Park, & Ram, 2004). In addition, we reviewed the resulting publications to identify their domain.

In the third assessment, we analyzed how the DSR publications we identified from the second assessment implemented ethical principles. Therefore, we analyzed whether the publications reported implementing Myers and Venable's (2014) six ethical principles. In addition, we extracted challenges hindering their implementation. In order to assure inter-coder reliability in all three assessments, the first and second authors labeled and analyzed all retrieved publications and discussed all deviations in order to achieve mutual agreement.

## 4 Results

In this section, we discuss our results. We first outline the publications we retrieved by executing our search strategy and then outline our results from the three assessment steps in detail.

### 4.1 Retrieved Publications

We executed the search strategy by applying the search strings to the selected databases. We depict how many publications we retrieved from each outlet in Table 3. Following Wilson and Djamasbi's (2015) approach, we present the individual search string results separately for each outlet. In total, we applied the search string to 3,819 publications across the 10 outlets between January, 2014, and May, 2019. This procedure provided a broad perspective of current trends regarding ethics and design science in IS research. We obtained 721 publications after applying the first search string (i.e., ethics and its synonyms) and 636 publications after applying the second search string (i.e., "design science"). To identify the papers that applied to both search strings, we manually downloaded all papers that we identified from both search strings. Next, we reviewed the papers from both sets manually by searching for our search terms using the search function from which we identified 94 publications that met the conditions of both search strings. We identified 53 papers in the IS Senior Scholars' basket of journals, 13 papers in the DESRIST proceedings, and 28 papers in the ICIS proceedings.

We further went into detail by analyzing the papers per publication outlet (see Table 3). Across all publications, we found that 18.88 percent of the papers mentioned ethics-related terms and 16.65 percent mentioned the term design science. We identified 94 publications from manually combining both search strings. Thus, only 2.46 percent of all publications we retrieved mentioned both ethics-related terms and design science. Interestingly, 29.26 percent of all *ISJ* publications mentioned the term ethics or its synonyms and 23.26 percent of all *EJIS* publications mentioned the term design science.

Furthermore, we visualized the number of publications over time (see Figure 2 on the left) and decomposed and visualized the results of the search string components over time (see Figure 2 on the right). The visualization reveals that publications that mentioned the term ethics or its synonyms and design science increased from 2014 to 2018. Since we searched only until May, 2019, we found no increase for 2019.

**Table 3. Distribution of Publications regarding Objectives of Ethics, Design Science, and their Combination**

| Outlet | Total number of publications (January, 2014 to May, 2019) | First search string: "ethic OR ethos OR morality OR moral" | | Second search string: "design science" | | Both search strings combined | |
|---|---|---|---|---|---|---|---|
| | | Amount | % | Amount | % | Amount | % |
| *EJIS* | 215 | 33 | 15.35% | 50 | 23.26% | 14 | 6.51% |
| *ISJ* | 188 | 55 | 29.26% | 26 | 13.83% | 6 | 3.19% |
| *ISR* | 274 | 35 | 12.77% | 16 | 5.84% | 3 | 1.09% |
| *JAIS* | 192 | 51 | 26.56% | 33 | 17.19% | 11 | 5.73% |
| *JIT* | 142 | 32 | 22.54% | 12 | 8.45% | 3 | 2.11% |
| *JMIS* | 268 | 42 | 15.67% | 37 | 13.81% | 7 | 2.61% |
| *JSIS* | 122 | 21 | 17.21% | 10 | 8.20% | 3 | 2.46% |
| *MISQ* | 325 | 58 | 17.85% | 34 | 10.46% | 6 | 1.85% |
| DESRIST | 189 | 13 | 6.88% | 189 | 100.00% | 13 | 6.88% |
| ICIS | 1,904 | 381 | 20.01% | 229 | 12.03% | 28 | 1.47% |

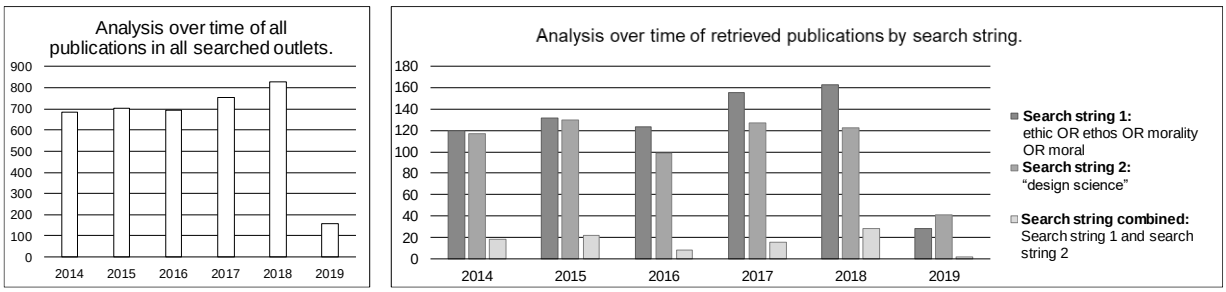


**Figure 2. Analysis of Publications over Time regarding Objectives of Ethics, Design Science, and their Combination**

## 4.2 First Assessment

In the first assessment, we reviewed the 94 publications and checked whether they engaged in ethical considerations. After removing papers that did not sufficiently engage in ethical considerations, 49 publications remained (see Figure 3). Subsequently, we analyzed the remaining 49 publications' bibliometric and descriptive information. Most of the 49 publications (i.e., 13 papers) that engaged in ethical discussions appeared in the ICIS proceedings. We somewhat expected this result since we retrieved the highest number of papers from ICIS overall. In addition, we identified nine publications in *JAIS*, seven publications in *EJIS*, and six publications in the DESRIST proceedings. Thus, about half of the publications in the various outlets engaged in ethical considerations. After we reviewed the publications' publication category, we found that most (i.e., 24 papers) were research papers. We found seven methods papers. Interestingly, we also found 15 editorials and two issue and opinion papers. Consequently, 18 out of the 49 publications (36%) that engaged in ethical considerations were editorials or issue and opinion papers.

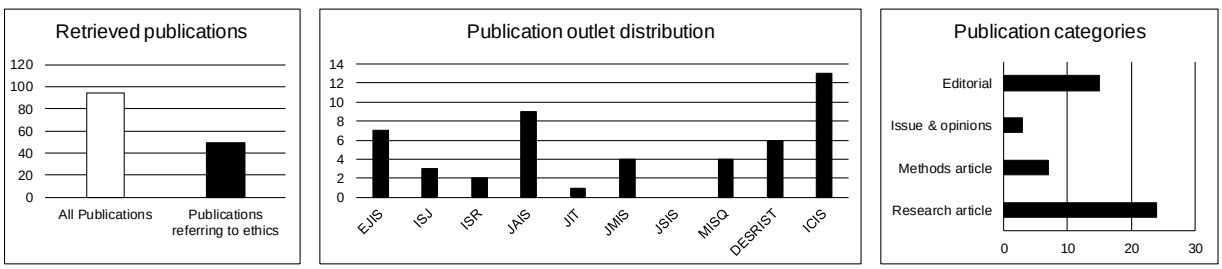


**Figure 3. Results of First Paper Assessment regarding Ethics, Publication Outlet, and Categories**

## 4.3 Second Assessment

In the second assessment (see Figure 4), we reviewed the 49 publications that engaged in ethical considerations and investigated whether they actually applied the DSR paradigm. In total, we identified 19 such publications that actively conducted DSR and designed an artifact (Gregor & Hevner, 2013). Subsequently, we identified the publications' research domain. To identify suitable research domains, we followed Wolfswinkel, Furtmueller, and Wilderom's (2013) grounded theory guidelines: first, we labeled all relevant study excerpts that comprised insights about the study domain. Second, we systematically differentiated, partitioned, and integrated these excerpts in iterative cycles in order to identify the high-level domains. By conducting this approach, we grouped the 19 research papers into seven research domains. By analyzing the results (see Figure 4), we found that most of the research papers that discussed ethics and applied the DSR paradigm came from the healthcare domain; developed, improved, or evaluated methods and frameworks for conducting DSR; or addressed the topics safety, security, and privacy.

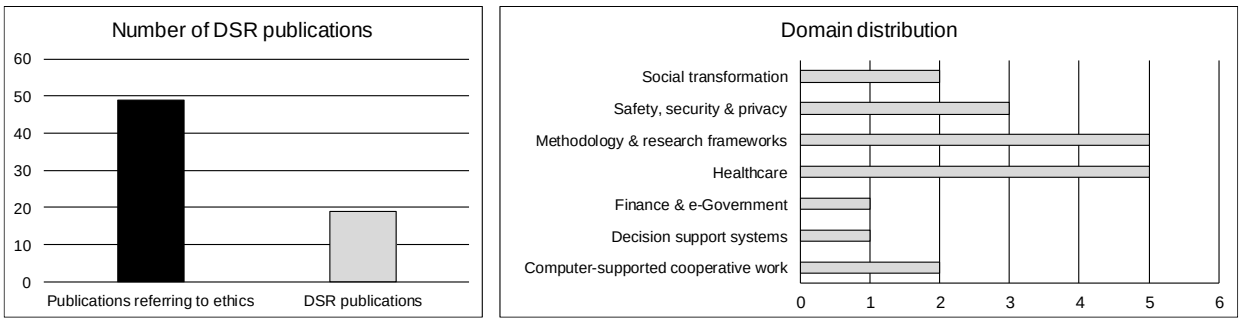


**Figure 4. Results of Second Paper Assessment regarding DSR and Research Domain**

## 4.4 Third Assessment

We assessed the resulting 19 publications that engaged in ethical considerations and followed the DSR paradigm in two steps: 1) we analyzed the implemented ethical principles and 2) we discussed the challenges hindering their implementation.

### 4.4.1 Analysis of Ethical DSR Principles

To assess the remaining 19 DSR publications that engaged in ethical considerations in more detail, we first analyzed the papers' authors. We found that no specific authors appeared constantly. Next, we investigated whether they discussed the six ethical principles that Myers and Venable (2014) proposed. To do so, the first and second authors reviewed the 19 DSR papers and labeled relevant study excerpts that addressed one of the six ethical principles. Subsequently, they extracted all labels and coded each paper as to whether it implemented each principle or not. In situations where labeling conflicts arose, they reviewed the publication again and discussed the publications to code the study based on mutual agreement. To outline the coding approach in more detail, we illustrate two example publications, their extracted study excerpts, and the assigned codes in Table A1 in the Appendix.

We illustrate our results from the coding exercise in Figure 5. Overall, 14 of the 19 DSR research papers considered whether their research may benefit or harm stakeholders. Another 11 publications stated that they received consent from all stakeholders involved in their research projects. Furthermore, 10 publications mentioned privacy concerns, and nine publications discussed the quality of the artifact. Overall, five publications stated that they ensured honest and accurate research, and four publications explicitly reported on the intellectual property of the research. Fewer DSR publications explicitly mentioned the other two principles (i.e., honesty and accuracy as well as property).

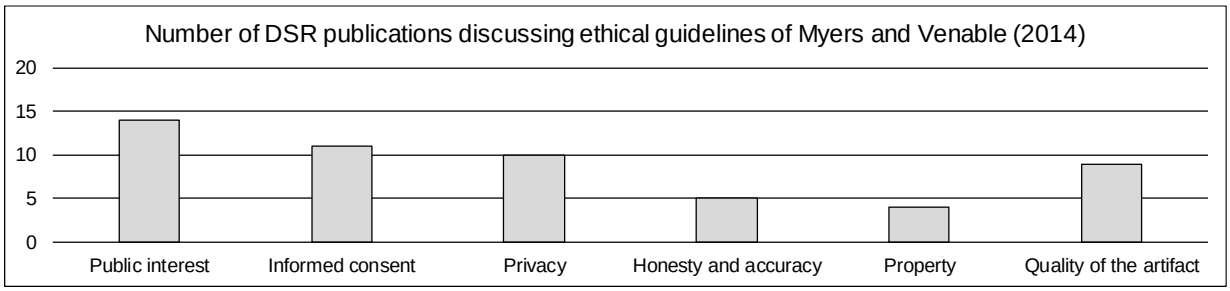


**Figure 5. Analysis of the Ethical Principles in the 19 Identified DSR Publications**

### 4.4.2 Challenges in Implementing Ethical Principles in DSR

We observed that DSR researchers face significant challenges that hinder them from addressing Myers and Venable's (2014) ethical principles. We report the challenges that we identified in our SLR in this section to showcase the complexity of implementing ethical principles and provide insights why the DSR studies we identified insufficiently realized them.

**Public interest:** the first ethical principle demands that researchers critically assess their DSR projects to identify what may benefit or harm any stakeholders. In the claim of this principle, it is apparently impossible for DSR researchers to consider all stakeholders since the term "critically assess" is rather imprecise. That impreciseness creates uncertainty for DSR researchers on how to proceed. On the other hand, it is not easy to implement the principle. For example, Germonprez et al. (2017) investigated corporate engagement in open source communities. However, one cannot easily identify all stakeholder groups in the body of open source communities due to their nature (e.g., people often contribute anonymously), which leads to multiple differentiating groups. Similarly, Gregor, Imran, and Turner (2014) investigated e-Government in Bangladesh. However, it might be too abstract to consider the whole country as one stakeholder group since the administration might have different interests than the population.

**Informed consent**: informed consent demands that any persons involved in a DSR project provide their consent for participation. Researchers cannot obtain such consent if they cannot reach a stakeholder or if they failed to consider a particular stakeholder. For example, researchers cannot easily receive consent for an artifact that the whole population might use in the future. Lee, Cho, and Lim's (2018) bright Internet initiative represents such a case in which the authors designed a new Internet infrastructure to protect innocent citizens from cybersecurity threats. Thus, the informed consent principle should provide more detailed boundary conditions.

**Privacy**: the third ethical principle demands adequate safeguards to protect privacy. Nevertheless, researchers cannot easily maintain it. Due to many artifacts' high technical complexity, one needs to design multiple components in such a way as to safeguard private information. As such, researchers could end up discussing an innovation's security components more than its theoretical contribution. For example, when conducting a DSR project to develop a healthcare platform (Keijzer-Broers & De Reuver, 2016), researchers need to outline how the platforms protect and preserve private information for several different software components—a highly challenging task that the DSR project may not focus on.

**Honesty and accuracy**: this principle addresses what true honestly actually means. The answer to this question depends highly on the context, the goal, and a DSR project's contribution. For example, does one achieve honesty via explaining how a piece of software code excels or can one simply explain what one intends an artifact to do? Since the existing principle currently misses a detailed prescription, the existing publications had limited guidance on how to discuss this principle.

**Property**: the fifth principle demands that researchers need to determine property rights in order to avoid legal harm to stakeholders. In DSR studies, researchers can assure legal determination of property rights through licensing their artifact and by agreeing to specific publication formats. However, that task becomes more complex when the usage of an artifact produces information or data (e.g., a healthcare platform that creates new data by analyzing patients) (Keijzer-Broers & De Reuver, 2016). However, it appears that the existing property principle does not deal with such issues yet.

**Quality of the artifact**: the sixth principle demands high-quality artifacts. Although researchers face manifold challenges in achieving such quality (e.g., correct data sampling, sensors, and software

development), DSR researchers themselves can probably best control how they apply this principle since rigorously evaluating a proposed artifact constitutes one key contribution of a DSR study (Venable, Pries-Heje, & Baskerville, 2016). However, all publications that conducted a completed DSR study contained an evaluation with confirming results. Nevertheless, one cannot guarantee quality holistically since others can always challenge evaluation methods and biases.

# 5 Discussion

In this paper, we investigate how thought-leading DSR publications address ethical principles and their implementation in DSR. In this section, we first discuss results we obtained from our SLR. Based on this foundation, we advance our research and propose two pathways toward ethical DSR. Finally, we conclude our study's limitations.

## 5.1 Implementing Ethical Principles in DSR

In the SLR, we analyzed the body of DSR literature in leading IS journals (IS Senior Scholars' basket of journals) and conferences (DESRIST and ICIS proceedings) since the publication about DSR ethical principles (Myers & Venable, 2014). Among the 94 publications that we initially identified in our search category, we found a discrepancy between publications that used the terms ethics and its synonyms and design science (94 publications) and publications that actually engaged in discussing ethics (49 publication or 52%). Overall, this ratio reflects two tendencies in recent literature. The first tendency refers to publications that mentioned ethics-related terms but did not discuss them in detail. The second tendency refers to publications that discussed the importance of ethics and the concrete implementation of ethical principles in order to explicitly address potential risks for individuals and society. When taking a more detailed look at the 49 publications that engaged in ethical considerations, we found that editorials, opinions, and issue papers comprised more than one-third (37%) of the papers. Primarily editorial board members write editorials in order to shift the discussion to future research directions, to identify research gaps, and to provide principles for the journal's scholars (Journal of thr Association for Information Systems, 2019), which indicates that current ethical discussions in DSR projects might still be in a conceptual stage while awaiting their concrete implementation in DSR.

### 5.1.1 DSR Studies Engaging in Ethical Considerations

We identified only 19 research papers that engaged in ethical considerations and explicitly developed an artifact in a DSR project. Looking at their domains, we found that publications conducting DSR in healthcare addressed ethical considerations very often, which appears reasonable since studies dealing with human health and wellbeing need to follow very high ethical standards in order to avoid harm to individuals. Similarly, ethical review boards at universities and laboratories have profound ethical expertise in the health sector, which fosters ethically conformant research in that discipline. This expertise will likely lead DSR researchers to identify ethical considerations in DSR projects earlier on since they may compare their DSR projects directly to research projects that normally receive ethical scrutiny.

In addition, many DSR publications developed methodology and research framework artifacts in our final sample. On the one hand, researchers should address ethics when developing methodology and research frameworks, which DSR projects will apply in the future. Such work has particular value since the most prominent frameworks for conducting DSR do not mention the term ethics so far. Thus, DSR methodology and research framework publications that acknowledge the need for ethical considerations represent a positive sign. On the other hand, it might be easier to argue for the necessity to address ethical considerations than to practically address them in DSR projects.

### 5.1.2 Distribution of Ethical DSR Principles

Diving deeper into how the publications in our sample implemented the six ethical principles, we found few publications that explicitly addressed them all. Comparing the individual principles, we found a tendency for the publications to address two differentiating groups of principles. The first group (which contained more than half of the DSR papers) discussed public interest, informed consent, privacy, and quality of the artifact. The second group (which contained a quarter of the papers) discussed honesty and accuracy as well as property. One reason for this difference might be the complexity of the latter two principles, which might be more difficult to report than the others. For example, it might be easier to argue that one has addressed the public interest than property. Another reason could be that the academic community so strongly reinforces

ethical principles such as honesty and accuracy that subscribing to them constitutes a prerequisite for membership in the community.

In conclusion, we found that contemporary DSR publications do not explicitly address several ethical principles. We found specific manifestations in different domains and an unbalanced distribution in the attention that studies have paid to the six ethical principles for DSR. Our findings raise the question of why DSR publications rarely report ethical considerations. While the answer to this question might appear simple at first glance, it creates important implications when viewed again.

As we examined peer-reviewed, high-quality journal and conference publications, they were all liable to their respective codes of conduct with which authors agreed before publication. Consequently, the authors implicitly covered ethical considerations in submitting their papers. However, this answer might fall short since not mentioning how they implemented ethical principles as part of a research project neither reinforces the importance of ethics in DSR nor supports the consolidation of ethical discussions in DSR. As a consequence, scholars and, in particular, less experienced researchers face difficulties in finding real-world examples that provide blueprints for answering elementary ethical questions about a DSR project (e.g., how to balance a DSR project's benefits against its potential risks (i.e., balancing ethical design trade-offs)). By adding metadata to papers, researchers could report their ethical considerations in a structured format. Such a practice could complement existing DSR knowledge bases. Thus, the DSR community may currently not leverage its potential to shape the ethical discussion in how socio-technical artifacts develop because it does not sufficiently report ethical considerations.

## 5.2 Towards Ethical DSR

To conduct and promote ethical considerations in DSR, we argue that both DSR researchers who conduct the research and DSR reviewers (e.g., editorial board members, associate chairs, reviewers for journals/conferences) who evaluate and publish researchers' work require more support. First, DSR researchers might need more prescriptive guidance on how to translate existing ethical principles into their individual DSR projects. Therefore, we propose that they use prescriptive knowledge structures from DSR. Second, DSR researchers and reviewers might need more support in reporting and evaluating a DSR project's ethical implications. Therefore, we extend established DSR conceptualizations by adding an ethical dimension and introduce the concept of an E-DSR process model. In this section, we outline both pathways to conduct and promote ethical DSR in more detail.

### 5.2.1 Prescriptive Knowledge Structures for Ethical Principles for DSR

To support DSR researchers to engage in ethical considerations, we argue for more prescriptive guidance in translating ethical principles to a specific research project. We propose to rely on established DSR theory to provide such guidance. An inherent strength of DSR is to transform descriptive Ω-knowledge into prescriptive λ-knowledge in order to describe design implications of innovative socio-technical artifacts (Gregor & Hevner, 2013). To articulate meaningful λ-knowledge, several scholars have provided frameworks on how to articulate general design theories (Gregor & Jones, 2007) but also on how to articulate more specific design principles (Chandra, Seidel, & Gregor, 2015; Van Aken, 2004). As a result, researchers have established commonly applied structures on what knowledge components design principles in DSR should include.

In this regard, design principles usually include three knowledge components. They should describe how 1) a specific property/action can 2) lead to the desired state 3) under certain boundary conditions and contexts (Chandra et al., 2015; Djamasbi & Strong, 2019; Gregor, Chandra Kruse, & Seidel, 2020; van Aken, 2004). Applying this structure to existing ethical principles (see Table 4) shows that some do not include all the three knowledge components, which might constitute one important reason why it may be difficult to translate and apply the existing ethical principles in a concrete DSR project.

To address this difficulty, prescriptive knowledge structures from DSR could help researchers articulate more applicable ethical principles that they can contextualize in a specific DSR project. Therefore, exploiting the DSR community's existing approaches to define design knowledge represents a promising point to start this endeavor (Gregor & Hevner, 2013). For example, researchers could start by transforming the broad landscape of existing ethical principles (e.g., the ethical principles that Mittelstadt (2019) identified from 84 public and private initiatives) into the prescriptive knowledge structures for DSR. The transformed ethical principles then include the required knowledge pieces that make them more applicable in a specific DSR project. Additionally, reviewers of DSR publications can refer to these ethical principles when assessing

DSR researchers' work. While we lack the scope here, we encourage the research community to consider how to elevate existing ethical principles to the next, prescriptive rung of knowledge. We suggest that doing so might prove to be an important step toward supporting DSR researchers to more effectively achieve ethical DSR.

**Table 4. Exemplary Analysis of Ethical Principles Regarding Prescriptive Knowledge Structures for DSR**

| | | | Analysis of ethical principles regarding prescriptive knowledge structures for DSR (i.e., design principles) | | |
|---|---|---|---|---|---|
| **Reference** | **Ethical principle** | **Description of principles** | **(How) Information about the property/action** | **(What) Desired states made possible through the property/action** | **(When) Boundary conditions for when the principle applies** |
| Myers & Venable (2014) | Public interest | Critically consider what benefit or harm may result for all stakeholders. | "Critically consider" is an indefinite action. | The source does not explicitly define the desired state (e.g., a stakeholder benefit/harm map). | The principle does not explicitly set boundary conditions but refers to "all" stakeholders. |
| EU (2019) | Stakeholder participation | In order to develop trustworthy AI systems, one should consult stakeholders who may directly/indirectly be affected by the system through its lifecycle. | Source does not explicitly define what "consulting" constitutes. | Although the source defines the goal of developing trustworthiness, it does not state how one can measure it. | The principle does not define concrete moments in time. |
| Microsoft (2019) | Make clear what the system can do | Help users understand what the AI system can do. | "Help users understand" does not define an explicit action and is rather abstract. | The source does not clearly state how one can operationalize and measure this state. | The principle does not clearly identify which users and what system capabilities. |

### 5.2.2 Ethical Design Science Research Process Model

To help researchers and reviewers embrace ethical principles in DSR publications, we suggest that they need a supportive approach to report their ethical considerations as part of their DSR projects and to evaluate the ethical conformity of DSR publications regarding ethical principles, respectively. To facilitate the communication about ethical considerations between DSR researchers and reviewers in a research cycle, we suggest that they require an explicit structure for reporting and evaluating these considerations. However, a major challenge for coherently reporting ethical considerations in DSR concerns the fact that researchers have no guidance on 1) how to apply and 2) where to report them when following established DSR process models.

To include ethical principles in a DSR project, we argue that researchers need to rigorously draw on existing ethical principles and to apply them in a DSR project's key activities. DSR projects' key conceptual activities comprise creating and accumulating design knowledge in the problem and solution space (Hevner, vom Brocke, & Maedche, 2019; Hevner et al., 2004). The problem space represents researchers' understanding of a problem that they address, while the solution space comprises the concepts about the solution artifact in terms of solution entities, search criteria, and build and evaluation actions (Drechsler & Hevner, 2018; Venable, 2006). The distance between both spaces iteratively diminishes with each design cycle while researchers continuously draw on and accumulate knowledge from the Ω-knowledge and λ-knowledge bases (Drechsler & Hevner, 2018; Hevner et al., 2019). We argue that, in this continuous process of executing DSR projects, researchers should explicitly account for ethical considerations when describing the problem space and should apply ethical principles when exploring the solution space. By doing so, other researchers could learn about ethical considerations from other research projects related to the corresponding problem space and simultaneously contribute their ethical considerations to the body of Ω-knowledge. Similarly, researchers could draw on existing ethical principles when exploring the solution space while also contributing new or adapted ethical principles to the body of λ-knowledge. We visualize

this relationship by extending Drechsler and Hevner's (2018) DSR knowledge utilization, production, and contribution framework (see Figure 6).

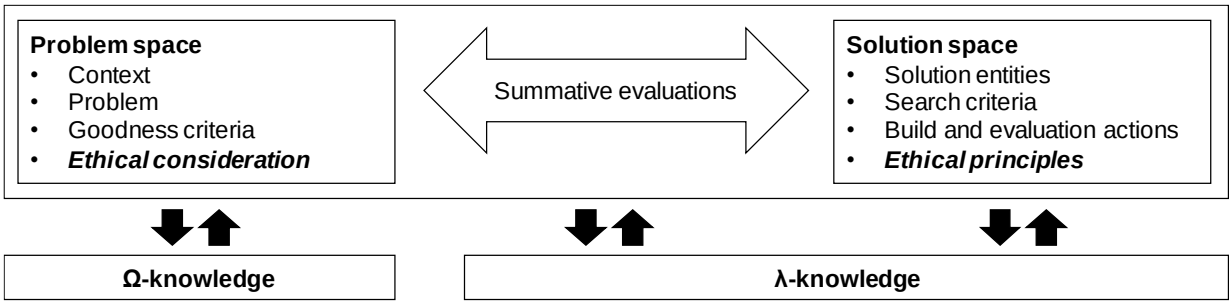


**Figure 6. DSR Knowledge Utilization, Production, and Contribution in DSR (Based on Drechsler & Hevner 2018) Extended with an Ethical Knowledge Dimension**

To help researchers report ethical considerations and how they apply ethical principles, we provide a structure by extending existing DSR process models. Therefore, we propose adding an ethical dimension to existing DSR process models—a reasonable idea due to their familiarity in the community, their big impact, and their proven applicability in prescriptive theory building (Gregor & Hevner, 2013; Gregor & Jones, 2007). Such an addition could encourage DSR researchers to report their ethical considerations as part of their DSR projects and support reviewers to evaluate and publish such projects. To facilitate such actions, we propose that an ethical dimension of a DSR process model should address two main aspects.

First, the ethical dimension should be integrated into the DSR process models as ongoing and continuous introspective reflections throughout an entire DSR project. DSR researchers should, in every phase of a DSR project, balance the scientific benefits of action with the implications and risks for each stakeholder group that the project may affect. On this path, ethical design trade-offs will occur, which researchers have to continuously assess against the artifact's functionality, scientific benefit, and ethical implications (Fischer, 2017). For example, when using AI-based algorithms to measure individual affective states (i.e., emotions), researchers have to reflect on the impact they may have on individuals and should potentially cut functionality that provides a potential for abuse (Benke et al., 2020; Tarafdar, Teodorescu, Tanriverdi, Robert, & Morse, 2020). We propose that, in every phase of a DSR project, researchers should assess whether the benefits outweigh the risks for those involved and should describe this assessment process. We further argue that such an explicit integration into established DSR process models would promote the need for researchers to continuously reflect on ethical design trade-offs.

Second, the ethical dimension should be integrated into existing DSR process models by adding an additional phase. This phase provides DSR researchers with a dedicated space to report on the ethical design trade-offs that occurred and to demonstrate how they addressed the trade-offs in a DSR project. Extending existing DSR process models in this way would make the ethical dimension an integral part of DSR processes. It would provide clear guidance for DSR researchers on where to demonstrate the ethical design trade-offs and for reviewers on where to search for them in order to decide on a DSR publication's ethical conformity with a journal's guidelines. Therefore, adding such a phase would reinforce a critical and transparent reflection on DSR research projects' ethical conformity, which would be necessary due to ethical principles' commensurability (i.e., ethical considerations may not share a common measurement standard). While the ethical design trade-offs may be subject to variation depending on the research domain, publication outlet, or contextual circumstances, the additional phase would ensure a common space for reporting and reviewing ethical considerations in DSR. The phase would also allow DSR researchers to compare their ethical endeavors and promote a learning process for the DSR research community (Myers & Venable, 2014). To provide explicit guidance, we instantiate the ethical dimension by adding it to Kuechler and Vaishnavi's (2012) DSR process model (see Figure 7) and propose an E-DSR process model.

Finally, the E-DSR process steps do not always mirror the way researchers present it in a paper. Therefore, we propose that the ethical dimension should also be added to established structural schemes that guide the way researchers present a DSR paper (Gregor & Hevner, 2013). For example, including an additional subsection to Gregor and Hevner's (2013) publication schema and to the structural schema for proposing a design theory (Gregor & Jones, 2007) could further help researchers demonstrate ethical design trade-offs that they experience and address in E-DSR.

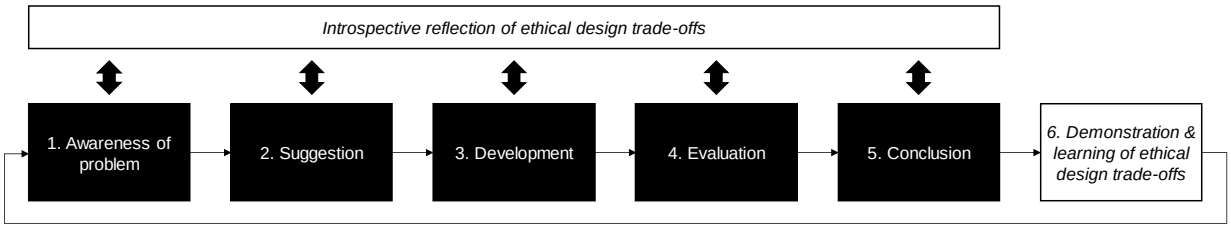


**Figure 7. Ethical Design Science Research (E-DSR) Process Model**

# 6 Limitations

Although we followed a rigorous structure throughout our research, some limitations apply that provide avenues for further research. First of all, the SLR method includes natural limitations. According to Webster and Watson (2002), SLRs may not be exhaustive. Although we followed a rigorous search strategy following Kitchenham's (2004) principles, we have missed important publications that appeared in different outlets to the ones we considered. Because the papers we investigated came from major outlets known for publishing high-quality DSR studies, we believe that they provide a sufficient sample to achieve our study goals (i.e., to identify a suitable sample of contemporary DSR publications that have implemented ethical principles). Second, we only analyzed DSR publications since 2014—the year in which Myers and Venable (2014) published their ethical principles for DSR. Consequently, our findings only apply to the most recently published DSR publications and do not generalize to DSR and related design-oriented disciplines as a whole. Therefore, this limitation provides further research opportunities to extend the review with publications from other design-oriented disciplines. Third, in this paper, we focus on Myers and Venable's (2014) ethical DSR principles. However, we could have analyzed the DSR papers we identified according to other highly valuable ethical principles (e.g., Mittelstadt, 2019). However, since Myers and Venable (2014) articulated their principles for DSR in particular, we selected them as most relevant. Fourth, we argue that researchers could address the insufficient extent to which they have implemented ethical principles via transforming existing ethical principles using prescriptive knowledge frameworks from DSR. To highlight Mittelstadt's (2019) work once again, we recognize that one could use other methods to implement ethics in AI-based technologies. Other interventions also have important influences on efforts to implement ethical principles (e.g., democratization of technology and laws). However, we focus on DSR. Fifth, besides methodological limitations, we recognize that other non-DSR-related challenges may drive our results. For example, space limitations in certain outlets could lead authors to not report ethical considerations that they conducted before, during, and after their research (also a known issue in other domains). Sixth, ethical considerations depend highly on the context and are sensitive to cultural differences. Researchers from different cultural backgrounds might have conflicting views on our findings. However, our findings can serve as a starting point for future discussions with the international and intercultural DSR community.

# 7 Conclusion

In this paper, we assess how contemporary DSR publications have implemented ethical principles. We found that not all DSR publications have discussed how they implemented ethical principles in-depth. Based on this finding, we critically examine potential challenges that hinder researchers from implementing Myers and Venable's (2014) six ethical principles. We conclude that existing ethical principles are too abstract in nature to apply to a broad range of contexts, which makes it difficult to apply them in specific DSR projects. In addition, we conclude that researchers and reviewers lack explicit structures on where to report and on how to evaluate ethical considerations in DSR projects. Therefore, we propose two pathways towards ethical DSR. First, we argue that the next generation of ethical principles for DSR should be articulated using prescriptive knowledge structures from DSR. Second, we propose extending established DSR process models with an ethical dimension and, therefore, introduce the concept of E-DSR.

# Acknowledgments

We thank the Editor-in-Chief Fiona Nah and Special Issue Editors Lionel P. Robert, Gaurav Bansal, and Tom Stafford for their support in publishing this manuscript and their suggestions for improvement. Especially, we thank the Special Issue and Senior Editor Nigel Melville as well as the anonymous reviewers for their thoughtful suggestions that helped us to improve the quality of this paper.

## References

André, E., Bayer, S., Benke, I., Benlian, A., Cummins, N., Gimpel, H., Hinz, O., Kersting, K., Maedche, A., Muehlhaeuser, M., Riemann, J., Schuller, B. W., & Weber, K. (2019). Humane anthropomorphic agents : The quest for the outcome measure. In *Proceedings of the Pre-ICIS Workshop "Values and Ethics in the Digital Age".*

Association for Information Systems. (2019). *Association of information systems ethics code.* Retrieved from https://aisnet.org/page/ISEthics

Baskerville, R. L., & Myers, M. D. (2015). Design ethnography in information systems. *Information Systems Journal*, *25*(1), 23-46.

Benke, I. (2020). Towards design principles for trustworthy affective chatbots in virtual teams. In *Proceedings of the 28th European Conference on Information Systems.*

Benke, I., Knierim, M. T., & Maedche, A. (2020). Chatbot-based emotion management for distributed teams: A participatory design study. In *Proceedings of the ACM on Human-Computer Interaction*.

Chandra, L., Seidel, S., & Gregor, S. (2015). Prescriptive knowledge in IS research: Conceptualizing design principles in terms of materiality, action, and boundary conditions. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (pp. 4039-4048).

Chatterjee, S., Byun, J., Dutta, K., Pedersen, R. U., Pottathil, A., & Xie, H. (Qi). (2018). Designing an Internet-of-things (IoT) and sensor-based in-home monitoring system for assisting diabetes patients: Iterative learning from two case studies. *European Journal of Information Systems*, *27*(6), 670-685.

Demetis, D., & Lee, A. S. (2018). When humans using the IT artifact becomes IT using the human artifact. *Journal of the Association for Information Systems*, *19*(10), 929-952.

Djamasbi, S., & Strong, D. (2019). User experience-driven innovation in smart and connected worlds. *AIS Transactions on Human-Computer Interaction*, *11*(4), 215-231.

Drechsler, A., & Hevner, A. R. (2018). Utilizing, producing, and contributing design knowledge in DSR projects. In *Proceedings of the Design Science Research in Information Systems and Technology.*

Ethics. (n.d.). In *Merriam-Webster*. Retreived from https://www.merriam-webster.com/thesaurus/ethics

EU. (2018). *AI strategic note*. European Commission. Retrieved from https://ec.europa.eu/jrc/communities/sites/jrccties/files/epsc_strategicnote_ai.pdf

EU. (2019). Ethics guidelines for trustworthy AI. HLEG AI, European Commission. Retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation

EU. (2020). *The ethics of artificial intelligence: Issues and initiatives*. European Parliamentary Research Service. Retrieved from https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf

Evensky, J. (2005). Adam Smith's theory of moral sentiments: On morals and why they matter to a liberal society of free people and free markets. *Journal of Economic Perspectives*, *19*(3), 109-130.

Feine, J., Adam, M., Benke, I., Maedche, A., & Benlian, A. (2020a). Exploring design principles for enterprise chatbots: An analytic hierarchy process study. In *Proceedings of the 15th International Conference on Design Science Research in Information Systems and Technology.*

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019a). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, *132*, 138-161.

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2020b). Gender bias in chatbot design. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, & P. B. Brandtzaeg (Eds.), *Chatbot research and design* (pp. 79-93). Berlin: Springer.

Feine, J., Morana, S., & Maedche, A. (2019b). Designing a chatbot social cue configuration system. In *Proceedings of the 40th International Conference on Information Systems.*

Feine, J., Morana, S., & Maedche, A. (2020c). Designing interactive chatbot development systems. In *Proceedings of the 41st International Conference on Information Systems.*

Feine, J., Morana, S., & Maedche, A. (2020d). A chatbot response generation system. In *Proceedings of the Conference on Mensch und Computer.*

Fischer, G. (2017). Exploring design trade-offs for quality of life in human-centered design. *Interactions*, *25*(1), 26-33.

Friedman, B., Kahn, P. H., & Borning, A. (2008). Value sensitive design and information systems. In K. Himma & H. Tavan (Eds.), *The handbook of information and computer ethics* (pp. 69-101). New York, NY: John Wiley & Sons.

Germonprez, M., Kendall, J. E., Kendall, K. E., Mathiassen, L., Young, B., & Warner, B. (2017). A theory of responsive design: A field study of corporate engagement with open source communities. *Information Systems Research*, *28*(1), 64-83.

Gnewuch, U., Morana, S., Adam, M. T. P., & Maedche, A. (2018). Faster is not always better: Understanding the effect of dynamic response delays in human-chatbot interaction. In *Proceedings of the 26th European Conference on Information Systems.*

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science for maximum impact. *MIS Quarterly*, *37*(2), 337-355.

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems*, *8*(5), 312-335.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). The anatomy of a design principle. *Journal of the Association for Information Systems*, *21*(6), 1622-1652.

Gregor, S., Imran, A., & Turner, T. (2014). A "sweet spot" change strategy for a least developed country: Leveraging e-government in Bangladesh. *European Journal of Information Systems*, *23*(6), 655-671.

Hagendorff, T. (2020). The ethics of AI ethics—an evaluation of guidelines. *arXiv*. Retrieved from http://arxiv.org/abs/1903.03425

Hassan, N. R., Mingers, J., & Stahl, B. (2018). Philosophy and information systems: Where are we and where should we go? *European Journal of Information Systems*, *27*(3), 263-277.

Hern, A. (2018). Cambridge Analytica scandal "highlights need for AI regulation". *The Guardian*. Retrieved from https://www.theguardian.com/technology/2018/apr/16/cambridge-analytica-scandal-highlights-need-for-ai-regulation

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, *28*(1), 75-105.

Hevner, A. R., vom Brocke, J., & Maedche, A. (2019). Roles of digital innovation in design science research. *Business and Information Systems Engineering*, *61*(1), 3-8.

Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian Journal of Information Systems*, *19*(2), 39-64.

ITU. (2005). *World summit on the information society*. World Summit on the Information Society (WSIS). https://www.itu.int/net/wsis/

Jee, C. (2019). Google has now cancelled its AI ethics board after a backlash from staff. *MIT Technology Review.* Retrieved from https://www.technologyreview.com/f/613271/google-has-now-cancelled-its-ai-ethics-board-after-a-backlash-from-staff/

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389-399.

Journal of the Association for Information Systems. (2019). JAIS manuscript categories—information for authors. *Assocation for Information Systems.* Retrieved from https://aisel.aisnet.org/jais/authorinfo.html

Keijzer-Broers, W. J. W., & De Reuver, M. (2016). Applying agile design sprint methods in action design research: Prototyping a health and wellbeing platform. In *Proceedings of the 11th International Conference on Design Science Research in Information Systems and Technology.*

Kitchenham, B. A. (2004). *Procedures for performing systematic reviews*. Retrieved from http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf

Kuechler, W. L., & Vaishnavi, V. K. (2012). A framework for theory development in design science research: Multiple perspectives. *Journal of the Association for Information Systems*, *13*(6), 395-423.

Lee, J., Cho, D., & Lim, G. (2018). Design and validation of the bright Internet. *Journal of the Association for Information Systems*, *19*(2), 63-85.

Leitch, S., & Warren, M. J. (2010). ETHICS: The past, present and future of socio-technical systems design. In *Proceedings of the IFIP International Conference on the History of Computing*.

Mason, R. O. (1986). Four ethical issues of the information age. *Computer Ethics*, *10*(1), 41-48.

Microsoft. (2019). *Guidelines for human AI interaction design*. Retrieved from https://www.microsoft.com/en-us/research/blog/guidelines-for-human-ai-interaction-design/

Mingers, & Walsham. (2010). Toward ethical information systems: The contribution of discourse ethics. *MIS Quarterly*, *34*(4), 833-854.

MIS Quarterly. (2019). *MISQ manuscript categories*. Retrieved from https://misq.org/categories/

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*, 501-507.

Moor, J. H. (1985). What is computer ethics? *Metaphilosophy*, *6*(4), 266-275.

Moore, G. E. (1993). *Principia ethica: With the preface to the second edition and other papers.* Cambridge, UK: Cambridge University Press.

Mumford, E., & Beekmann, G. J. (1994). *Tools for change & progress: A socio-technical approach to business process re-engineering*. Leiden: The Netherlands: CSG Publications.

Mumford, E., & Ward, T. B. (1968). *Computers: Planning for people*. London, UK: Batsford.

Myers, M. D., & Venable, J. R. (2014). A set of ethical principles for design science research in information systems. *Information and Management*, *51*(6), 801-809.

Peukert, C., & Kloker, S. (2020). Trustworthy AI: How ethicswashing undermines consumer trust. In *Proceedings of the 15th International Conference on Wirtschaftsinformatik.*

Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, *32*(3), 229-267.

Rietz, T., Benke, I., & Maedche, A. (2019). The impact of anthropomorphic and functional chatbot design features in enterprise collaboration systems on user acceptance. In *Proceedings of the 14th International Conference on Wirtschaftsinformatik.*

Robert, L. P., Jr., Bansal, G., & Lütge, C. (2020). *ICIS 2019 SIGHCI Workshop panel report: Human-computer interaction challenges and opportunities for fair, trustworthy and ethical artificial intelligence. AIS Transactions on Human-Computer Interaction, 12*(2), 96-108.

Shneiderman, B. (2020). Human-centered artificial intelligence: Three fresh ideas. *AIS Transactions on Human-Computer Interaction*, *12*(3), 109-124.

Smith, H. J., & Hasnas, J. (1999). Ethics and information systems: The corporate domain. *MIS Quarterly*, *23*(1), 109-127.

Tarafdar, M., Teodorescu, M., Tanriverdi, H., Robert, L. P., Jr., & Morse, L. (2020). Seeking ethical use of AI algorithms: Challenges and mitigations. In *Proceedings of the 41th International Conference on Information Systems.*

Ulrich, W. (2002). Critical systems heuristics. In H. G. Daellenbach & R. L. Flood (Eds.), *The informed student guide to management science*. London, UK: Thomson Learning.

United Nations. (2016). *The sustainable development agenda 2030*. Retrieved from https://www.un.org/sustainabledevelopment/development-agenda/

Vakkuri, V., & Abrahamsson, P. (2018). The key concepts of ethics of artificial intellligence: A keyword

based systematic mapping study. In *Proceedings of the IEEE International Conference on Engineering, Technology and Innovation.*

Van Aken, J. E. (2004). Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules. *Journal of Management Studies*, *41*(2), 219-246.

van den Hoven, J. (2008). Moral methodology and information technology. In K. E. Himma & H. T. Tavani (Eds.), *The handbook of information and computer ethics*. New York, NY: John Wiley & Sons.

Venable, J. R. (2006). The role of theory and theorising in design science research. In *Proceedings of the 1st International Conference on Design Science in Information Systems and Technology.*

Venable, J. R. (2009). Identifying and addressing stakeholder interests in design science research: An analysis using critical systems heuristics. In G. Dhillon, B. C. Stahl, & R. Baskerville (Eds.), *Information systems—creativity and innovation in small and medium-sized enterprises* (IFIP AICT vol. 301). Berlin: Springer.

Venable, J. R., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A framework for evaluation in design science research. *European Journal of Information Systems*, *25*(1), 77-89.

Vincent, J. (2019). The problem with AI ethics. *The Verge*. Retrieved from https://www.theverge.com/2019/4/3/18293410/ai-artificial-intelligence-ethics-boards-charters-problem-big-tech

Webster, J., & Watson, R. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), 13-23.

Weiner, N. (1950). *The human use of human beings: Cybernetics and society*. London, UK: Free Association Books.

Wilson, E. V., & Djamasbi, S. (2015). Human-computer interaction in health and wellness: Research and publication opportunities. *AIS Transactions on Human-Computer Interaction*, *7*(3), 97-108.

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45-55.

# Appendix: Illustration of Coding Approach

**Table A1. Illustration of Coding Approach for Two Publications**

| **Publications** | Lee et al. (2018) | Chatterjee et al. (2018) |
|---|---|---|
| **Summary** | The authors investigated the development and deployment of the bright Internet on a global scale, which seeks safe infrastructure, freedom of expression for innocent citizens, and the protection of privacy. To fulfill these conflicting goals, they proposed five design principles based on prevention motivation theory and analogical social norm theory. | The authors designed and implemented a wireless sensor system that helped diabetes patients to capture their daily activities. The system tracked activity data and provided motivational messages in order to change patients' behavior. |
| **Domain** | Safety, security, and privacy | Healthcare |
| **Public interest** | The authors conducted their study to develop a new infrastructure for the bright Internet in order to protect innocent citizens from cybersecurity threats. (*code: yes*) | The authors conducted their study to support the chronic disease management of diabetes patients. In addition, a university's ethics board approved the research. (*code: yes*) |
| **Informed consent** | As a conceptual study, the publication did not involve participants. (*code: no*) | The authors do not explicitly state whether they obtained informed consent, although we assume that they did since they implemented the system in two homes. (*code: no*) |
| **Privacy** | The authors sought to reduce origins of cybersecurity threats. The publication respecifies design principles to fulfill goals of origin responsibility, deliverer responsibility, identifiable anonymity, global collaboration, and privacy protection through maintaining the freedom of anonymous expression and a legitimate level of privacy protection for innocent citizens. (*code: yes*) | The authors explicitly mention that the work will benefit several new directions with respect to data privacy and applicability. (*code: yes*) |
| **Honesty and accuracy** | Not explicitly mentioned. (*code: no*) | Not explicitly mentioned. (*code: no*) |
| **Property** | Not explicitly mentioned. (*code: no*) | Not explicitly mentioned. (*code: no*) |
| **Quality of the artifact** | The authors evaluated the initial design by the justification of principles, consistency between principles, and design variables in three constructs of technologies, policies, and global collaborations. (*code: yes*) | The authors conducted a quantitative evaluation study using two case studies from home implementations with real patients of the designed system. (*code: yes*) |

## About the Authors

**Ivo Benke** is a PhD student at the Institute of Information Systems and Marketing (IISM) at the Karlsruhe Institute of Technology (KIT). He received a bachelor's and master's degree in Industrial Engineering and Management from the KIT. His research focuses on emotions, the trustworthy design of affective chatbots for enterprise collaboration, and design science research.

**Jasper Feine** is a PhD student at the Institute of Information Systems and Marketing (IISM) at the Karlsruhe Institute of Technology (KIT). He received a bachelor's and master's degree in Industrial Engineering and Management from the KIT. His research focuses on the effective and efficient design of social chatbots and design science research.

**John R. Venable** is Associate Professor and Discipline Lead, Business Information Systems in the School of Management at Curtin University, Perth, Western Australia. He has held academic positions in Information Systems and Computer Science in the USA, Denmark, New Zealand, and Australia. His research work focuses on design science research and information systems development and planning methods. His work has been published in leading international journals such as the *European Journal of Information Systems, Journal of Information Technology, Information & Management, The Information Systems Journal, Information Technology & People,* and *Communications of the Association for Information Systems*.

**Alexander Maedche** is a full professor at the Karlsruhe Institute of Technology (KIT) and head of the research group "Information Systems & Service Design" at the Institute of Information Systems and Marketing (IISM) and the Karlsruhe Service Research Institute (KSRI). His research work focuses on designing interactive intelligent systems. His work has been published in leading international journals such as the *Management Information Systems Quarterly*, *Journal of the Association for Information Systems*, *International Journal of Human-Computer Studies*, *Business Process Management Journal*, *Information & Software Technology*, *IEEE Intelligent Systems*, *SIGMOD Record*, and *AI Magazine*.

Transactions on Human – Computer Interaction