



# A Code of Digital Ethics: laying the foundation for digital ethics in a science and technology company

Sarah J. Becker<sup>1</sup> · André T. Nemat<sup>1</sup> · Simon Lucas<sup>2</sup> · René M. Heinitz<sup>1</sup> · Manfred Klevesath<sup>2</sup> · Jean Enno Charton<sup>2</sup>

Received: 19 August 2021 / Accepted: 8 December 2021 / Published online: 11 January 2022  
© The Author(s) 2022

## Abstract

The rapid and dynamic nature of digital transformation challenges companies that wish to develop and deploy novel digital technologies. Like other actors faced with this transformation, companies need to find robust ways to ethically guide their innovations and business decisions. Digital ethics has recently featured in a plethora of both practical corporate guidelines and compilations of high-level principles, but there remains a gap concerning the development of sound ethical guidance in specific business contexts. As a multinational science and technology company faced with a broad range of digital ventures and associated ethical challenges, Merck KGaA has laid the foundations for bridging this gap by developing a Code of Digital Ethics (CoDE) tailored for this context. Following a comprehensive analysis of existing digital ethics guidelines, we used a reconstructive social research approach to identify 20 relevant principles and derive a code designed as a multi-purpose tool. Versatility was prioritised by defining non-prescriptive guidelines that are open to different perspectives and thus well-suited for operationalisation for varied business purposes. We also chose a clear nested structure that highlights the relationships between five core and fifteen subsidiary principles as well as the different levels of reference—data and algorithmic systems—to which they apply. The CoDE will serve Merck KGaA and its new Digital Ethics Advisory Panel to guide ethical reflection, evaluation and decision-making across the full spectrum of digital developments encountered and undertaken by the company whilst also offering an opportunity to increase transparency for external partners, and thus trust.

**Keywords** Digital ethics · Data · Algorithms · Business ethics · Ethical principles · Operationalization of ethics

## 1 Introduction

The digital transformation has resulted in drastic change at every level of society (Floridi 2014). Enabled by the increased capabilities, miniaturisation and proliferation of computer hardware, the possibilities for collecting, connecting and analysing vast amounts of data from a wide variety of sources have grown exponentially, along with the complexity and ubiquity of the algorithmic systems employed to analyse and utilise this data.

Building upon the foundations laid by computer and information ethics since the second half of the Twentieth Century (Moor 1985; Bynum 2018; Floridi 2013), the young

field of data ethics (Floridi and Taddeo 2016) has been tackling the ethical challenges posed by both the data expansion and the accompanying algorithmic advances. The former include, for example, privacy concerns, as personal data from different areas of life can be combined and recombined to reveal ever more comprehensive and intimate insights into the habits and preferences of individuals and groups (Mittelstadt and Floridi 2016). The latter concern bias, transparency and responsibility issues arising around the development and function of algorithms (Rochel and Évêquoz 2021; Mittelstadt et al. 2016), as well as the increasing inscrutability and sophistication of machine learning approaches that have precipitated a recent flourish of enquiry into the ethics of artificial intelligence (AI) or more broadly automated decision-making-systems (Hagendorff 2020; Mökander et al. 2021b).

✉ Sarah J. Becker  
sarah.becker@transforming-healthcare.com

<sup>1</sup> Universität Witten/Herdecke, Institute for Digital Transformation in Healthcare, Düsseldorf, Germany

<sup>2</sup> Merck KGaA, Bioethics & Digital Ethics, Corporate Affairs, Darmstadt, Germany

## 2 Digital ethics challenges for science and technology companies

Whilst some principles—such as responsibility, transparency, beneficence and justice—feature almost universally in digital ethics, others apply more specifically to certain types of organisations. For example, in the case of companies seeking to develop and deploy novel digital technologies, it may be particularly important to weigh the opportunity costs of foregoing a potentially profitable new product, service or innovation against the potential risks of a mis- and overuse, especially in areas of regulatory uncertainty. Balancing these principles has significantly impacted regulatory adoption and consumer acceptance of innovations (e.g., as previously observed with genetically modified crops in the European Union) and seems to be important for developing new digital technologies too, as indicated by a recently stalling consumer uptake of digital health solutions due to mistrust in companies and lack of understanding (Safavi and Kalis 2020). Regulatory and cultural differences between countries (ÓHéigeartaigh et al. 2020) may exacerbate some of these challenges for companies operating internationally. This makes a strong case for considering ethical and societal expectations when planning business strategies (Handelsblatt Research Institute 2015; Institute for Business Ethics 2018), to allow companies to “identify and leverage new opportunities that are socially acceptable” whilst avoiding socially unacceptable courses of action that may be “rejected, even when legally unquestionable [and lowering] the opportunity costs of choices not made or options not grabbed for fear of mistakes” (Floridi et al. 2018).

Another specific challenge facing companies and similar organisations that pursue practical digital innovation are that ethical principles need to be operationalised for business purposes to provide useable tools in a variety of contexts. These range from business case analyses and panel evaluations to guidance for individual employees and external communications aiming to build trust. Any tools developed should also be universal and flexible enough to be applicable and adaptable to the organisation’s various sectors, projects and stakeholders, which may all feature different perspectives.

As a multinational science and technology company operating across healthcare, life science and performance materials, Merck KGaA (henceforth Merck) has, like other corporations, experienced digital changes and challenges to both its internal workings and the ways in which it conducts and develops its business. AI-based approaches now feature for example in drug discovery and supply chain integrity protection (Merck 2020a), and big data applications are used in human resources and—perhaps most prominently—are at

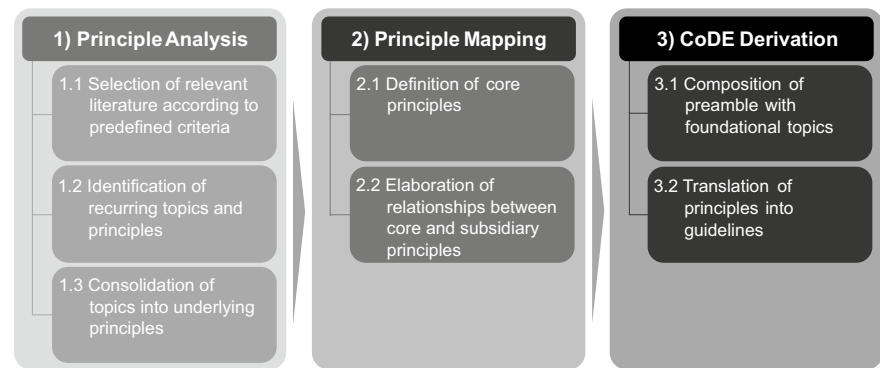
the heart of Syntropy, a partnership between Merck and Palantir Technologies Inc. specialising in integrating healthcare data from different sources to give scientists and research centres access to a collaborative technology platform for cancer research (Merck 2020b).

Due to the rapid nature of digital developments, comprehensive legal regulations have not yet been developed in all relevant areas (Gordon 2021). Given this regulatory gap, companies such as Merck that wish to operate responsibly and guided by ethically sound principles, need to be proactive about finding, communicating and implementing suitable standards for innovations that are not yet covered in detail by regulatory frameworks. Moreover, many ethical “should” questions go beyond the scope usually provided in legal regulations, which provide practitioners mainly with answers to “could” questions. However, finding answers to the ethical questions are important for building and maintaining trust in a company’s operations (Morley et al. 2021b). Merck has a strong track record of proactively seeking and implementing ethical guidance, especially in the biomedical field which has also often progressed rapidly beyond the scope of current regulation. Here, the Merck Bioethics Advisory Panel, which consists of disclosed external experts from diverse backgrounds in bioethics, philosophy and law, has provided practical guidance to the company’s healthcare and life science businesses in areas of high regulatory uncertainty. Based on the panel’s advice, Merck has published principles to guide the company’s research and development process in the areas of genome editing (Sugarman et al. 2018), stem cell research, and fertility research, in accordance with ethical standards.

Whilst the Syntropy partnership clearly belongs to the biomedical sphere, its business model, which involves enabling the integration and exchange of large amounts of highly sensitive patient data, is raising ethical challenges that are not fully covered by prevailing bioethical principles and fall outside the core expertise of the Merck Bioethics Advisory Panel. To be able to specifically address issues of digital ethics, Merck has, therefore, installed a new expert committee, the Merck Digital Ethics Advisory Panel, which operates as a separate yet interconnected advisory body (Merck 2021). Additionally, Palantir has a long history (since 2012) of drawing upon the advice of its Palantir Council of Advisors on Privacy and Civil Liberties (PCAP) as a group of independent experts in privacy law, policy, and ethics to understand and address the complex issues encountered in the course of its work.

## 3 Seeking a principle-based Code of Digital Ethics

To develop a scientifically valid foundation to guide ethical decision-making throughout the company, we sought to develop a Code of Digital Ethics for Merck that is both

**Fig. 1** Course of the project

ethically rigorous and suited for operationalisation across the full scope of digital ethics challenges arising at Merck. Such an internal guideline should function as a common ground-work from which to derive suitable strategies and procedures for different business sectors and individual employees to enable them to navigate areas of high regulatory uncertainty. It should also provide a clear structure for the assessment of ethical questions by the new Merck Digital Ethics Advisory Panel and other decision-makers, and help to build trust with customers and business partners.

Due to the recent rapid increase and the pervasive nature of digital developments, many different organisations have started to develop practical ethical guidance in this sphere. This has led to a plethora of guidelines, ranging from individual companies' big data or AI codices (Deutsche Telekom 2018; SAP 2018; Telefónica 2018), to recommendations for specific societal sectors like the healthcare system (Deutscher Ethikrat 2017; Mittelstadt and Floridi 2016), to high-level frameworks that target society as a whole at national (Die Bundesregierung 2018; Datenethikkommission 2019; UK Government 2018) or even international (EGE 2018; ICDPPC 2018; CEPEJ 2018) levels.

Recognising that “the sheer volume” of individual principles proposed in such a frontier situation “threatens to overwhelm and confuse” (Floridi and Cowls 2019), there have been several attempts to compare this abundance of heterogeneous and yet overlapping documents, to extract common core principles and recurring topics from them and to synthesise these into unified frameworks of digital or AI ethics (Floridi et al. 2018; Jobin et al. 2019; Fjeld et al. 2020; Hagendorff 2020). Such meta-analyses offer valuable orientation but cannot necessarily provide the specific perspective needed for individual organisations with their unique digital challenges, needs and values, or for translating high-level principles into robust practice (Mittelstadt 2019; Floridi 2019; Hickok 2020; Blackman 2020). We, therefore, decided to pursue a tailor-made approach that combines an analysis of the most relevant data and AI ethics guidelines with the extraction of key principles suitable for Merck and their translation into guidelines that can serve as a foundation

for the diverse operationalisation requirements within the organisation.

## 4 Crafting the CoDE

Creating the CoDE took three steps, (1) principle analysis, (2) principle mapping and (3) CoDE derivation (see Fig. 1). The first step involved the selection of relevant literature according to predefined criteria and the identification and consolidation of the topics and ethical principles contained in these documents, using a reconstructive social research approach (Vogd 2009; Özbilgin 2006; Bohnsack 1999). During the second step of the project, we defined the most important core principles and mapped the remaining principles to them in subsidiary groups. All principles were then, in the third and final step, translated into the guidelines that constitute the CoDE.

### 4.1 Principle analysis

In our analysis, we considered recent<sup>1</sup> literature that fulfilled three criteria: First, all documents included had to refer to the ethical handling of data and/or AI, machine learning or other aspects of algorithmic systems.<sup>2</sup> Secondly, each document had to contain normative statements about data and/or algorithmic systems. Thirdly, the editor had to be from Europe. Across Europe, there has been especially strong commitment to stringent scrutiny of the ethical challenges

<sup>1</sup> Considering the recent emergence of the field of data ethics (Floridi and Taddeo 2016), we chose to limit our search to the five years prior to the commencement of this project (2015–2019).

<sup>2</sup> The terms “artificial intelligence” (AI) and “machine learning” are often used interchangeably or to refer more broadly to advanced algorithmic systems, and a majority of guidelines explicitly uses “AI” as an umbrella term. However, as the German Data Ethics Commission (DEK) has pointed out, “[AI] [...] represents only a special manifestation and thus a subset of algorithmic systems” (DEK 2019, p. 34). The broader term “algorithmic systems” appears therefore better suited to comprise the full variety of algorithmic applications that may pose ethical challenges in the context of this analysis.

around digital developments, at the level of both data and algorithmic systems. Attempts to devise appropriate, innovative and trust-building ways of tackling these have been developed by individual countries (e.g. Agenzia per l'Italia Digitale 2018; Datenethikkommission 2019; Deutscher Ethikrat 2017; Ekspertgruppe om dataetik 2018; Schweizerische Eidgenossenschaft 2020; UK Government 2018; Villani 2018) as well as at EU level (e.g. EGE 2018; EDPS Ethics Advisory Group 2018; AI HLEG 2019; European Commission 2020), with pioneer projects like the EU General Data Protection Regulation and the recently proposed first-ever legal framework on artificial intelligence, the EU Artificial Intelligence Act, also likely to affect and inspire regulatory development beyond Europe (EU Commission 2021; Mökander et al. 2021a). Merck has a long history of operating within a European values framework and a clear commitment to upholding these values. A clear focus on European guidelines was, therefore, expected to provide the best match for the perspective and requirements of an internationally operating EU-based science and technology company like Merck, with its demonstrated commitment to developing and adhering to rigorous ethical standards in business contexts.

As a baseline, we used the compilations of recommendations, guidelines, frameworks and sets of principles on data and algorithm ethics (henceforth referred to as “guideline documents”) by the Harvard Internet Institute (later published by Fjeld et al. 2020) and Algorithm Watch’s crowd-sourced AI Ethics Guidelines Global Inventory (<https://inventory.algorithmwatch.org>). This yielded 32 and 83 publications, respectively, and was complemented by internet searches for the keywords “digital ethics”, “data ethics” and “corporate digital responsibility”, yielding two additional publications. After eliminating 15 duplicates from this collection of 117 documents, we excluded a further 60 guidelines that did not originate from European sources, leaving a total of 42 European guideline documents (see Online Resource 1).

These documents were subjected to the reconstructive social research approach in four steps: First, we examined the documents for explicitly named principles, recording all text passages with such explicit mentions.<sup>3</sup> We also included certain recurring topics that could not be directly identified as principles but were related to them and considered relevant due to their frequency. This resulted in an initial list of 29 recurring principles and topics (see Fig. 2) that either represented, or bore a close connection to, ethical principles. As a second step, we formulated preliminary definitions

of the topics thus uncovered, guided by the scope of the recorded passages. In a third step, we used these definitions to examine the documents a second time, to uncover and extract passages that mentioned the corresponding topics implicitly. Implicit mentions were discussed in the team and only recorded as referrals to the previously agreed topics if a match was agreed unanimously.

We then reduced the list of 29 topics to 20 central digital ethics principles in the fourth and final step (see Fig. 3). This reduction process aimed to remove some ambiguities and overlaps discovered upon further scrutiny of the preliminary definitions and the newly recorded implicit text passages: Some topics could not be directly identified as ethical principles but might be assigned to one (for example “data and AI usage” and “data collection” to the principle of proportionality). Others were not suitable for such assignment and instead dropped from further analysis (“research practices” and “societal ethics”). The principles “equality” and “fairness” were combined under the principle of “equality” since the latter, although lexically subordinate to “fairness”, constitutes a more specific and thus easier to operationalise principle with particular relevance for digital innovations and businesses. Finally, some topics were of such a basic nature or of such overarching importance that it appeared more appropriate to introduce them in a foundational preamble to the main principles. The topic “respect law”, for example, was moved to the preamble of the CoDE since in a business context, compliance with the applicable law is considered a fundamental prerequisite that needs to be met before further ethical principles are discussed. And while a specific aspect of the topic “governance mechanisms”—reliability—found translation into the final set of principles, the topic’s broad scope also warranted its inclusion in the preamble. Similarly, the topics “human rights”, “ethical design” and “trustworthiness” were also placed in the preamble.

## 4.2 Principle mapping

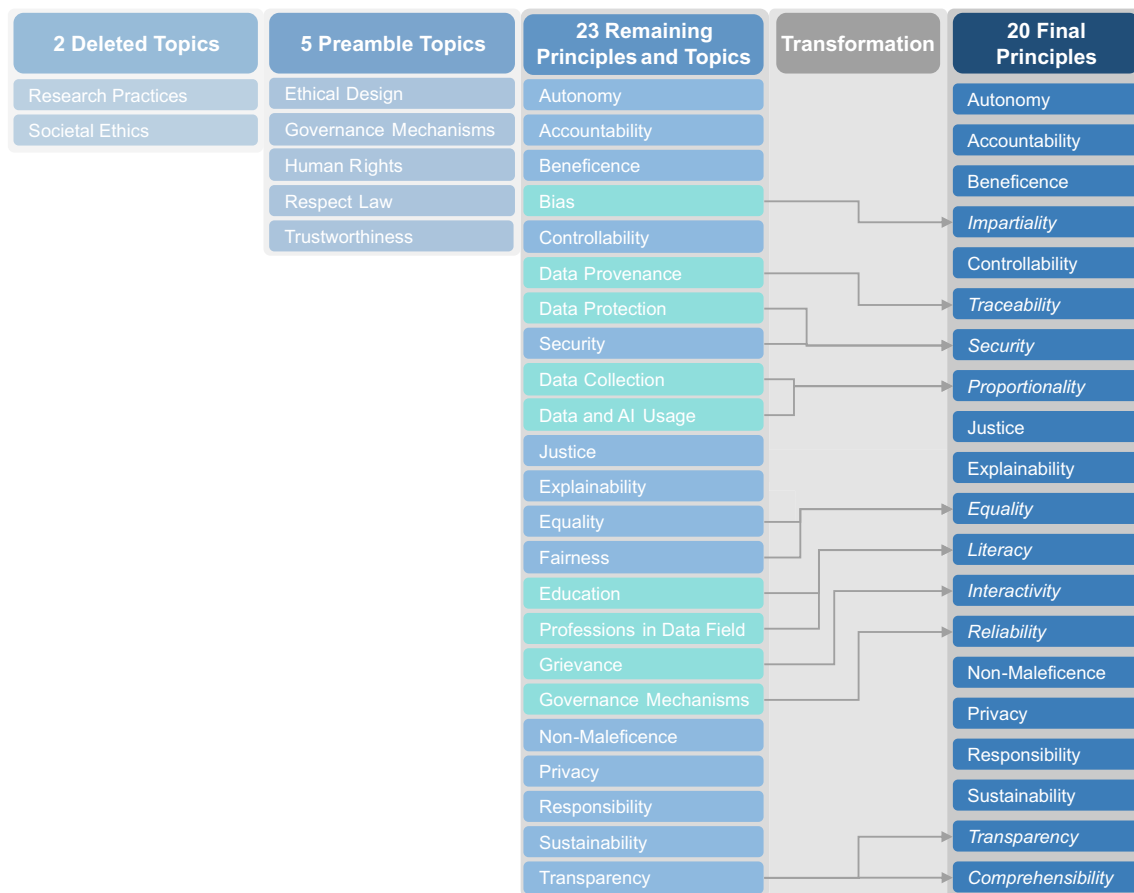
We next strove to facilitate understanding and navigation of the identified principles, as well as their suitability for later operationalisation, by refining their preliminary definitions, elucidating their relationships to each other, and structuring them further in a hierarchical model with core principles that have the remaining principles mapped to them as subsidiaries.

The final definitions of the 20 principles were derived in an iterative process by further analysing the text passages from the 42 guideline documents that had previously been recorded as explicit or implicit referrals to the principles. Care was taken to ensure that no nuances were lost where terms from the initial list of 29 topics and principles had been combined or transformed. The full definitions of all

<sup>3</sup> The following statement, for example, explicitly refers to the principle of transparency: “Big Data applications should be transparent so that the persons concerned can see which of their personal data are processed in which way.” (Bitkom 2015, p. 82).

**Fig. 2** Frequency distribution of the initially identified 15 principles and 14 further topics, ordered from highest number of mentions to the lowest

Principles and Topics Initially Identified			
15 Principles		14 Further Topics	
	Numbers of Guidelines that Mention this		Numbers of Guidelines that Mention this
Transparency	39	Bias	24
Beneficence	25	Governance Mechanisms	21
Explainability	23	Data Protection	21
Autonomy	23	Ethical Design	20
Responsibility	23	Education	18
Security	21	Respect Law	18
Controllability	17	Data Collection	16
Equality	16	Human Rights	15
Accountability	16	Data and AI Usage	9
Privacy	15	Data Provenance	7
Fairness	14	Research Practices	6
Non-Maleficence	13	Societal Ethics	6
Sustainability	9	Grievance	5
Trustworthiness	6	Professions in the Data Field	4
Justice	5		



**Fig. 3** The process of transforming the initial 15 principles and 14 further topics to 5 preamble topics and 20 final principles



20 principles are listed in the appendix of Merck's Code of Digital Ethics.

Five particularly far-reaching principles were selected as core principles as they represent the essence of larger thematic clusters that are arguably suitable as a foundation of digital ethics: autonomy, non-maleficence, beneficence, justice and transparency. Of these, the first four appear as classical principles of biomedical ethics and have been championed in this field since the 1970s (Beauchamp and Childress 2019). Their affinity to digital ethics has previously been highlighted by Floridi et al. (2018), who state that “[o]f all areas of applied ethics, bioethics is the one that most closely resembles digital ethics in dealing ecologically with new forms of agents, patients, and environments”. Although limitations of their applicability to digital ethics have also been pointed out, referring to the comparative recency and heterogeneity of digital developments (Mittelstadt 2019), the four principles of biomedical ethics feature heavily in many digital ethics guidelines, albeit adjusted to the specific digital challenges discussed (e.g. Floridi et al. 2018; Jobin et al. 2019).

For the purpose of developing Merck's CoDE, choosing an integrated methodology around Beauchamp and Childress' principled ethics seemed plausible because the four-principle-approach has proven convenient to analysing a broad spectrum of ethical dilemmas in medicine and bioscience and has become one of the most influential frameworks in applied ethics generally over the past decades. Moreover, Merck has been successfully utilizing principle-based thinking and ethical analysis to guide business decisions in areas of regulatory uncertainty related to its healthcare and life science operations, e.g. by implementing ethical guidance of the Merck Bioethics Advisory Panel.

However, the four principles clearly cannot cover all ethical issues raised by data use and algorithmic systems. Therefore, we selected transparency as a fifth core principle. Transparency had by far the most mentions in the guidelines analysed (39 of 42, see Fig. 2) and works well as the core of a final cluster to structure related principles and to complement the clusters around the biomedical ethics principles. Although transparency has traditionally been regarded as an integral part of the autonomy principle in the field of bioethics, its reach beyond the individual patient perspective and its outstanding relevance for ethics in data and AI warrant its appreciation as a distinct principle in digital ethics. For example, while the asymmetry of information between a patient and a physician is an important factor influencing the autonomy principle in medical ethics, it does not fully capture the importance of transparency with regards to complex data ecosystems and algorithmic systems, which for the most part operate invisibly and are unintelligible to the vast majority of observers, yet affect far more people. The

importance of transparency is reflected by how much consumer confidence depends on comprehension of, and trust in, digital offerings, including how and by whom data and results will be used (Safavi and Kalis 2020). Thus, we arrive at similar conclusions as Floridi et al. 2018, who chose explicability as a fifth key principle, which shares many features with the transparency principle as defined by us.

Indeed, these shared features can serve as an example to illustrate how we mapped the remaining 15 principles uncovered in our analysis to the five core principles. We sought to assign additional principles to each core principle to highlight and address its key aspects and characteristics. Assignment was initially undertaken independently by two team members, already resulting in a high level of agreement. Any remaining discrepancies were discussed by the entire team and resolved consensually. Initial attempts to map the principle explainability (which is synonymous with explicability) to one of our core principles resulted in uncertainty whether autonomy or transparency would provide a better fit. According to our final definition, “In the context of digital solutions, explainability means that users can understand how the results of algorithm-based processes are achieved [,] that users know the reason for their decision wherever they are directly or indirectly affected by automated decisions [and that] users receive sufficient information about the models used and the architecture of the algorithmic system”. This definition fits well with the core principle of transparency because explainability serves to increase transparency, but it also increases the autonomy of users by enabling them to make better-informed evaluations of the algorithms affecting them. We ultimately chose to assign explainability to autonomy, due to our definition's focus on strengthening users' decision-making abilities, but the strong connection to transparency obviously remains. The final mapping of the CoDE's principles is shown in Fig. 4.

Within each cluster of principles, descriptions follow a clear pattern. Each core principle (see Fig. 5) is introduced by (1) listing the other principles assigned to it; followed by (2) a definition of the core principle; (3) risks that may result from not respecting it; and (4) a description of the relationship between the core principle and the principles assigned to it. Subsidiary principles (see Fig. 6) are introduced by (1) stating the reference level(s) addressed by the principle—data, algorithmic systems or both—; followed by (2) a definition of the principle in the context of digital solutions; (3) the risks involved if it is not taken into account by organisations; (4) examples of solutions to help organisations counteract these risks and to preserve this principle; and (5) an explanation of why the principle supports the associated core principle.

Core Principles	Subsidiary Principles
Justice	Impartiality
	Equality
	Proportionality
Autonomy	Explainability
	Privacy
	Literacy
Beneficence	Sustainability
	Security
	Responsibility
Non-Maleficence	Reliability
	Controllability
	Accountability
Transparency	Traceability
	Interactivity
	Comprehensibility

Fig. 4 Final structure of the principles

### 4.3 CoDE derivation

In the third and final step of this project, we sought to transform the principles into guidelines for ethical decision-making around digital applications and to present them in a CoDE that clearly expresses how Merck intends to handle its data and algorithmic systems and any associated services and products. The target audience of the CoDE should be everyone who develops, operates, uses or decides on the use of Merck’s digital offerings. Scrutiny of the existing guideline documents analysed by us revealed considerable heterogeneity. The level of reference, underlying principles and recommended measures or organisational elements were frequently not easily identifiable. Indeed, the translation of “lofty” principles (Mittelstadt 2019) into robust and operationalisable guidelines that can enable truly informed decisions on the ground has been identified as a key challenge for digital ethics (Floridi 2019; Blackman 2020; Morley et al. 2021a). We decided to tackle this challenge with a strong dual focus on clarity and operationalisability.

Towards this end, we first took advantage of the clear structure previously established for the 20 principles by mirroring this structure in the CoDE guidelines. By firmly grounding our guidelines on relevant principles well established in the current digital ethics literature and elucidating both their meaning and their relationships to each other, we strove to arrive at guidelines that are relatable and understandable from the various perspectives found inside a company such as Merck whilst also maintaining a strong connection to the academic discourse on digital ethics. We

hope that this approach offers a more rigorous foundation for developing a strong practice of digital ethics in a company than a code of conduct compiled with less reflection and reference about the origin, meaning, and relationship of the principles underlying its guidelines. Moreover, this clear foundation offers Merck the opportunity to learn from, and respond to, new developments in both the academic discourse on digital ethics and the company’s own efforts to operationalise digital ethics throughout its practices, based on these guidelines.

We then selected three criteria to ensure operationalisability for business purposes and suitable to Merck’s diverse activities and needs in the digital sphere. Further analysis of existing digital ethics guidelines (ITI 2017) as well as consultation of best practice recommendations from the established field of corporate social responsibility (FMLS 2020) convinced us that for successful operationalisation, guidelines should (1) avoid oversimplified closed normative statements (“dos and don’ts”), (2) instead strive to be universal (i.e. not limited to specific applications, technologies, ventures or measures) and (3) be open to multiple perspectives (i.e. not limited to addressing specific stakeholders). At the same time, the guidelines should be written *from* the perspective of Merck as a private-sector company with digital offerings and follow a consistent structure that clearly expresses the commitment of the company to each principle discussed and elaborates on specific intentions for implementing these commitments.

Figure 7 illustrates the result of employing these criteria, using as an example the core guideline autonomy and its subsidiary guidelines. Each core guideline consists of a central statement that shows what Merck wants to achieve, followed by an elaboration of the content of the statement, the reference level(s) addressed, and an explanation of which subsidiary guidelines, derived from which principles, support the organisation in following this specific core guideline. The subsidiary guidelines follow the same structure, and the contents of all statements are derived from the principles on which these guidelines are based.

The full CoDE, including the preamble which explains the structure of the guidelines and introduces the CoDE’s aim, building on the five fundamental topics identified in the analysis step (4.1), can be found in Online Resource 2.

## 5 Discussion

Our goal was to identify principles suitable for the development of a Code of Digital Ethics for Merck and to transform these into clear guidelines that lay the foundation for operationalising ethical reflection, evaluation and decision-making across the full spectrum of digital developments encountered and undertaken by the company. The CoDE

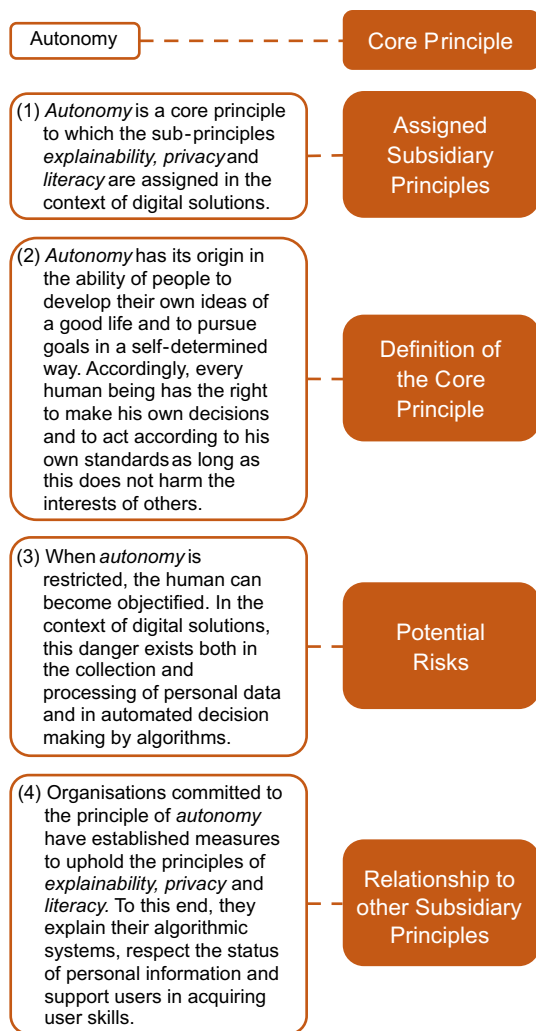


Fig. 5 Structure of the core principle autonomy

should furthermore serve to clearly communicate to customers, business partners, and other external stakeholders the principles that underlie Merck's decision-making around its digital offerings, policies and initiatives.

A simple compilation of high-level principles would likely have been too unwieldy and unstructured for these purposes, which is why we chose to further cluster and hierarchically order the individual principles. Even a cursory glance at other guideline documents and meta-analyses of such guideline documents shows that there are many ways to define, label and structure the principles, values and recurring topics at the heart of digital ethics. Our final five-by-four matrix of five core and 15 subsidiary principles resembles the structuring proposed by Floridi et al. 2018 but works with a somewhat different set and mapping of principles tailored to the needs of Merck's businesses. Others have chosen similar but longer (Jobin et al. 2019) or shorter (AI HLEG 2019) lists of core principles or hierarchical structures that

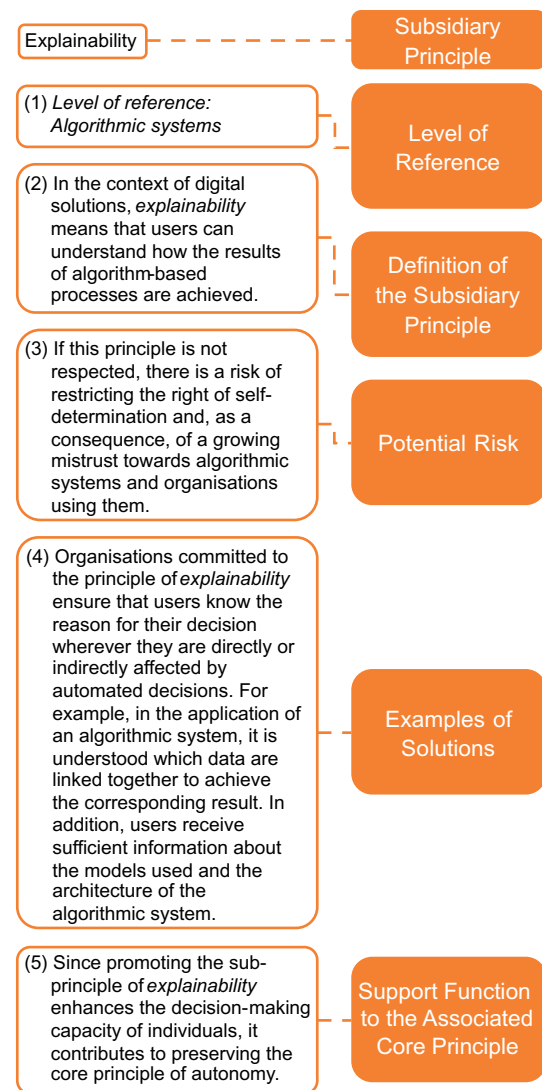


Fig. 6 Structure of the subsidiary principle explainability

appear quite different (Fjeld et al. 2020; Zeng et al. 2018), or they have foregone attempts at hierarchical structuring entirely (Hagendorff 2020). We believe that the nested structure chosen by us, with its clear overview at the top level of core guidelines and the details and relational information added by the subsidiary guidelines and levels of reference is well suited for a code that aims to be broadly accessible for diverse stakeholders and yet sufficiently nuanced to guide ethical decision-making in complex business contexts.

The very heterogeneity of digital developments has been proposed as an obstacle to finding common principles or rules that will be applicable across all potential ethical facets of the digital transformation and has highlighted the importance of devising and combining tailor-made solutions for different stakeholders, to enable them to share responsibility (Mittelstadt 2019; Blackman 2020; Deutscher Ethikrat



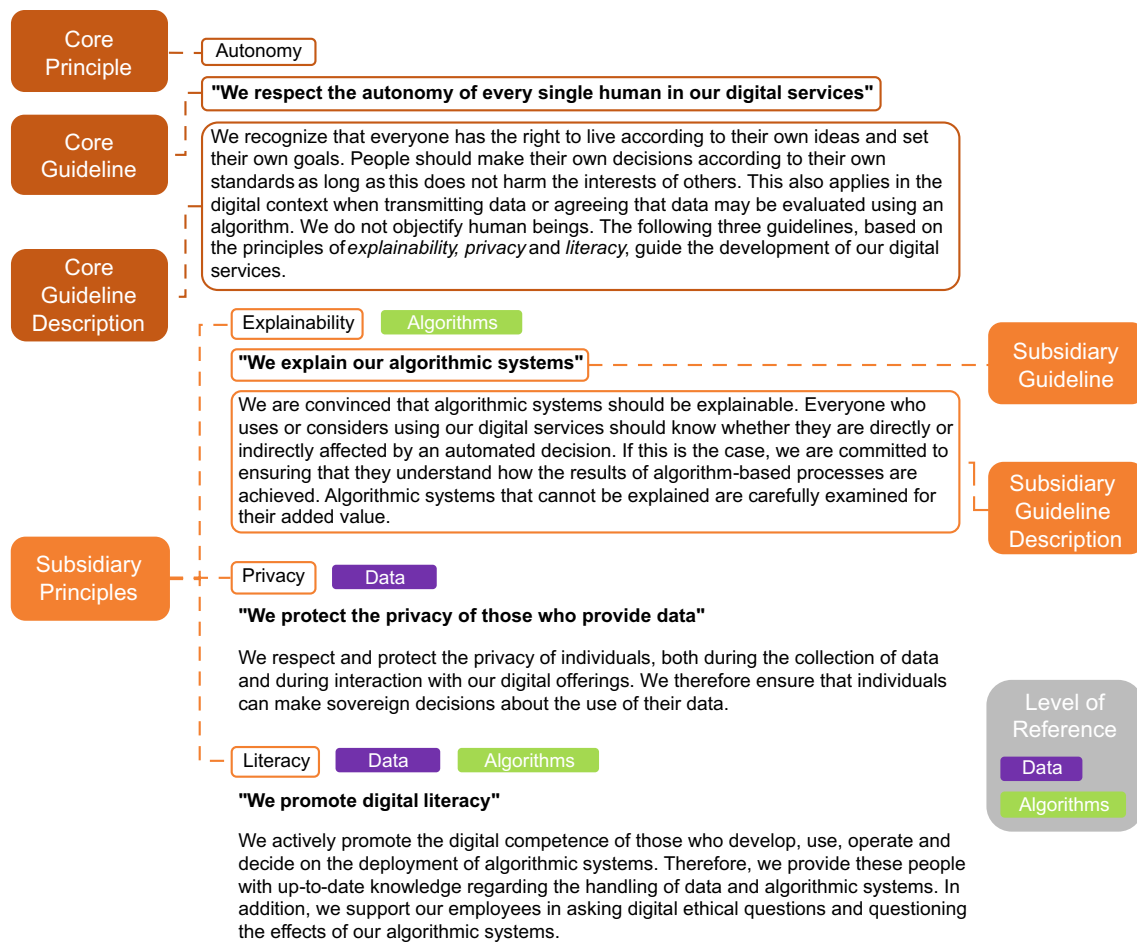


Fig. 7 Final structure of the guidelines

2017). This makes it all the more important to be clear about the process by which certain principles are defined or given priority, as well as the relationships between principles and the reference levels—data and/or algorithmic systems—addressed by them. Such clarity of structure, which we have strived to achieve in the CoDE, helps different actors and stakeholders to understand each other and to find common ground whilst recognising important differences in their perspectives and needs. As a multinational company operating across several sectors and comprising different businesses and partnerships, Merck relies on creating such understanding and common ground between different parties.

By selecting a scientific methodology for creating the CoDE and fully documenting the process we also wish to offer a transparent baseline for other companies that may face similar challenges to obtain ethical guidance for digital innovation in a complex business environment. At the same time, our approach creates the opportunity to participate in the academic discourse on digital ethics and to learn from any feedback and further developments in the field that may inform the ongoing operationalisation and refinement of our

own ethical guidance in the digital sphere. We consider such engagement essential to appropriately continue and further develop a process of reflection on, and implementation of, digital ethics in a multifaceted science and technology company such as Merck.

Given the variety of viewpoints, ventures and challenges encountered in the context of digital innovation we designed the CoDE as a multi-purpose tool—a robust framework for consistent ethical analysis and decision making that can be adapted to diverse business cases and against different national, international and culturally varied backgrounds. The three criteria applied during the formulation of the CoDE's guidelines—avoidance of closed normative statements; preference for universality; and openness to multiple perspectives—have yielded a document that we hope will be easy to operationalise in such varied situations. Towards this end, Merck's new Digital Ethics Advisory Panel is currently utilising the CoDE in a first test case for operationalisation to tackle ethical questions around patient data sharing that have arisen at Syntropy, with promising initial results. Other operationalisation efforts currently under development include the development of a tool

that will enable standardized use of the CoDE in evaluating the criticality of data analysis projects.

We thus believe that the CoDE constitutes a suitable tool to accompany digital innovation and associated business decisions by sound ethical reasoning. Actively employing this tool, and clearly explaining the selection and meaning of the principles that guide digital offerings, can serve to create trust, transparency and, perhaps, competitive advantage.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00146-021-01376-w>.

**Author contributions** Not applicable.

**Funding** No funding was received to assist with the preparation of this manuscript.

**Availability of data and material** A list of the guidelines examined is attached to the supplementary information (Online Resource 1). The CoDE is attached to the supplementary information (Online Resource 2).

**Code availability** Not applicable.

## Declarations

**Conflict of interest** S.B., A.N., R.H. are employed by idigiT GmbH. S.L., M.K. and J.E.C. are employed by Merck KGaA. idigiT GmbH has been commissioned by Merck KGaA to consult the company on digital ethics. In her role as advisor to Merck KGaA, Darmstadt, S.B. has been a compensated private consultant. For this editorial, the authors acted as independent scholars and did not receive payment for their contributions. Editorial assistance, including writing, was provided by Nora Schultz and was funded by Merck KGaA.

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agenzia per l'Italia Digitale (2018) AI White paper. <https://ai-white-paper.readthedocs.io/en/latest/doc/raccomandazioni.html>. Accessed 9 Mar 2021
- AI HLEG-Independent High-Level Expert Group on AI (2019) Ethics guidelines for trustworthy AI. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Accessed 9 Mar 2021
- Beauchamp T, Childress J (2019) Principles of biomedical ethics: marking its fortieth anniversary. *Am J Bioethics* 19(11):9–12. <https://doi.org/10.1080/15265161.2019.1665402>
- Blackman R (2020) A practical guide to building ethical AI. *Harv Bus Rev*. <https://hbr.org/2020/10/a-practical-guide-to-building-ethical-ai>. Accessed 29 Nov 2020
- Bohnsack R (1999) *Rekonstruktive Sozialforschung Einführung in Methodologie und Praxis qualitativer Forschung*. Leske + Budrich, Opladen
- Bynum T (2018) Computer and information ethics. The Stanford Encyclopedia of Philosophy Archive (Summer 2018 Edition). <https://plato.stanford.edu/archives/sum2018/entries/ethics-computer/>. Accessed 26 Nov 2020
- CEPEJ-European Commission for the Efficiency of Justice (2018) European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. <https://rm.coe.int/ethical-ai-charter-en-for-publication-4-december-2018/16808f699c>. Accessed 9 Mar 2021
- Datenethikkommission (2019) Gutachten der Datenethikkommission. [https://datenethikkommission.de/wp-content/uploads/191128\\_DEK\\_Gutachten\\_bf\\_b.pdf](https://datenethikkommission.de/wp-content/uploads/191128_DEK_Gutachten_bf_b.pdf). Accessed 9 Mar 2021
- Deutsche Telekom AG (2018) KI-Leitlinien. <https://www.telekom.com/resource/blob/532444/87e1e54df08cce6f4483985bd25250b6/dl-180710-ki-leitlinien-data.pdf>. Accessed 5 Mar 2021
- Deutscher Ethikrat (2017) Big Data und Gesundheit—Datensouveränität als informationelle Freiheitsgestaltung. <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-big-data-und-gesundheit.pdf>. Accessed 9 Mar 2021
- Die Bundesregierung (2018) Strategie Künstliche Intelligenz der Bundesregierung. [https://www.bmbf.de/files/Nationale\\_KI-Strategie.pdf](https://www.bmbf.de/files/Nationale_KI-Strategie.pdf). Accessed 9 Mar 2021
- EDPS Ethics Advisory Group (2018) Towards a digital ethics. [https://edps.europa.eu/sites/edp/files/publication/18-01-25\\_eag\\_report\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf). Accessed 9 Mar 2021
- Ekspertgruppe om dataetik (2018) Data for the benefit of the people. Recommendations from the Danish Expert Group on Data Ethics. <https://em.dk/media/12190/dataethics-v2.pdf>. Accessed 9 Mar 2021
- European Commission (2020) Whiter paper on artificial intelligence—a European approach to excellence and trust. [https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf). Accessed 9 Mar 2021
- European Commission (2021) Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206>
- European Group on Ethics in Science and New Technologies (EGE) (2018) Statement on artificial intelligence, robotics and 'Autonomous' systems. [http://ec.europa.eu/research/ege/pdf/egi\\_ai\\_statement\\_2018.pdf](http://ec.europa.eu/research/ege/pdf/egi_ai_statement_2018.pdf). Accessed 9 Mar 2021
- Federal Ministry of Labour and Social Affairs (FMLS) (2020) International frameworks: guide for global business. <https://www.csr-in-deutschland.de/EN/What-is-CSR/Background/International-Frameworks/international-frameworks-article.html>. Accessed 16 Dec 2020
- Fjeld J, Achten N, Hilligoss H, Nagy A, Srikumar M (2020) Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Center Res Publ No*. <https://doi.org/10.2139/ssrn.3518482>
- Floridi L (2013) *The ethics of information*. Oxford University Press, Oxford

- Floridi L (2014) *The fourth revolution: how the infosphere is reshaping human reality*. Oxford University Press, Oxford
- Floridi L (2019) Translating principles into practices of digital ethics: five risks of being unethical. *Philos Technol* 32(2):185–193. <https://doi.org/10.1007/s13347-019-00354-x>
- Floridi L, Cowls J (2019) A unified framework of five principles for AI in society. *Harvard Data Sci Rev* 1(1):1–15. <https://doi.org/10.1162/99608f92.8cd550d1>
- Floridi L, Taddeo M (2016) What is data ethics? *Philos Trans R Soc A* 374(2083):1–5. <https://doi.org/10.1098/rsta.2016.0360>
- Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B, Valcke P, Vayena E (2018) AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds Mach* 28(4):689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gordon JS (2021) AI and law: ethical, legal, and socio-political implications. *AI Soc* 36:403–404. <https://doi.org/10.1007/s00146-021-01194-0>
- Hagendorff T (2020) The ethics of AI ethics: an evaluation of guidelines. *Minds Mach* 30(1):99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Handelsblatt Research Institute (2015) *Datenschutz und big data—Ein Leitfaden für Unternehmen*. [https://www.umweltdialog.de/de-wAssets/docs/2014-Dokumente-zu-Artikeln/leitfaden\\_unternehmen.pdf](https://www.umweltdialog.de/de-wAssets/docs/2014-Dokumente-zu-Artikeln/leitfaden_unternehmen.pdf). Accessed 5 Mar 2021
- Hickok M (2020) Lessons learned from AI ethics principles for future actions. *AI Ethics* 1:41–47. <https://doi.org/10.1007/s43681-020-00008-1>
- ICDPPC—International Conference of Data Protection and Privacy Commissioners (2018) Declaration on ethics and data protection in artificial intelligence. [https://edps.europa.eu/sites/default/files/publication/icdppc-40th\\_ai-declaration\\_adopted\\_en\\_0.pdf](https://edps.europa.eu/sites/default/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf). Accessed 9 Mar 2021
- Institute for Business Ethics (2018) Business ethics and artificial intelligence. <https://www.ibe.org.uk/uploads/assets/5f167681-e05f-4fae-ae1bef7699625a0d/ibebriefing58businessethicsandartificialintelligence.pdf>. Accessed 5 Mar 2021
- ITI—Information Technology Industry Council (2017) AI policy principles. <https://www.itic.org/public-policy/ITIAIPolicyPrinciple.sFINAL.pdf>. Accessed 9 Mar 2021
- Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. *Nat Mach Intell* 1(9):389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Merck KGaA (2020a) Annual report 2019. [https://www.merckgroup.com/en/annualreport/2019/downloads/merck\\_annual\\_report\\_2019.pdf](https://www.merckgroup.com/en/annualreport/2019/downloads/merck_annual_report_2019.pdf). Accessed 11 Dec 2020.
- Merck KGaA (2020b) Corporate responsibility report 2019. [https://www.merckgroup.com/en/cr-report/2019/servicepages/downloads/files/entire\\_merck\\_crr19.pdf](https://www.merckgroup.com/en/cr-report/2019/servicepages/downloads/files/entire_merck_crr19.pdf). Accessed 11 Dec 2020
- Merck KGaA (2021) Merck announces formation of merck digital ethics advisory panel. News Release, 8 Jan 2021. <https://www.merckgroup.com/press-releases/2021/jan/en/Digital-Ethics-Advisory-Panel-EN.pdf>. Accessed 19 Feb 2021
- Mittelstadt BD (2019) Principles alone cannot guarantee ethical AI. *Nat Mach Intell* 1(11):501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- Mittelstadt BD, Floridi L (2016) The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics* 22(2):303–341. <https://doi.org/10.1007/s11948-015-9652-2>
- Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The ethics of algorithms: mapping the debate. *Big Data Soc* 3(2):1–21. <https://doi.org/10.1177/2053951716679679>
- Mökander J, Axente M, Casolari F, Floridi L (2021a) Conformity assessments and post-market monitoring: a guide to the role of auditing in the proposed European AI regulation. *Minds Mach*. <https://doi.org/10.1007/s11023-021-09577-4>
- Mökander J, Morley J, Taddeo M, Floridi L (2021b) Ethics-based auditing of automated decision-making systems: nature, scope, and limitations. *Sci Eng Ethics* 27:44. <https://doi.org/10.1007/s11948-021-00319-4>
- Moor JH (1985) What is computer ethics? *Metaphilosophy* 16(4):266–275. <https://doi.org/10.1111/j.1467-9973.1985.tb00173.x>
- Morley J, Elhalal A, Garcia F, Kinsey L, Mökander J, Floridi L (2021a) Ethics as a service: a pragmatic operationalisation of AI ethics. *Minds Mach* 31(2):239–256. <https://doi.org/10.1007/s11023-021-09563-w>
- Morley J, Kinsey L, Elhalal A, Garcia F, Ziosi M, Floridi L (2021b) Operationalising AI ethics: barriers, enablers and next steps. *AI Soc*. <https://doi.org/10.1007/s00146-021-01308-8>
- ÓhÉigeartaigh SS, Whittlestone J, Liu Y, Zeng Y, Liu Z (2020) Overcoming barriers to cross-cultural cooperation in AI ethics and governance. *Philos Technol* 33(4):571–593. <https://doi.org/10.1007/s13347-020-00402-x>
- Özbilgin MF (2006) Relational methods in organization studies: a review of the field. In: Kyriakidou O, Özbilgin MF (eds) *Relational perspective in organizational studies*. Edward Elgar Publishing Limited, Cheltenham, pp 224–264
- Rochel J, Évêquoz F (2021) Getting into the engine room: a blueprint to investigate the shadowy steps of AI ethics. *AI Soc* 36:609–622. <https://doi.org/10.1007/s00146-020-01069-w>
- Safavi K, Kalis B (2020) How can leader make recent digital health gains last? Re-examining the Accenture 2020 Digital Health Consumer Survey. US Findings. Accenture [https://www.accenture.com/\\_acnmedia/PDF-130/Accenture-2020-Digital-Health-Consumer-Survey-US.pdf](https://www.accenture.com/_acnmedia/PDF-130/Accenture-2020-Digital-Health-Consumer-Survey-US.pdf). Accessed 11 Dec 2020
- SAP SE (2018) SAPs guiding principles for artificial intelligence. <https://www.sap.com/docs/download/2018/09/940c6047-1c7d-0010-87a3-c30de2ffd8ff.pdf>. Accessed 5 Mar 2021
- Schweizerische Eidgenossenschaft (2020) Künstliche Intelligenz: Wie geht der Bund damit um? <https://www.sbf.admin.ch/sbfi/de/home/dienstleistungen/publikationen/publikationsdatenbank/s-n-2020-6/s-n-2020-6f.html>. Accessed 9 Mar 2021
- Sugarman J, Shivakumar S, Rook M, Loring JF, Rehmann-Sutter C, Taupitz J, Rheinhard-Rupp J, Hildemann S (2018) Ethical considerations in the manufacture, sale, and distribution of genome editing technologies. *Am J Bioethics* 18(8):3–6. <https://doi.org/10.1080/15265161.2018.1489653>
- Telefónica SA (2018) AI principles of telefónica. <https://www.telefonica.com/documents/1258915/0/Infographics-AI/>. Accessed 5 Mar 2021
- UK Government (2018) Data ethics framework. <https://www.gov.uk/government/publications/data-ethics-framework/dataethics-framework>. Accessed 9 Mar 2021
- Villani C (2018) For a meaningful artificial intelligence towards a French and European strategy. French Parliament. [https://www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf). Accessed 9 Mar 2021
- Vogd W (2009) *Rekonstruktive Organisationsforschung: Qualitative Methodologie und theoretische Integration—eine Einführung*. Verlag Barbara Budrich, Opladen
- Zeng Y, Lu E, Huangfu C (2018) Linking artificial intelligence principles. arXiv: 1812.04814

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.