

# DATA ETHICS SKILLS: PROFESSIONALISING DATA ETHICS IN THE PUBLIC SECTOR

**Natalia Domagala**

*Head of Data Ethics at the Central Digital and Data Office, Cabinet Office, UK  
natalia.domagala@digital.cabinet-office.gov.uk*

## Abstract

Data ethics is a cross-cutting discipline that critically assesses the wider societal impact of technology and data and produces actionable recommendations for other data professionals. It involves thinking about fairness, accountability, the law, moral dilemmas, and risks involved in the creation of technology and data products and policies. Due to the complexity and multidisciplinary demands of this field, there is a need for a designated data ethics position in the public sector that would be responsible for providing research and expertise on data ethics, and acting as a champion and change agent. This paper outlines what the role of the data ethicist would entail, for instance: enabling and empowering others to implement data ethics best practice in their work; demonstrating how to apply ethical principles in practice; communicating effectively to explain and raise awareness of data ethics issues; listening, convening, advising and mediating between various parts of the organisation. It also entails creating a space for others to ask questions, express concerns, and discuss ethical dilemmas. The paper connects the practical experience of data ethics professionals and the emerging theory and research on data ethics with primary data gathered during workshops and a comprehensive survey with data ethics practitioners in all sectors. The research identifies key skills for effective data ethics work and provides a description of a designated Data Ethicist role for the Digital, Data and Technology (DDaT) Professional Capability Framework in the United Kingdom.

**Keywords** – data ethics; digital ethics; data skills; public sector data ethics; data ethics skills

## 1 Introduction

Data ethics is a cross-cutting discipline that critically assesses the wider societal impact of technology and data and produces actionable recommendations for other data professionals. It involves thinking about fairness, accountability, the law, moral dilemmas, and risks involved in the creation of technology and data products and policies.

Due to the complexity and multidisciplinary demands of this field, there is a need for a designated data ethics position in the public sector that would be responsible for providing research and expertise on data ethics, and acting as a champion and change agent. This multidisciplinary role needs to sit across data science and policy teams and link all strands of data work while considering the big picture and their long-term societal impacts. This paper aims to identify key skills necessary for working in data ethics. The following questions will guide the research: (1) What is the purpose of having a designated data/AI ethics person within an organisation and what their role should entail?; (2) How do data ethics skills differ from data science skills? (3) What specific skills are needed when working in data ethics? The research presented in this paper is based on the practice of working in data ethics in the public sector and it outlines the process of professionalising data ethics in the UK government.

### 1.1 The case for data ethics skills

Data ethics ought to be an essential part of the future professional ethos for data science, as stressed by the Royal Society (2019, p. 12). Their recommendations include working with the relevant educational institutions to identify key skills needed for data scientists, in particular those related to data ethics. The Guide to Responsible Tech (All Tech Is Human, 2020) provides an overview of emerging data ethics professions in the private sector. The roles featured in the guide include job titles such as AI Ethicist; Activation lead, ethics and society; Chief AI Ethics Officer; Chief Ethical and Humane Use Officer; Ethical AI Lead. The guide affirms that AI and data ethics jobs are essential to the successful product and project development in the majority of big tech companies and that there is a need to employ data ethicists with a diverse skillset and social science background. Developing data ethics strategies and products in the public sector requires specialist ethics knowledge alongside data literacy and organisational

savviness. As observed by Domagala (2021a), in the UK Civil Service the established data professions such as analysts and data scientists are well supported through upskilling programmes, courses, and communities of practice. Nevertheless, there has been a lack of a well-defined ethics component of the existing data science training and communities and little clarity on what skills are needed when working in data ethics.

The OECD Good Practice Principles for Data Ethics in the Public Sector (2021) emphasise that the effectiveness of data ethical frameworks and guides requires cross-organisational cooperation. They suggest that designated institutional roles and responsibilities should be a part of the successful data governance arrangement (p. 5). Furthermore, the OECD recommend recognising and introducing measures to mitigate ethical risks at different levels to support the impact of non-binding data ethics guidance. Proposed initiatives include reinforcing data ethics considerations in the range of policy levers available to governments, as well as in the accountability mechanisms public officials are subject to; and integrating data ethical considerations into contractual agreements or the terms and conditions of partnerships with third parties or external actors (p. 7). Introducing a data ethics profession with responsibilities for embedding and implementing ethical considerations across these areas would support delivering the recommendations from the OECD. In the UK, the National Data Strategy published by the Department for Digital, Culture, Media and Sport (2020) emphasised the significance of professionalising data ethics: ‘we will (...) support data scientists and data policymakers to build lasting capability for ethical data use’. This commitment was reiterated further in response to the call for evidence where the government pledged to lead by example, including scoping data ethics as a role in the Digital, Data and Technology Capability Framework (Department for Digital, Culture, Media and Sport, 2021). Similarly, the consultation response to the strategy stressed that data ethics should be at the heart of government data operations and indicated the need for improving ethical norms and standards to guide data sharing across the government. Recommendations included improving accountability structures for data ethics tools, making Data Ethics a recognised DDaT Capability Framework profession for all government teams, and introducing skills programmes focused on data ethics (ibid).

This paper supports these recommendations to professionalise data ethics in the public sector and outlines the research process of creating the Data Ethicist role for the Digital, Data and Technology (DDaT) Professional Capability Framework in the United Kingdom. The DDaT Capability Framework was published in 2017 to help improve consistency across data, digital, and technology job

roles in the UK government. It sets out a common language to describe roles in these fields and defines the skills needed for each. The framework facilitates government transformation and personal development for DDaT professionals, and it helps identify skill gaps in the wider public sector. This research confirms that including a designated data ethicist role as a part of the Data job family in the UK DDaT Capability Framework would enable data engineers and scientists to innovate responsibly and respond to the ongoing demand for implementing data ethics best practices on the ground. Data ethicist is a critical functional role that has been extremely successful in recent years in the private sector and is now an instrumental function in any high-risk data and AI products. Considering that many data science projects and processes in the government have a direct impact on the public, there is a much stronger case to provide sufficient oversight and advice on data ethics in the public sector.

## 2 Methodology

The approach to research on data ethics in this paper derives from the ethos of open government and ways of working defined in the UK Open Government Playbook (2020). In the spirit of collaboration, external stakeholders were invited to participate in an interactive workshop to share their expertise and feedback with the team working on data ethics skills. Due to the timing of this research amid the Covid19 pandemic, the workshop was conducted online alongside an anonymous survey that was openly shared with data ethics practitioners and the general public.

To decide what skills should be a part of the Data Ethicist role, the team surveyed 50 data and AI ethics practitioners from the public sector (44% of respondents), private sector (20% of respondents), academia (18%), media (2%) and civil society (16%). The survey contained skills from the Data job family of the UK DDaT Capability Framework and additional skills focused on data ethics in a random order (Appendix A, B). In addition to the survey, the team ran a workshop with 25 data ethics experts and professionals from the UK public sector, featuring guest speakers working as data ethics specialists and consultants in the private sector in the UK and the US. The aim of the session was to work with practitioners to crowdsource the information on the purpose of having a designated data/AI ethics position within an organisation, define what their role should entail, and identify a set of skills needed when working in data ethics (Appendix C). The findings from the workshop were qualitatively analysed using a thematic coding technique. Subsequently, they were combined with the survey results to serve as an evidence base to draft the description of the Data Ethicist role.

The UK-based practitioner focus is a potential limitation of this study. The skills and organisational conditions presented in this work are specific to the UK and might not be fully transferable to other cultural contexts. This paper adds to the existing scholarship by providing a first-hand, step by step account of developing a data ethics role for an established skills framework in the public sector, including detailed descriptions of the collaborative research process. It should serve as an example for public and private sector practitioners willing to introduce a Data Ethicist role in their organisations.

### 3 Key findings

#### 3.1 Data Ethicists within an organisation

Survey responses and workshop contributions emphasised the need for a designated data/AI ethicist to connect and mediate between different parts of the organisation and provide practical support on embedding ethical considerations in data projects. Such support would include helping teams embed data ethics in their project delivery and choose data ethics tools and frameworks that are most suitable for their projects. Other suggestions included: championing and coordinating data ethics policy across the organisation, translating and sharing ideas between teams, and building lasting stakeholder relationships across legal, technical, digital, and policy teams. Responses emphasised the importance of building and cultivating external connections, especially with academics and civil society groups working on data ethics, and communicating new developments in the field across the government.

Based on these results, this paper suggests the following summary of the duties and purpose of having a data ethicist within an organisation. This description forms the Data Ethicist role in the DDaT Capability Framework:

As a data ethicist, you will:

- Provide research and expertise on data ethics, and act as a champion and change agent.
- Enable and empower others to implement data ethics best practice in their work. This could include providing data ethics training, advising data science teams, demonstrating how to apply ethical principles in practice through examples and case studies, and driving best practices in data ethics.
- Communicate effectively to explain and raise awareness of data ethics issues as well as listen, convene, advise and mediate between various parts of the organisation. You will create a space for

people to ask questions, express concerns, and discuss ethical dilemmas.

#### 3.2 How do data ethics skills differ from data science skills?

On the topic of how data ethics skills differ from traditional data science skills, survey respondents and workshop participants offered a range of definitions. They described data ethics skills as focused on understanding the society and broader social and economic impacts of technologies, and the norms it helps create and shape. In contrast, data science focuses on specific project delivery aspects and technological capabilities rather than social implications. Insights from the survey included the following definitions: ‘data science is gathering and drawing conclusions from data. Data ethics is making sure those conclusions and the methods for collecting them are moral’ and ‘data ethics is about knowing what you should do, data science about what you could do’.

Data ethics is a cross-cutting discipline operating across data science, engineering, governance, project management, service design, and other functions within an organisation. Responses stressed that data ethics is a big-picture discipline that should include a holistic view of the data lifecycle and the risks and harms introduced at each phase. Data science focuses on extracting value at each stage of the data lifecycle and the technical aspects such as coding, cleaning, and analysing data. Furthermore, there was an emphasis on data ethics as a space for negotiation where it is necessary to rethink objectives, integrate different perspectives, and constantly redefine complex problems that emerge. The findings stressed the importance of social science or humanities background to facilitate the practical application of societal norms in data work. Other responses highlighted the significance of thinking about fairness, accountability, law, and the ability to interpret relevant data and AI laws and regulations. The respondents indicated that data ethics skills are generally not taught in many data science courses, thus reaffirming the gap identified in this paper.

Based on the findings from the workshop and the survey, and the existing definitions of data ethics, this paper provides the following short description for the DDaT Capability Framework: ‘Data ethics is a cross-cutting discipline that critically assesses the wider societal impact of technology and data and produces actionable recommendations for other data professionals. It involves thinking about fairness, accountability, the law, moral dilemmas, and risks involved in the creation of technology and data products and policies’.

### 3.3 Skills required to be a Data Ethicist

The majority of survey responses heavily favoured social science and humanities skills as opposed to traditional data science skills from the DDaT Capability Framework. These findings further support the hypothesis that the gap in the data job family could be addressed by including a role focused on data ethics. Findings from the workshop and the survey led to identifying the following skills as essential when working in data ethics within the public sector:

- Communications skills
- Knowledge of/ background in social sciences
- Engagement and collaboration
- Ability to draw together, analyse, and critically evaluate information
- Data literacy
- Product development and management skills
- Empathy and Inclusivity.

These skills were further refined to comply with internal DDaT Capability standards and harmonised with the existing skills catalogue to avoid duplication and encourage the use of the skills that are currently listed in the wider framework (Domagala, 2021b):

- **Communication skills (data).** You can communicate effectively across organisational, technical and political boundaries, understanding the context. You know how to make complex and technical information and language simple and accessible for non-technical audiences. You can advocate on behalf of a team and communicate what it does, to create trust and authenticity. You know how to respond to challenges.
- **Applied knowledge of social sciences.** You have the knowledge of or background in social sciences (anthropology, economics, sociology, philosophy, psychology, race theory etc). You are a critical thinker with a social science or humanities background. You can operationalise and implement the wealth of theoretical knowledge from social sciences and use it to inform data projects, products, and policies, and to evaluate and challenge assumptions made in data science projects. You are familiar with domain knowledge of existing schools of thought and exemplar models of data ethics in practice.
- **Stakeholder relationship management.** You know how to identify, analyse, manage and monitor relationships with and between stakeholders. You can communicate with stakeholders clearly and regularly, clarifying mutual needs and commitments through consultation and consideration of impacts while focusing on user needs.
- **Analysis and synthesis (data ethics).** You have the ability to draw together, analyse, and critically evaluate information. This includes the ability to quickly read and interpret complex documents from a range of sources and distill to what is relevant. You can turn research data into clear findings that inform decisions. You know how to involve colleagues in analysis and synthesis to increase consensus and challenge assumptions. You are able to analyse project information and data to exercise sound judgement of ethical risk and potential pitfalls and provide recommendations on how to prevent that. You actively follow the latest research and academic developments in data ethics and draw from them in your work. In critically evaluating project plans, you help teams define the outcomes and ethical considerations of their projects, and integrate ethical diagnostics and assessment in their work.
- **Bridging the gap between the technical and non-technical (data ethics).** You can translate technical concepts relating to software engineering, delivery management and service management so they are understood by all. You can mediate between people and mend relationships, communicating with stakeholders at all levels. You need to have an understanding of how technology and data products and services are built. You are familiar with the technical jargon and have a sufficient knowledge of data to hold meaningful conversations with data science experts on issues such as minimising bias in data, data gathering, collecting, cleansing, triangulating, reusing etc. You are able to effectively support data scientists and engineers in implementing data ethics.
- **Product development (data ethics).** Working in data ethics involves creating frameworks, guides, and other resources for practitioners. You are able to develop data ethics tools and translate theoretical principles into practice. You know how to use a range of product management principles and approaches. You can capture and translate user needs into deliverables. You know how to work with a range of specialists in multidisciplinary teams. You are familiar with feedback gathering, evaluation mechanisms, and product promotion. In evaluating the existing data ethics tools, you can identify best practices and map how data ethics is being implemented within an organisation.
- **Empathy and Inclusivity.** You have an inclusive approach to consensus building and ability to incorporate disparate views of underrepresented groups in evolving products and policy work. You are involved in the wider organisational diversity and inclusion agenda and you actively draw on your multidisciplinary background and lived

experience to understand the consequences of data systems on a diverse range of stakeholders. You have a thorough understanding of social issues, types of bias, and discrimination different groups can face, and you are able to use this knowledge to inform your data ethics work.

- **Ethics and Privacy.** You understand the ethical considerations of potential data science approaches, and the legislation applicable in this area, i.e. GDPR, DPA etc. You are aware of existing data and AI ethics frameworks within the Government and externally.
- **Problem Solving.** You can identify and understand problems, analysing and helping to identify the appropriate solution. You can classify and prioritise problems, document their causes and implement remedies.
- **Facilitating decisions and risks.** You can make and guide effective decisions, explaining clearly how the decision has been reached. You have the ability to understand technical complexity and risks, run collaborative design activities, influence others and build consensus.

To comply with the requirements of the DDaT Capability Framework, the core skills described above have been divided into practitioner and expert levels to correspond with other roles in the Framework. Furthermore, to help practitioners identify skills that need development and help career progression in data ethics, there are currently two levels available as a part of this role. The Data Ethics Lead is a junior role that supports the Head of Data Ethics to enable others to understand data ethics and implement best practice in their work. The Head of Data Ethics is a senior position that leads on the development and implementation of organisational data ethics policies to empower colleagues and drive lasting change.

## 4 Conclusion

The Data Ethicist role developed based on the expert engagement outlined in this paper has been added to the DDaT Capability Framework and launched on the 7th of September 2021. The role is currently available on the internal skills platform Knowledgehub and as a downloadable PDF on gov.uk (Domagala, 2021b). Data Ethicist is a part of the Data job family of the DDaT Capability Framework, and the description should inform the recruitment of data ethics professionals. As with all content on the Digital, Data and Technology Capability Framework, this role is subject to iteration and improvement. The role was published in its current form to support DDaT attraction, recruitment and retention activities across the government. Further research is required in the emerging field of professionalising data

ethics, especially on the implementation of data ethics roles in various capability frameworks and on the evaluation and impact of data ethics skills.

## Acknowledgements

I would like to thank S. Roberts and I. Fudge for their invaluable support during the data ethics skills workshop. 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

All Tech Is Human. (2020). *Guide to Responsible Tech: How to Get Involved & Build a Better Tech Future*.

<https://www.scribd.com/document/476272088/Guide-to-Responsible-Tech-How-to-Get-Involved-Build-a-Better-Tech-Future>

Department for Digital, Culture, Media and Sport. (2021). *Government response to the consultation on the National Data Strategy*. GOV.UK.

<https://www.gov.uk/government/consultations/uk-national-data-strategy-nds-consultation/outcome/government-response-to-the-consultation-on-the-national-data-strategy>

Department for Digital, Culture, Media and Sport. (2020). *National Data Strategy*. GOV.UK.

<https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>

Department for Digital, Culture, Media & Sport, Cabinet Office (2020). *Open Government Playbook*. GOV.UK.

<https://www.gov.uk/government/publications/open-government-playbook>

Digital, Data and Technology Profession. (2020). *Digital, Data and Technology Profession Capability Framework*. GOV.UK.

<https://www.gov.uk/government/collections/digital-data-and-technology-profession-capability-framework>

Domagala, N. (2021a). Data Ethics in Practice: Challenges and Opportunities for a Data Ethics Policy Function in the Public Sector. *Journal of Leadership, Accountability and Ethics*, 18(2).

<https://articlegateway.com/index.php/JLAE/article/view/4253/4042>

Domagala, N. (2021b). *What skills do you need for working in data ethics?* Gov.Uk.

<https://dataingovernment.blog.gov.uk/2021/09/07/what-skills-do-you-need-for-working-in-data-ethics/>

OECD. (2021). *Good Practice Principles for Data Ethics in the Public Sector*. <https://www.oecd.org/gov/digital->

[government/good-practice-principles-for-data-ethics-in-the-public-sector.pdf](https://royalsociety.org/-/media/policy/projects/dynamics-of-data-science/dynamics-of-data-science-skills-report.pdf)

The Royal Society (2019). *Dynamics of data science skills: How can all sectors benefit from data science talent?* Available at: <https://royalsociety.org/-/media/policy/projects/dynamics-of-data-science/dynamics-of-data-science-skills-report.pdf>

- Quality assurance, validation and data linkage abilities
- Data engineering and manipulation
- Other...

## Appendix

### Appendix A: Data ethics skills survey

1. According to you, which of the following skills are the most important in data ethics? Choose up to six.
  - Ability to quickly read and interpret complex documents from a range of sources and distill to what is relevant
  - Knowledge of/ background in social science (anthropology, economics, sociology, philosophy etc)
  - Problem-solving skills
  - Experience translating technical information for a non-technical audience, and vice versa
  - Domain knowledge of existing schools of thought and exemplar models of data ethics in practice
  - Communication skills
  - Data modelling, data cleansing, and data enrichment skills
  - Data literacy
  - Project management skills
  - Networking and engagement skills; working collaboratively across multidisciplinary teams and stakeholder from data science and other areas (policy, social science)
  - Proven IT and mathematical skills
  - Broad policy analysis and evaluation experience
  - Ability to draw together, analyse, and critically evaluate information
  - Sound judgement on relevant information, stakeholders, and value-driven activities
  - Applied maths, statistics, and scientific practices
  - Logical and creative thinking skills
  - Understanding analysis across the product life cycle

2. How do data ethics skills differ from data science skills? (optional)
3. What additional skills do people working in data ethics need beyond technical knowledge?
4. Is there anything that you would like to add? (optional)
5. What is your role? (optional)
6. Where do you work?

### Appendix B: Top skills from the survey

	% (responders were allowed to choose up to 6 skills)
Networking and engagement skills; working collaboratively across multidisciplinary teams and stakeholder from data science and other areas (policy, social science)	66%
Ability to draw together, analyse, and critically evaluate information	60%
Experience translating technical information for a non-technical audience, and vice versa	58%
Knowledge of/ background in social science (anthropology, economics, sociology, philosophy etc)	54%
Domain knowledge of existing schools of thought and exemplar models of data ethics in practice	44%
Data literacy	36%
Logical and creative thinking skills	30%
Sound judgement on relevant information, stakeholders, and value-driven activities	30%
Communications skills	28%
Ability to quickly read and interpret complex documents from a range of sources and distill to what is relevant	24%

### Appendix C: Workshop questions

- What is the purpose of having a designated data/AI ethics person within an organisation?
  - *What should their objectives be?*
- What should the role of the 'data/AI ethicist' entail?

- *What should they do on a daily basis?*
- *What would their job description be?*
- How do data ethics skills differ from data science skills?
  - *How is working in data ethics different from working in data science?*
  - *What other skills are needed beyond the core data science skills?*
- What additional skills and knowledge do people working in data ethics need beyond technical knowledge?
  - *How about humanities or social science knowledge/background?*
  - *Do they need any additional knowledge of philosophy/psychology/anthropology/history etc?*
- What should various job levels/ types of data ethics roles be? What career paths are available in data/AI ethics (e.g. trainee data ethicists vs chief data ethicist)?
  - *Is there a separate career ladder in data ethics or are people working in data ethics a part of other ladders (e.g. data science, policy)?*
  - *What could the career ladder be for data ethics?*
  - *How an intern in data ethics could progress within an organisation?*