

Data Ethics in Practice: Challenges and Opportunities for a Data Ethics Policy Function in the Public Sector

**Natalia Domagala
Cabinet Office**

The overarching aim of data ethics is to promote responsible and sustainable use of data and its products for the benefit of people and society. Through an analysis of various approaches to data and AI ethics in the public sector, this paper aims to identify key goals and challenges for a data ethics policy function and provide recommendations. The paper connects the practical experience of data ethics professionals and the emerging theory through combining a literature review with primary data gathered during workshops with practitioners in the United Kingdom. It finds that key challenges for data ethics in the public sector include the lack of accountability of data ethics tools, lack of skills and awareness, the saturated landscape of ethical guidance, and insufficient diversity in the field. Opportunities and recommendations include: embedding data ethics in data science processes; providing a platform that collates the available ethical guidance; developing data ethics courses for public sector data scientists; establishing a data ethics community; increasing transparency in hiring and introducing extensive diversity measures.

Keywords: data ethics, digital ethics, ethical principles, public sector data ethics, data ethics policy

INTRODUCTION

In the world dominated by technological progress, there is an overarching consensus that considering data and AI ethics policy development is necessary. Nowadays, ethical oversight constitutes the key component of safe and responsible deployment of data-driven technologies in most public sector institutions. In general, data ethics is a branch of applied ethics that studies and evaluates moral problems and describes the value judgements related to data, algorithms, and corresponding practices, in order to formulate and support morally good solutions (e.g. right conducts or right values) (Floridi, Taddeo 2016, p. 1; Gov.uk 2018). The overarching aim of data ethics is to promote responsible and sustainable use of data for the benefit of people and society (Tranberg et al 2018, p.8) and ensure that knowledge obtained through data is not used against the legitimate interests of an individual or group while identifying and promoting standards, values and responsibilities that allow us to judge whether decisions or actions are appropriate, ‘right’, or ‘good’. Nevertheless, for data ethics policy to serve its purpose, it needs to be understood and its function ought to be well-defined and given an appropriate mandate. This paper aims to investigate what are the obstacles and opportunities in the field of public sector data ethics through a combination of the literature review and primary evidence gathered through data ethics surveys and workshops in the United Kingdom. The following questions guide the research: (1) what should the function of data ethics policy be in the public sector? (2) What is the role of ethical principles? (3) What are the current challenges in data ethics in the public sector? This paper begins with defining data ethics policy and

principle-based approaches. Subsequently, key challenges are identified based on the literature review and additional data gathered through surveys and workshops. Finally, the paper provides a set of recommendations for data ethics practitioners in the public sector.

DATA ETHICS POLICY

Data Ethics Policy Function

It can be argued that establishing the role and aims of data ethics policy depends on the organisation, sector, and cultural background. Hasselbalch (2019) defined data ethics policy initiatives in the context of European public policy as ‘moving beyond issues of mere compliance with data protection law to increasingly focus on the ethics of big data, especially concerning private companies’ and public institutions’ handling of personal data in digital forms’. What is apparent in this definition is the natural split between the data ethics function focused on the private sector, which is inherently linked with regulation, and data ethics focused on examining the governments’ use of data. This paper analyses the latter. Quoting the work of the moral philosopher James H. Moor (1985) on conceptual vacuums caused by the societal implications of computer technologies, Hasselbalch (2019) adopts his definition of policy vacuums as ‘core problems and challenges, revealing “conceptual muddles” (Moor 1985, p. 266), uncertainties and the emergence of new values and alternative policies (p. 267)’. Thus, with Moor’s framework as the starting point, Hasselbalch (2019) defines data ethics policy initiatives as ‘moments of sense-making and negotiation created in response to the policy vacuums that arise when situations and settings are amended by computerised systems’. This enables data ethics to become the means of evaluating potential solutions and confronting ethical challenges caused by the rapid developments in data and technology. Hasselbalch (2019) emphasises how the policy vacuums invite questions on the governance of technologies that are interlinked with their impact on the personal, social, cultural, and economic activities. This indicates that in addition to providing the negotiation space that prompts policymakers to make choices on their value preferences, data ethics policy function ought to present suggestions on what these potential solutions could be.

In their practical paper on how to implement data ethics in the public sector, Taylor, Leenes, and van Schendel (2017, p. 24) argue that data ethics oversight for government must be multi-level and composed of overlapping checks and balances and that ethics is rooted in principles that can serve as a backdrop to such a sector-specific system of checks and balances. In practice, ethical needs should be evaluated within the context and function of each particular application. Thus, Taylor, Leenes, and van Schendel (2017) distinguish a model for data ethics oversight that consists of three elements: 1. Background principles; 2. Accountability framework/control loop; 3. Enforcement/oversight (p. 24). The first component of the model is a set of principles that serve as a baseline for responsible data science and innovation and help outline the core values involved in a particular data science application area. The second layer consists of the accountability scheme that enables the organisation to define what it does, perform the task, monitor it, remedy any discrepancies between what should occur and what is actually occurring; and explains and justifies its actions (p. 27). The model distinguished various types of reporting that organisations can choose from: extensive (impact assessment) or marginal (description of goals and methods); and internal or external reporting. Finally, although the first two parts of the model could already contribute to establishing a data ethics function, Taylor, Leenes, and van Schendel (2017, p. 27) propose oversight and enforcement (of both internal and external norms) as an opportunity to strengthen it even further. Here, the choices are between internal versus external oversight; ad hoc versus structural oversight; and determining the level/distance to the oversight organisation.

Based on these studies (Taylor, Leenes, and van Schendel 2017; Hasselbalch 2019), data ethics policy in the public sector could serve the following functions: provide means for negotiation and value judgements related to ethical challenges in data and technology; evaluate and propose potential solutions to ethical dilemmas; establish a set of centralised ethical principles and provide support on their implementation in specific contexts; determine the types of accountability frameworks and oversight needed and set them up.

Ethical Principles

Sociocultural complexity of data ethics led to the emergence of macroethics manifested by the presence of a principle-based approach that addresses the ethical impact and implications of data science within a holistic, action-oriented framework. Floridi and Taddeo (2016) perceive macroethics as the optimal approach that will lead to data ethics providing solutions that can maximize the value of data science for the society. Most frameworks are designed to voluntarily address risks of harm caused by new technologies, and point to design, business, policy, social and cultural processes that support a human-centric distribution of power (Hasselbalch 2019). Usually, the end goal is for ethical principles at a high level to permeate down to every aspect of business operation and conduct. Thus, principles should be seen as fundamental starting points to guide deliberation and action - their purpose is to point decision-makers in the direction of the relevant values to be taken into consideration when a particular decision or course of action is being contemplated. Therefore, principles are at the very core of any data ethics policy function, yet due to their high-level nature, the implementation process is context-dependent. This is in contrast with rules which are designed to indicate clearly to a decision-maker what must be done (or not done) in each case. Principles are designed to help reach a decision, whereas rules offer a specific diagnostic of actions to be taken or refrained from (Laurie, Seth 2013).

There is no universally accepted ethical system or agreed set of ethical principles. Nevertheless, the analysis of ethical frameworks developed by the UK, Australian, Canadian, Danish, German, Finnish governments and ethics commissions as well as the OECD, G20, European Commission conducted for the needs of this paper revealed that there are many similarities and common themes underpinning most frameworks, including: transparency, accountability, openness; human and user-centrism; privacy and security; democracy, justice, fairness, sustainability; understanding limitations of data and AI; explainability and contestability; links with legislation; data quality and skills.

Similarly, in their comparative analysis of thirty-six AI principles documents, Fjeld et al (2020) identified the eight main themes that constitute the 'normative core' of the ethics landscape: Privacy, Accountability, Safety and Security, Transparency and Explainability, Fairness and Non-discrimination, Human Control of Technology, Professional Responsibility, and Promotion of Human Values. In the analysis of the UK landscape of ethical principles, Floridi (2019) flagged that differences between various frameworks tend to be due to working and vocabulary rather than content, thus confirming the existence of common themes in the context of the UK.

METHODOLOGY

Given the need to connect the practical experience of data ethics professionals and the emerging theory and research on data ethics, this paper combines the literature review with primary data on challenges and opportunities for data ethics practitioners gathered through a survey focused on the redevelopment of the existing UK Data Ethics Framework. The team surveyed a sample of 32 officials working with data (data analysts, data scientists, data policymakers, statisticians, researchers) to understand their data ethics practices and knowledge of existing ethical guidelines. Additional data was collected during a series of data ethics workshops with attendants from the wider public sector in the UK, academia, civil society, and the tech industry. The team convened conversations on various ethical principles and methods of their implementation with the Data Ethics Framework as a case study and collected further comments through post-workshop forms. It is important to acknowledge that although when writing this paper, I benefited from my experience of working in data ethics policy in the public sector in the United Kingdom, this study was conducted independently from my role and the views, opinions, and findings of the paper do not represent those of institutions or organisations I am associated with in professional capacity unless explicitly stated. This paper adds to the existing scholarship by providing a combination of first-hand experience of setting up the data ethics policy function in the government with a range of views from public sector practitioners, and the academic literature on data and AI ethics.

CHALLENGES IN DATA ETHICS

Lack of Enforcement and Accountability of Data Ethics Tools, Principles, and Guidelines

One of the biggest challenges in data ethics policy is the lack of enforcement of the ethical frameworks, tools, and principles that generally operate on a voluntary basis. There is a lack of accountability in adhering to these principles and no concrete consequences that follow a failure to adhere. For Whittaker et al (2018), ethics guidelines fall into the category of ‘trust us’ form of non-binding corporate self-governance (p. 30) and this voluntary commitment to implement ethical practices should not be perceived by the public as a sufficient reassurance that their data is indeed used and processed ethically. Comparative study of AI ethics principles conducted by Fjeld et al (2020) supports that claim - although their analysis identified that all ethics guidelines analysed stated human rights adherence and respect as the foundational requirements, there is little evidence of implementation of these concepts in the real world (p. 66).

Furthermore, there is a risk of voluntary commitments substituting enforceable legal rules that could lead to data ethics principles not being released. Effectiveness of ethical principles and guidelines has been widely questioned by academics and professionals. In their controlled study, McNamara, Smith, and Murphy-Hill (2018) analysed the impact of ethical guidelines as a basis for ethical decision-making for software engineers. They found that the effectiveness of ethical codes is very limited and they fail to influence the behaviour of tech professionals. Similarly, Hagendorff (2019) challenges the efficacy of ethical guidelines due to the disconnect between the broad, overarching principles and their desired implementation in a diversified set of scientific, technical and economic practices. Therefore, the distance from the practices that ethics aims to govern results in the emergence of superficial guidelines that encourage the devolution of ethical responsibility to others, and fail to respond to the complexities of specific processions (2019, p. 9). The lack of effective implementation examples accompanying the release of ethical principles, as well as no further instructions on how, where, and by whom such guidelines may be operationalised has been criticised by Crawford et al (2019, p. 20). Their report points out that vague definitions of often conflicting ideals that constitute ethical guidelines often prevent their operationalisation in AI product development, deployment, and auditing frameworks; and the lack of accounts of methods used (ibid). This is aligned with Mittelstadt’s (2019) observations that developments in AI ethics lack common aims and fiduciary duties, professional history and norms, proven methods to translate principles into practice, and robust legal and professional accountability mechanisms. Thus, such lack of formal governance, oversight, and enforcement is at the very core of the reasons that prevent AI and data ethics principles from being as successful and impactful as in the case of principle-based approaches in more established professions such as medicine.

When asked whether ethical guidelines should be mandated within the UK Civil Service, a group of data scientists and data policymakers who participated in the workshops focused on the refresh of the Data Ethics Framework provided a range of options. 38% was against mandating the framework, 21% supported the idea to mandate it, and 41% responded with other solutions, including embedding ethical guidance within existing processes, driving bottom-up change, and promoting ethics as good practice rather than enforcing it. These findings indicate that there currently is a lack of a clearly defined way forward for ethical guidelines. In the case of the UK, the question of whether the Data Ethics Framework ought to remain voluntary or become mandated remains open.

The Landscape of Ethical Guidance Is Saturated

There is a variety of national, international, public, and private sector data and AI ethical guidelines available. This might decrease the engagement with ethical frameworks caused by an inability for public servants and private users to locate the most appropriate framework. As reported by the Committee on Standards in Public Life (2020), over 70 AI ethics codes have been published over the past three years in the UK. For Floridi (2019), this landscape saturation creates a number of issues: it generates inconsistency and confusion regarding which set of principles is preferable; it puts pressure on actors working in data and digital to produce their own guidelines at the risk of appearing to be left behind, further contributing to the redundancy of information; and creates a “market of principles and values”, where private and public actors

may shop for the kind of ethics that is best retrofitted to justify their current behaviours instead of implementing or improving new behaviours by benchmarking them against the public, ethical standards (p. 186). He coins the term ‘digital ethics shopping’ to describe the latter and warns that ethics shopping causes incompatibility of standards and hinders the ability to understand whether two organisations follow the same ethical principles in developing AI solutions, which results in a lower chance of comparison, competition, and accountability (p.187). In the case of the UK government, the Committee on Standards in Public Life (2020) reported that currently there are three main sets of ethical principles endorsed by UK government bodies: the AI Guide, the Data Ethics Framework, and the OECD Principles on Artificial Intelligence endorsed by the Centre for Data Ethics and Innovation and adopted by the government. The Committee concludes, similarly to Floridi’s study (2019), that this multiplicity of principles and codes confuses the landscape and can undermine attempts to make any set of ethical principles authoritative; and that there is a lack of clarity in how these principles work together (p.32). Therefore, the task is to collate, standardise, and transpose the existing ethical standards into a meaningful overarching approach. One potential short-term solution can be creating portals that consolidate multiple guidance and offer short assessment tools that help determine which ones are the most appropriate to use when. The recent publication of the Data ethics and AI guidance landscape could be the first step towards this consolidation. Floridi (2019) recommends that the strategy to mitigate the risks of digital ethics shopping is to establish clear, shared, and publicly accepted ethical standards. This could be one of the key tasks of a centralised data ethics policy function in the public sector.

Lack of Skills and Awareness

Designing and implementing data ethics strategies and initiatives requires a broad range of skills, including data science, policy, and specialised ethics knowledge. Assembling multidisciplinary teams with access to necessary subject matter expertise can be challenging, yet it is essential to establish robust practices for the appropriate data use. When using data or designing new processes or tools to do so, it is essential to have the ability to recognise if necessary information or expertise is lacking and address these gaps throughout any data project. In terms of specific data skills, the Royal Society (2019, p. 12) recommends that data ethics training and codes of practice should have a key role in developing data science as a profession. In the longer term, relevant professional bodies ought to work with employers and educational institutions to identify the key skills needed for data scientists and consider how to address accreditation. For the Royal Society, data ethics is the backbone of the potential professional ethos for data science. In the UK Civil Service, there are numerous established professions including analysts and data scientists with upskilling programmes and courses available. For instance, the Data Science Accelerator builds data science capability amongst the public sector analytical professions (statistics, economics, operational research and social research). The programme introduces them to data science techniques including machine learning, natural language processing, geospatial analysis, and advanced visualisation (Gov.uk, 2020). In addition, the data science community has been established to develop and share best practice in the field and bring together data scientists working across the government to provide space for discussion and idea generation. Although these initiatives are a step in the right direction, there is a lack of a well-defined ethics component of the existing data science training and communities. Furthermore, developing an effective data ethics course is a demanding task. As outlined by Stoyanovich and Lewis (2019), the existing data ethics courses in academia rely exclusively on texts rather than on software development or data analysis. Thus, technical students can dismiss them as a distraction from the ‘real’ material. To mitigate this risk and develop thoughtful and engaging content, Stoyanovich and Lewis (2019) postulate for balance ‘between texts and coding, between critique and solution, and between cutting-edge research and practical applicability’. They argue that finding this balance is particularly challenging in the emerging field of responsible data science where there is a need to consolidate varying methodologies of engineering and social science. Although their work focuses on academia and teaching data science at universities, the assumption in this paper is that similar challenges apply when developing ethics courses for practitioners.

Another crucial challenge for the public sector data ethics is popularising and ensuring the take-up of ethical guides and frameworks. A survey conducted amongst data scientists and data policymakers across the UK Civil Service prior to the refresh of the Data Ethics Framework revealed that the lack of awareness is the most prominent obstacle in promoting data ethics tools and practices. 37,5% of respondents haven't heard of any of the major data and AI ethics guidelines produced by different departments. In terms of the Data Ethics Framework, it has been used only by 18,8% of survey respondents. Not knowing that such Framework exists was the main reason for this low uptake as reported by 57.7% of surveyed officials, followed by time constraints (11,5%). Other reasons included the lack of perceived relevance of the Framework, the lack of understanding of how to use it; ethical frameworks not being viewed as a priority. On the other hand, out of the group that did use the Framework, the majority (66.7%) was advised to use it by a colleague or a manager or found out about it through the word of mouth (16.7%). These findings indicate that the most effective manner of raising awareness and encouraging the use of available ethical guidelines has been through informal and/or professional networks, rather than through the established corporate channels or communities of practice. Furthermore, the lack of awareness of ethical tools is the biggest obstacle hindering their usefulness. Therefore, the challenge is to promote the existing guidelines to increase their outreach beyond informal recommendations.

Insufficient Diversity in the Field

Crawford et al (2019) criticise the emerging AI ethics principles and statements that despite being primarily Western entities driven by predominantly white, male, wealthy individuals, are labelled as products of 'global consensus' on AI ethics. They argue that this contributes to promoting a majoritarian view of ethics that can further reinforce the dynamics in which AI bias and misuse negatively affect the people whose voices are not present in these debates. The wider lack of diversity in AI as analysed by West, Whittaker, Crawford (2019), manifests itself through inequalities in the workplace, industry and academia, in disparities in hiring and promotion, and finally, in AI technologies that reflect and amplify biased stereotypes; and resurfacing of biological determinism arguments (p. 32). Nevertheless, fairness and non-discrimination principles calling for designing AI systems to maximise fairness and promote inclusivity are present in thirty-six AI ethics documents in the dataset analysed by Fjeld et al (2020). This indicates that, in line with the earlier criticisms of ethical principles being disconnected from practice and implementation, although in theory there is a widespread consensus that diversity in AI and data ethics is essential, in practice the industry is still falling behind.

There are currently no studies or statistics available on the diversity of data ethics professionals, especially in the UK public sector, thus it is not possible to fully estimate the scale of the problem. Nevertheless, the available statistics on the wider tech sector or the civil service workforce confirm that there is an urgent need for addressing the diversity gap. As reported by the Alan Turing Institute (2020), in spite of the increase in availability of data science and machine learning roles in recent years, women occupy only a minority of these new positions. In the UK, women represent 47% of the workforce, yet they hold less than 17% of all available tech jobs (ONS 2017). Alan Turing Institute (2020) suggests that this growing evidence indicates the gender imbalance that affects the wider tech sector extends to data science and AI. Moreover, the existing statistics most often assume gender as binary ('male' or 'female'), thus erasing all other forms of gender identity and in consequence contributing to their further discrimination (West, Whittaker, Crawford 2019). In the wider Civil Service in the UK, the latest 2019 statistics reported that 53.9% of all employees are women, 12.7% are from an ethnic minority background, 11.7% declare themselves as having a disability, 4.9% identify as being lesbian, gay, bisexual or recorded their sexual orientation as 'other' (Civil Service Statistics 2019). This data is a total from all departments and all professions (including data science and analytical professions, as well as data policymaking) hence it is not conclusive when estimating the scale of the diversity gap in the government data science and data ethics work. However, it does signal that the relative lack of diversity might permeate to the specific professions, especially the tech and data ones that have historically had low diversity figures.

To improve workplace diversity, West, Whittaker, Crawford (2019) recommend increasing transparency in hiring practices and compensation levels across all roles and job categories, broken down

by race and gender; changing hiring practices to maximise diversity (targeted recruitment beyond elite universities, more equitable focus on under-represented groups, creating pathways for contractors, temps, and vendors to become full-time employees), and incentive structures for hiring and retention of under-represented groups; increasing the number of people of colour, women and other under-represented groups at senior leadership levels in AI-related roles (p. 4). To address bias and discrimination in algorithmic systems, they postulate for increasing transparency, rigorous testing, and researching bias and fairness through social analysis of how AI is used in context (ibid).

OPPORTUNITIES AND RECOMMENDATIONS

In light of the challenges outlined above, this paper aims to provide the following practical recommendations and opportunities for a successful data ethics policy function in the public sector.

Firstly, to ensure that ethical frameworks and guidance are being actively used, due to their voluntary nature, they ought to be firmly embedded in the public sector data science work and processes. The opportunity in the public sector is to present data ethics as an essential step of any data science process and normalise ethical assessments. Ensuring that ethical frameworks are built into the project process and follow its implementation cycle would increase their effectiveness and uptake. In the long run, the task of the government data ethics policy function might be to scope potential legislative measures to support the implementation of data ethics in the public sector.

Secondly, now a variety of data ethics guidance and standards exist, the task is to collate, standardise and transpose them into a meaningful overarching approach. To help navigate this saturated landscape of the public sector data ethics, the centralised policy function should assist other bodies with identifying the specific guidelines tailored to their needs - for instance, a procurement-focused framework or a health-specific framework. The centralised data ethics function might provide a platform that gathers all the available guidance and issues instructions on how to differentiate between them. It might also, in line with Floridi's (2019) recommendation, establish a set of centralised ethical principles and provide support on their implementation in specific contexts.

Thirdly, to address the lack of data ethics skills, the centralised policy function could develop and provide data ethics training for civil servants working on data through the centralised learning platform; or embed data ethics modules in the existing data science courses. To increase the awareness of the existing ethical guidelines, the central data ethics policy function ought to convene a data ethics working group within the existing data science circles or establish a separate data ethics community that would enable practitioners and policymakers to exchange learnings, provide case studies of using ethical guidelines, and in the long run, serve as the data ethics champions within their departments.

Addressing the lack of diversity in the field of data ethics is a complex effort that is intertwined with the wider recruitment and promotion policies. On a micro level, the public sector data ethics policy function should recognise if it lacks diversity and expertise, and establish regular communication and outreach channels with a diverse range of stakeholders from all relevant sectors. In line with recommendations from West, Whittaker, and Crawford (2019), there ought to be more transparency in hiring and processes and compensation. When recruiting for data and AI ethics positions, there should be an emphasis on diversity considerations and attracting candidates from under-represented groups. Data ethics community ought to contribute to the government-wide diversity efforts and participate in creating potential incentive structures within the wider tech and data policy functions.

A potential limitation of this paper is that although the research has drawn from academic literature, it presents data ethics policy from a particular, Eurocentric practitioner's perspective in the United Kingdom. Thus, challenges and opportunities identified in this work might not be fully applicable in different cultural contexts. Furthermore, this research focuses on data ethics in the public sector and further work is needed to investigate this issue in the private sector.

ACKNOWLEDGEMENT

I would like to thank S. Roberts, A. Scieluna, R. Want, and S. Krouwel for their invaluable work and support during the refresh of the Data Ethics Framework, and the Data for Policy 2020 team for enabling my participation in the conference. Although when writing this paper I benefited from my experience of working in data ethics policy in the public sector in the United Kingdom, this study was conducted independently from my role and the views, opinions, and findings of the paper do not represent those of institutions or organisations I am associated with in professional capacity unless explicitly stated.

REFERENCES

- Australian Human Rights Commission. (2019). *Human Rights and Technology*. Retrieved from https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights2019_DiscussionPaper.pdf
- Civil Service Statistics. (2019). *Statistical bulletin V2*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/836373/Statistical-bulletin-Civil-Service-Statistics-2019-V2.pdf
- Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, K., . . . Whittaker, M. (2019). *AI Now 2019 Report*. New York: AI Now Institute. Retrieved from https://ainowinstitute.org/AI_Now_2019_Report.html
- Danish Expert Group on Data Ethics. (2018). *Data for the Benefit of the People*. Retrieved from <https://dataetiskraad.dk/sites/default/files/2020-02/Recommendations%20from%20the%20Danish%20Expert%20Group%20on%20Data%20Ethics.pdf>
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Sriksumar, M. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI*. Berkman Klein Center Research Publication No. 2020-1. Retrieved from <https://ssrn.com/abstract=3518482> or <http://dx.doi.org/10.2139/ssrn.3518482>
- Floridi, L. (2019). Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. *Philosophy & Technology*, 32(185). Retrieved from <https://link.springer.com/article/10.1007/s13347-019-00354-x>
- Floridi, L., & Cowsls, L. (2019, June 22). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*. <https://doi.org/10.1162/99608f92.8cd550d1>
- Floridi, L., & Taddeo, M. (2016). What is data ethics? *Phil. Trans. R. Soc. A*, 374(20160360). <http://dx.doi.org/10.1098/rsta.2016.0360>
- Gov.uk. (2018). *Data Ethics Framework* [online]. Retrieved from <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>
- Hagendorff, T. (2019, October 11). *The Ethics of AI Ethics—An Evaluation of Guidelines*. arXiv:1903.03425 [Cs, Stat]. Retrieved from <http://arxiv.org/abs/1903.03425>
- Hasselbalch, G. (2019). Making sense of data ethics. The powers behind the data ethics debate in European policymaking. *Internet Policy Review*, 8(2). DOI: 10.14763/2019.2.1401
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nat Mach Intell*, 1, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Laurie, G., & Sethi, N. (2013). Towards Principles-Based Approaches to Governance of Health-related Research using Personal Data. *Eur J Risk Regul.*, 4(1), 43–57. doi:10.1017/s1867299x00002786
- Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. The Alan Turing Institute. <https://doi.org/10.5281/zenodo.3240529>
- McNamara, A., Smith, J., & Murphy-Hill, E. (2018). Does ACM’s Code of Ethics Change Ethical Decision Making in Software Development? In G.T. Leavens, A. Garcia, & C.S. Păsăreanu (Eds.), *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering*

- Conference and Symposium on the Foundations of 15 Software Engineering - ESEC/FSE 2018* (pp. 1–7). New York: ACM Press.
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nat Mach Intell*, *1*, 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016, December). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2).
- Moor, J.H. (1985). What is computer ethics? *Metaphilosophy*, *16*(4), 266-275. doi:10.1111/j.1467-9973.1985.tb00173.x
- Office for National Statistics. (2017, February). *UK labour market statistics*. Retrieved from <https://www.ons.gov.uk/releases/uklabourmarketstatisticsfeb2017>
- Stoyanovich, J., & Lewis, A. (2019). Teaching Responsible Data Science: Charting New Pedagogical Territory. In *Proceedings of ACM Conference (Conference '17)* (14 pages). ACM, New York, NY, USA. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>
- Taylor, L., Leenes, R., & van Schendel, S. (2017). *Public Sector Data Ethics. From principles to practice*. Tilburg University TILT – Tilburg Institute for Law, Technology, and Society.
- The Alan Turing Institute. (2020). *Women in data science and AI*. Retrieved from <https://www.turing.ac.uk/research/research-projects/women-data-science-and-ai#:~:text=As%20the%20national%20institute%20for,step%20towards%20fulfilling%20that%20responsibility>
- The Committee on Standards in Public Life (CSPL). (2020). *Artificial Intelligence and Public Standards: report*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF
- The Royal Society. (2019). *Dynamics of data science skills: How can all sectors benefit from data science talent?* Retrieved from <https://royalsociety.org/-/media/policy/projects/dynamics-of-data-science/dynamics-of-data-science-skills-report.pdf>
- Tranberg, P., Hasselbalch, G., Kofod Olsen, B., & Søndergaard Byrne C. (2018). *DATAETHICS – Principles and Guidelines for Companies, Authorities & Organisations* (1st Edition). Retrieved from <https://dataethics.eu/wp-content/uploads/Dataethics-uk.pdf>
- West, S.M., Whittaker, M., & Crawford, K. (2019). *Discriminating Systems: Gender, Race and Power in AI*. AI Now Institute. Retrieved from <https://ainowinstitute.org/discriminatingystems.html>
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., . . . Schwartz, O. (2018). *AI Now Report*.
- Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). *Ethical and societal implications of algorithms, data, and artificial intelligence: A roadmap for research*. London: Nuffield Foundation. Retrieved from <https://www.nuffieldfoundation.org/sites/default/files/files/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf>
- Wirtz, B.W., & Weyerer J.C. (2019). Artificial Intelligence in the Public Sector. In A. Farazmand (Ed.), *Global Encyclopedia of Public Administration, Public Policy, and Governance*. Springer, Cham.