Brokering Ethics: A Retrospective Study on the Use of a Data Ethics Framework

Sophia Worth, Georgia Panagiotidou, Elena Simperl

King's College London
Department of Informatics
WC2B 4BG London, U.K.
sophia.worth@kcl.ac.uk, georgia.panagiotidou@kcl.ac.uk, elena.simperl@kcl.ac.uk

Abstract

Recent years have seen a surge of interest in structured tools and interventions for responsible data and technology development and deployment. As part of this phenomenon, data ethics frameworks have emerged as a basic tool that go beyond lists of 'ethics principles', aiming to guide the practical application of these ethical considerations. This study offers an in-depth insight into how organisations are making use of the Data Ethics Canvas - a popular, openly licensed data ethics framework, developed by a non-profit, and used across a diverse range of organisations. Through a reflexive thematic analysis of 22 interviews with tool users, we draw key reflections to understand its role in organisational settings. We find that participants aim to use the tool for data ethics assessment, literacy, and as a foundation or inspiration for governance initiatives. We argue that the versatility of the tool offers an opportunity for 'brokering' among stakeholders who approach data practices from different perspectives. Still, limited guidance and a need to overcome organisational barriers leave significant responsibility on individuals to deliver outcomes from interactions with the tool. We further discuss how far these findings can be used to reflect on wider use, re-design, re-framing and evaluation of interventions to generate responsible data practices. We offer recommendations including the need for clear and robust guidance of users' engagements with tools, and for tools to account for a landscape where responsibilities for data practices and their AI applications increasingly converge.

1 Introduction

Across different sectors and domains, organisations have posed ubiquitous solutions focused on how to 'operationalise' data and artificial intelligence (AI) ethics. This is complimented by a surge in academic literature focused on the same objective. Such efforts aim to address concerns that decision-makers struggle to interpret and effectively address their responsibilities in their practices with data (Hirsch et al. 2024; Morley et al. 2020).

As part of this, data ethics frameworks (DEFs) have emerged as a practice-oriented tool that "offer instructions on how to identify and mitigate the ethical risks posed by [data] systems across key normative values such as fairness, transparency, and accountability" (Qiang, Rhim, and Moon 2024). DEFs have been developed in the context of a wider range of similar tools, such as responsible or ethical AI and data playbooks, guidelines, checklists and codes, that seek to achieve a similar function in guiding practical decision-making, with often overlapping design features (Prem 2023). The landscape of data practices is rapidly evolving with the changing landscape of computing technologies, and particularly the explosion of interest in use of AI. Data practices play a foundational role in the systemic challenges emerging from these applications of AI (Sambasivan et al. 2021). As such, the landscapes of data and AI ethics and governance approaches are highly intertwined (Khan 2025).

Data and AI ethics frameworks and other similar tools have been published by academics (Franzke, Muis, and Schäfer 2021), tech companies (Google 2023), (local) government departments (NSW Government 2024), non-profits focused on technology (Nethope 2020), and even private companies not focused on tech (Edquist et al. 2022). Investment in such approaches is continuing. For example, the UK government is currently in the process of updating their Data Ethics Framework (GDS 2025), which they have advised public servants to apply in their work since 2018, to help them to "use data appropriately and responsibly when planning, implementing, and evaluating a new policy or service" (GDS 2020).

Despite this continued investment in data and AI ethics frameworks, there has been limited research investigating how they are used and their impact in real world scenarios (Berman, Goyal, and Madaio 2024; Figueras, Verhagen, and Pargman 2022; Qiang, Rhim, and Moon 2024). Critical literature has offered valid concerns about about the potential for simplistic artefacts or tools to distract from urgent needs for systemic changes to data and AI ecosystems, by narrowing the scope of consideration (Munn 2023; Wang and Blok 2025). Meanwhile, existing studies have demonstrated that research can help to establish expectations for the role of ethics tools, and differentiate tools from one another (Qiang, Rhim, and Moon 2024; Yildirim et al. 2023). We consider that further systematic investigation of the practical benefits and challenges of engaging with tools is useful to build a realistic assessment of their role in working towards more responsible data practices, and informing this debate.

In this study, we explore how the Data Ethics Canvas[1] ('the Canvas'), a popular data ethics framework published by UK-based non-profit organisation the Open Data Institute, has been used in practice since June 2021. Previous similar studies have explored the application of frameworks focused specifically on AI applications (Qiang, Rhim, and Moon 2024; Yildirim et al. 2023), or have been more narrowly focused on singular, immediate outcomes like user satisfaction or literacy from single interaction with tools (Franzke, Muis, and Schäfer 2021; Vakkuri, Kemell, and Abrahamsson 2019). This study offers a different perspective, exploring the role of a data-oriented tool, through an in-depth case study based on 22 interviews with tool users. We use our findings to explore how its specific design features led to different types of interaction with the Canvas, developing a holistic understanding of participants' objectives in using the tool, the nature of their engagements, perceived impact, and barriers to change.

This study contributes to an urgent agenda to better understand the practical role and impact of data and AI ethics assessment tools. We make the following contributions: 1) we identify key factors to consider in redesign of this and similar tools, highlighting a need to navigate trade-offs between *guidance* and *engagement* of diverse stakeholders in data ethics challenges; 2) we make recommendations for reframing such tools, proposing a need for designers to outline clearer instructions for tool users. We suggest that this should be developed from research and evaluation of the tools. This should aim to help users to engage in more meaningful use with sustainable impact, and prevent overestimation of the impact of ethical assessment activities.

2 Background

2.1 Understanding the Emergence of Data Ethics Frameworks

For the development and deployment of a single data-driven approach or technology, decision-making about data practices is diffused across many different actors, including many mundane tasks (Leonelli 2016; Sadek et al. 2024; Sambasivan et al. 2021). This includes various points across dataset creation, governance, selection, analysis, and potentially further model training tasks, evaluation, and ultimately implementation (Longpre et al. 2024). Decisions that may appear small can have major downstream impacts. For example, over-reliance on poorly representative datasets in modelling, which can emerge from a lack of domain expertise in those making "data decisions" (Sambasivan et al. 2021), can lead to model applications with real world impacts, including harms like misdiagnosis of disease.

Agendas for 'responsible AI', 'AI ethics', and 'data ethics', among others, all seek to address practitioners' challenges in understanding and mitigating impacts from data and AI applications (Hirsch et al. 2024; Morley et al. 2020). These agendas are by no means new. However, industry, government, the third sector and academia have all responded to public urgency related to the increasing impact of data-driven technologies, and particularly AI systems, with a surge of proactive efforts in recent years. A 2019 article by authors from non-profit Data and Society (Metcalf 2019) quipped that "ethics is arguably hottest product in Silicon Valley today" (p. 449). While this tech landscape continues to evolve, a canon on responsible data and tech practices is emerging.

As part of these efforts towards more 'ethical' practice, organisations across all sectors have produced tools and interventions that aim to support decision-makers across different points of the 'lifecycle' of an AI system or data project (Lu et al. 2022; Morley et al. 2020). These include organisational approaches like committees or boards (Sorell 2024), training courses (Franzke, Muis, and Schäfer 2021), but also project or product level tools like frameworks (Hollanek 2024), toolkits (Holstein et al. 2019), audits (Schiff, Kelley, and Camacho Ibáñez 2024), impact assessments (Bogucka et al. 2024), and many more. There are also industry level approaches that aim to standardise and improve practices across many different organisations, like standardised data and model documentation approaches (Micheli et al. 2023). These different tools can be understood as interventions which are "designed to influence the norms, decisions, and daily practices of [decision makers working with data and AI] to mitigate the potential for [data and AI] systems to introduce or exacerbate societal inequities" (Berman, Goyal, and Madaio 2024, p. 1).

Those assessing the implications of AI systems often turn to to these tools as 'solutions' to the challenges of responsible technology development and deployment. For example, in November 2023, the UK government launched an official committee review in response to public survey findings demonstrating "distinct distrust in the regulation of advanced technology", a key recommendation of which was to develop a new guidance framework for "Ethics, Transparency and Accountability Framework for Automated Decision-Making" for use in the public sector (DSIT 2023). This turn to such solutions has been termed a 'toolification' of ethics by some researchers (Hollanek 2024).

2.2 How Have Data Ethics Frameworks Been Studied So Far

Critics of frameworks have argued that they can give a false sense that data and AI ethics can be limited to a step-by-step process or box ticking exercise (Hollanek 2024), which is problematic given the subjective and contentious nature of ethics claims. For example, Munn (2023) asked: "what is fair and who gets to decide it?" (p. 873), further arguing that "this is difficult work which requires engaging with social and political questions and prototyping, testing, and rejecting different designs: there are no shortcuts" (Munn 2023). Further, frameworks are often designed with an idealised understanding of a 'data project' where responsibilities can be encompassed under the remit of teams and employees, which does not tackle the need for infrastructural and system change at the scale of organisations and ecosystems (Sadek et al. 2024).

Still, there *is* responsibility that lies at the level of individuals and organisations, and as such it is important to ex-

[1] https://theodi.org/insights/tools/the-data-ethics-canvas-2021/

plore the role of interventions that address these responsibilities. Despite their pervasiveness and ambitious claims, there is limited research to understand the way that such frameworks are used in practice and their impacts on data work on decision-making (Figueras, Verhagen, and Pargman 2022; Vakkuri, Kemell, and Abrahamsson 2019; Yildirim et al. 2023; Qiang, Rhim, and Moon 2024).

Although relatively little is known about the practical adoption, use and impact of data and AI ethics frameworks, various studies explore ethics frameworks as artefacts that represent the current narratives and scope of the organisational ethics work they intend to instruct. They address limitations such as the critiquing the scope of issues and principles addressed (Greene, Hoffmann, and Stark 2019; Jobin, Ienca, and Vayena 2019). For example, authors argue that frameworks often fail to: prompt consideration of whether a given project should be taking place in the first place (Hollanek 2024); highlight systemic justice issues including labour rights and environmental sustainability (Corrêa et al. 2024; Hollanek 2024); and address wider conceptions of stakeholder engagement beyond involvement of expert professionals (Greene, Hoffmann, and Stark 2019). Particularly within academic domains, such insights have led to emergence of redesigned frameworks such as those focused on the 'ethical demands' of specific domains, addressing specific regulatory demands (Constantinides et al. 2024), are more focused on justice or feminist principles (Varon and Peña 2021), or look specifically to integrate sustainability considerations (Samarawickrama 2022).

A relatively small number of studies have explored users' experiences of implementing data and AI ethics frameworks, following the wider trend across the landscape of data and AI ethics tools of an absence of evaluation studies (Berman, Goyal, and Madaio 2024). This means that few existing frameworks are evaluated. Existing studies on their implementation typically investigate the immediate subjective outcomes such as 'improvements to literacy' from singular experimental interactions with tools - as opposed to real-world interactions of participants with specific tools (e.g., (Constantinides et al. 2024; Franzke, Muis, and Schäfer 2021; Qiang, Rhim, and Moon 2024; Vakkuri, Kemell, and Abrahamsson 2019).

Further, as Qiang, Rhim, and Moon (2024) concluded, 'one size does not fit all', meaning that different 'ethics' frameworks will provide different benefits and challenges to their users. Given that most data and AI ethics frameworks are not evaluated, and offer broad remits for their use, Qiang, Rhim, and Moon (2024) further argue that the "suitability and expected benefits of existing frameworks" should be better clarified "for better adoption of [data and] AI ethics in the industry" (p. 1976), through in depth investigation on the use of single tools. Therefore, research investigating adoption and use of tools in practice, rather than under controlled experimental settings, is important to understand the nature of their role in real world settings. This should include exploration of the types of users engaging with ethics frameworks, applications they are applied to, tasks that users aim to achieve and how far user needs are met, and approaches to evaluation and iteration. These qualitative investigations can develop the basis for more targeted evaluations, through establishing aspects such as the common 'users', 'situations', 'objectives' of a tool, needed for more robust effectiveness studies (Berman, Goyal, and Madaio 2024; Olsen 2007; Qiang, Rhim, and Moon 2024). While a recent study (Yildirim et al. 2023) has offered case study exploration into the use of a human-AI guideline published by Google, we have not identified similar explorations focused on the use of *data* ethics frameworks.

3 Methodology

We completed a retrospective qualitative study, conducting 22 semi-structured interviews with individuals who have used the Data Ethics Canvas (the Canvas) in the period since June 2021, when the framework was last updated. This research approach was approved by the Research Ethics Committee at King's College London.

3.1 Selected Case Study: The Data Ethics Canvas

The Data Ethics Canvas was published in 2017 by the Open Data Institute (ODI), a UK-based non-profit organisation. Over the past 8 years, the Canvas has undergone several minor updates but has remained similar in format and content. The ODI recorded 3,525 downloads in 2024, with the majority in English, and between 150 and 500 downloads in each of French, Spanish, Italian, Dutch and German. The tool is cited within many repositories of similar tools[2], within public sector governance guidelines and strategies (Met 2021; LOTI 2022) and has been used by non-profits (ODI 2024) and large private sector organisations including Barclays Bank and Co-op (ODI 2021; Co-op 2019).

We selected the Data Ethics Canvas for study due to its popularity across multiple sectors, its international orientation, and our specific access to study the tool.

The Canvas is a single page PDF segmented into four key areas for ethical practice (while not explicitly named as such, these could be entitled 'data sources', 'people', 'understanding impact' and 'identifying mitigations'), and containing lists of questions, which act as prompts for organisations' thinking, under each area of practice (see Figure 1).

The approach of 'prompting' considerations rather than offering explicit instructions is a common practice within data ethics frameworks (Häußler 2021). The Canvas' Creative Commons license of the tool allows others to use, adapt, and build upon the work, so long as the organisation that designed the tool is referenced. Like many similar tools (Qiang, Rhim, and Moon 2024), the Canvas gives minimal instruction over the intended context its use, explicitly stating that it can be used by anyone involved in working with data at any stage of their work, and across all different types of work with data. While there is a user guide to support organisations with the use of the tool, we chose to focus exclusively on the Canvas itself as many of the known applications of the tool do not include the use of the guide.

We have outlined numerous aspects that vary among different tools. Other differences include: whether they are generic or specific to a certain application or context

[2]e.g., https://oecd.ai/en/catalogue/tools

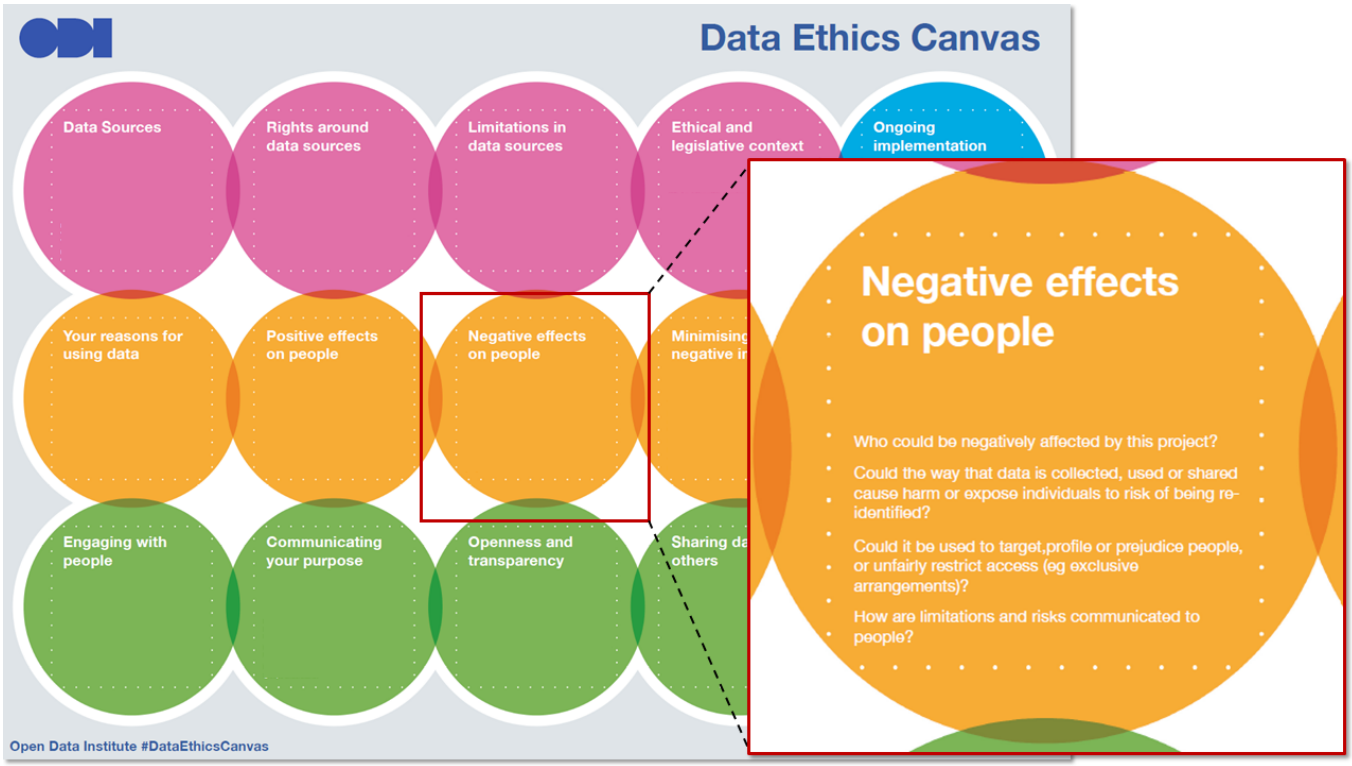

Figure 1: The Open Data Institute's Data Ethics Canvas, updated version published in June 2021. In each circle there are between 2-5 questions (see magnified image). The tool is designed to be viewed at a scale larger than A4 or digitally. Questions listed within the Data Ethics Canvas and copyright information are available in Appendix.

(Solanki, Grundy, and Hussain 2023; Tirado, Mulholland, and Fernandez 2024); the simplicity versus density of the tools; the context of the organisation designing the tool; how individualistic versus collective conceptions of responsibility are, among other facets. This means that while we study the design features of this one particular tool, there is scope to explore in further research how far these findings generalise or differ across the wider landscape of tools, by comparing these findings to examinations of other tools.

3.2 Sampling

We identified participants via purposive and snowball sampling (Parker, Scott, and Geddes 2019), aiming to generate a diverse sample of participants from organisations across different sectors, sizes of organisations, application domains, and uses of the tool. We first contacted staff members at the ODI who were known users of the tool, who mostly used it to support other organisations via functions such as training and consultancy. To identify further users, we asked each participant to share recruitment information with other users of the Data Ethics Canvas, and additionally conducted a search on Google, Twitter, LinkedIn, and Google Scholar using the keyword "Data Ethics Canvas" to identify all online publications or communications that mentioned the specific tool. Authors of these online publications were contacted directly by the lead author with recruitment information.

To achieve this sampling approach, the Open Data Institute acted as a 'gatekeeper' organisation, sharing recruitment information with potential participants. This role was important because, as empirical research has demonstrated, gaining insight into private organisations' internal data practices can be challenging (Scheuerman 2024; Collett 2024), and this likely helped to generate sufficient trust for tool users to share their experiences.

Participants were recent or present users of the tool, meaning they had used tool as a basis for their own decision-making or supporting decision-making of other organisations over the past four years. Each participant completed their work at a different organisation, except for seven participants who were employed at the ODI.

3.3 Interview Approach

Semi-structured interviews of 30-60 minutes were completed between January and April 2025. They addressed descriptive elements (who, what, where, why, approaches to iteration and evaluation) about their use of the tool, but also qualitative elements (their perceptions of success and limitations in their own work). We did not ask participants to speak about specific project details, and we anonymised the names of any specific projects or clients that were mentioned. Discussions instead focused on understanding the role that the tool played in mediating this work. This was important as organisations are at times reluctant to share specifics of their decision-making within projects due to concern about repu-

tational risk (Scheuerman 2024), yet the approach still allowed participants to share rich insights on their experiences.

3.4 Analysis

We used a reflexive thematic analysis method (Braun and Clarke 2021). A single author coded the transcripts, which were transcribed verbatim. 136 codes were generated inductively, and grouped into themes and subthemes via further analytic discussion with co-authors.

4 Findings and Discussion

Of the 22 interviewees, seven were from the Open Data Institute (non-profit), two were from other non-profit organisations, eight were from private organisations and four were from public sector organisations. We identified more prospective participants from private than non-profit organisations, although four private sector participants used the Canvas in training or consultancy roles to support public sector organisations using the tool.

As a result of the sampling approach, participants tended to be champions of use of the tool within their organisations, in that they were often those involved in procuring or leading data ethics work and were thus typically in middle or senior roles. Participants were largely consultants, ethics trainers, project or product managers, leads working on the implementation of data projects or tools, or leads directing technical teams and projects.

Further, 12 participants shared their experiences of supporting other organisations to use the tool, 10 participants shared their experiences of using the tool for their own internal practices, and one had experience of both. We have not specified which seven participants in the study used the tool as part of their role at the Open Data Institute, to avoid re-identification of individuals.

In this section, we outline and discuss three central findings about how participants have made use of the Data Ethics Canvas, sharing critical reflections about the implications of these findings. We found: (1) that a core function of the tool appears to be in *engaging* users in the scope, practices, and terminologies associated with data ethics, (2) that when applied within project settings, the tool primarily supports *assessment* and *agreement* of the ethical demands, to different degrees depending on the circumstances in which the tool is applied, and (3) that the tool is used beyond the project level, as a mediator for building *organisational data ethics maturity*, suggesting that this is primarily, but importantly not limited to the earlier stages of maturity.

4.1 Theme 1: Engaging With Data Ethics Practices

The Open Data Institute's website states that the Data Ethics Canvas "can help you to identify and manage ethical issues", which is a broad and ambitious objective. We identify a more specific function of the tool, in *engaging people* with data ethics practices. By bringing people in to discussions and activities focused on data ethics, the tool is then useful for a set of wider activities at the project or organisational levels, discussed in Sections 4.3 and 4.4.

Here, we describe two main ways the tool supports this engagement, by: (1) helping to define the scope of data ethics *"concepts and practices"* (P7), and (2) providing an *"artefact"* (P19) to form the basis of this engagement.

A Basic Definition of Data Ethics Practices The Data Ethics Canvas defines the scope of data ethics consideration. Many participants described the tool as offering an introduction to the broad 'lens' of data ethics. They highlighted decision-makers within their organisations, or within client organisations *"not knowing what data ethics is"* (P5), or needing support to make existing queries or discussions about the possible implications of their work *"concrete"* (P21, P22). For this, the tool provides shared, non-specialist language (P14, P21) and *"is there to guide you to ask the relevant questions"* (P21).

However, while many tools are more detailed, more specific to certain types of application, for example the Human-AI Guidelines examined in a study by Yildirim et al. (2023), or offer more specific structure for how they should be used, the Data Ethics Canvas offers a relatively simple and versatile framing. Participants describe seeing the tool as *"agnostic"* (P5) or *"basic"* (P6) in its definition of data ethics practices. This manifested in the understanding that it is non-specific to different philosophical frameworks (P2, P7), cultural values (P4), and motivations for engaging with data ethics work (P5, P9): *"I think what [the ODI] want the Data Ethics Canvas to do is guide people through their own ethical framework. In the sense of 'what is important to you, and are you doing that thing that is important to you'. Without telling them what's right and wrong."* (P4).

Further, participants suggested that the format of *questions* rather than instructions within the Canvas was important for avoiding *"running the risk of imposing anything on partners. So it's not saying you definitely need to do this or you definitely need to do that. By asking questions it still allows people to find their own approach to it. But concrete"* (P19).

Still, even in setting this 'basic' scope of ethical consideration, participants often alluded to the fact that they are using this tool as a benchmark to set the *standards* for appropriate practice. This included the role of the tool in creating legitimacy for addressing ethical considerations. For example one participant mentioned the tool in conversation with other stakeholders, saying: *"when they are asking why we are questioning things, we can point them to [the Data Ethics Canvas]. We're not just making this stuff up"* (P18). Other participants (external to the ODI) were more explicit about the importance of the institution behind the tool in creating this legitimacy *"the brand of ODI as a non profit helps to reassure people, particularly in your smaller enterprises and startup world"* (P16). In all, participants suggested that simple, adaptable, and, for some, reputable *content* of the tool was important for its ability to form a basis for engagement.

An Accessible Anchor for Engagement Participants made clear that the utility of the tool went beyond the in-

Sector	Public sector, Private sector, Non-profit
Type of data project	Data analytics, Data sharing or portability, AI/prediction applications, Synthetic data development
Domain	Healthcare, Finance, Environment, Energy, Public infrastructure, Public services delivery

Table 1: Summary table on applications of the Data Ethics Canvas by interviewees - sectors, types of data project, domains.

formation provided in its contents. Instead, value was also derived from the physical form of the tool as an *"artefact"* (P2), *"anchor"* (P7) or *"hook"* (P7) to generate engagement with the data ethics lens. The layout of the tool appears central to its capacity to engage participants with data ethics, and support consideration of its role within their contexts.

Multiple participants specified the importance of the Canvas being on a single page (P11, P14, P21). This was noted as important for facilitating a lower investment for engagement, as compared with other similar forms of tool: *"[A] playbook, just by virtue of being called a book, I'm already put off because it's a lot of reading and a lot of work. The canvas is basically an off the shelf, one-page solution and that is what makes it efficient"* (P11). Other aspects of the format that participants commented on were similarly oriented around the tool as visually appealing due to its colourful layout, and simple structure. As one participant noted: *"You've got the main concept, the main themes, it's colour coordinated, so it's easy to follow"* (P21).

The nature of structure in the tool was a further subject of debate. Participants tended to acknowledge that the tool offers structure, in the form of four colour-coded 'subsections' and 15 'bubbles', each containing a heading and brief set of questions. However, the Canvas does not give a clear indication of which order to engage with the different components, leaving this largely up to interpretation for the user. This fact appears to contribute to its versatility: *"you can take as much or as little because of course not everything on on here will be relevant to whatever you're doing"* (P21).

Several participants also discussed design/visual drawbacks of the tool, including small text making it less accessible (P3, P21), and a desire for a digital format that makes it easier to interact with the outputs of the tools (P3, P11, P15). Both of these features offer challenges for designers to tackle meanwhile maintaining its benefits for engagement.

Discussion of Theme 1 Findings The benefits that participants highlighted for the tool's role in engagement closely align to barriers for data and AI ethics practices recognised in wider literature. For example, previous research examining these barriers highlights a lack of understanding or 'concreteness' of ethical responsibilities in work with data (Morley et al. 2020; Khan et al. 2023), a need for legitimacy in asking challenging questions about data or engaging in 'ethics assessment' (Widder 2024), and a desire to introduce these considerations in a way that is low-resource due to low existing prioritisation of these factors (Schiff et al. 2021). These findings suggest a potential role for the tool in mediating some known barriers to responsible practice, by playing a role in *brokering* different perspectives and needs on responsibilities in uses of data. We propose that various design features of the tool appear to support its function in engagement, and that if the Data Ethics Canvas was less aesthetic, more confrontational, or more complex, it may have been less functional in achieving engagement among stakeholders.

These demands suggest **a potential trade-off between guidance and engagement, that tool users and designers should be aware of**, where more prescriptive and comprehensive approaches may have gained less buy-in from users. However, this trade-off also touches on contentious aspects of such tools.

As discussed in Section 2, critics argue that 'toolkitification' (Hollanek 2024), the reliance on simplified tools, suggests ethics can be 'solved' on an organisational level. Given the clear need for systemic changes that cannot, and will not, be addressed at the scale of individuals and teams (Crawford 2021; Wang and Blok 2025), such approaches must not drive attention away from the need for mandating responsible practices on greater scale. 'Basic' tools such as the Data Ethics Canvas will, critics point out, determine and limit the scope of ethical consideration by pointing attention towards or away from certain values or facets of what is 'ethical'(Hollanek 2024). In practice, it appears more appropriate for such tools to be used as a starting point for interpreting the scope of responsibility, rather than an encompassing guide, which would require far greater care and deliberation.

We explore further in Sections 4.2 and 4.3 the role for this 'engagement' when the tool is used at project and organisational levels, and the further scope for guidance of tool users.

4.2 Theme 2: Project Level Uses - Assessment

The Data Ethics Canvas is intended to be used within 'projects'. This is outlined on the website page hosting the tool[3]. The tool is structured around prompts for both "identifying and managing" the "adverse impacts on people and society" from "projects that use data". The majority of participants we spoke to had used the tool within their own projects, or supported other organisations to do so.

Broadly, participants reported that engagement with the Canvas tended to support an improved understanding, or 'assessment' of the ethical demands of the project. However, the extent of this understanding, and the follow up actions to address these demands, varied significantly with the role of mediators such as expert facilitators and organisational champions, as well as other factors including the domain of the application. Their role was pivotal since: *"organisations are way more nuanced than you can see from the outside, and so the approach needs to be tailored, ideally by someone who works within the organisation"* (P2).

[3]https://theodi.org/insights/tools/the-data-ethics-canvas-2021/

Identification and Agreement of Ethical Demands Participants all suggested that the tool had helped to identify at least some ethical demands associated with projects, and actions to mitigate these. For example: *"with four or five different questions [in each section of the Canvas] we may have considered one or two aspects, but not the rest"* (P13).

A few participants use the tool in ad hoc and light-touch ways, and not in collaboration with others, for project assessment work: *"really as a reference point to plug any gaps at provoking questions"* (P16). However, the majority of participants use the tool in more structured, collaborative and in-depth ways, frequently in a workshop format. Deliberation of the prompts among different stakeholders was crucial to their sense of success in this ethical assessment.

We identified two reasons for the importance of this deliberation: firstly, due to the need for diverse input to understand the challenges and likely solutions; and secondly due to a need for *agreement* of the issues and next steps: *"we want everybody going out the room to feel like they've made a consensual recognition of some of the challenges and that they all agree that these are the good ways to go."* (P1). Participant 14 explained the importance of including *"the quiet people on the front line"*, who may be able to identify different challenges than other team members. The vast majority of participants described using the tool only with individuals *internal* to the organisation undertaking the project, although in a couple of cases participants had examples where some external stakeholders were included (P3, P5). Some participants suggested that the tool changed perspectives of those otherwise not engaged with 'ethics': *"we had very technical people that were there who are only interested in the data and building the tools. They were not interested in ethics at all. And so with the exposure to the Canvas, they were like, OK, these are things we need to start thinking about when we're doing the data analysis and presenting and visualising the data, et cetera"* (P13).

Multiple participants suggested that 1-3 hour workshops had been insufficient to build a suitable assessment of a project - and particularly to be able to identify the necessary follow up actions (P3, P10). Participant 7 specified that it *"depends how much time you spend with it as to whether you get any granular detail or action off the back of it"*. Further still, many identified the importance of specific efforts made on the part of organisational 'champions' of the tool, or by external facilitators, for generating an effective assessment. This included contextualisation of the tool for the situation. For example participant 7, who had multiple years of experience deploying the tool, highlighted that they: *"got rid of [sections] or incorporated them in different ways, or held them in my back pocket for a discussion piece"* - while others adapted the questions for the domain, including introducing specific case studies and context specific language (e.g., P2, P8, P11).

In all, while the versatility of the tool may support the engagement of diverse users, there is also ambiguity over *how* to engage with it. Notably, responsibility for ensuring in depth and effective engagement formats tends to fall to those facilitating and championing interaction with the tool.

Ad Hoc Implementation of Follow Up Actions Although participants mostly presented optimism over the role of the Data Ethics Canvas in supporting *assessment* of ethical demands within projects, many were concerned about whether these would be addressed: *"it's valuable to give you a list of things to think about at specific stages within a project. But beyond that, it's very difficult to then understand what you need to do to action that if you're not someone who has the specific knowledge of how to do that thing"* (P2). Those facilitating workshops with the tool were often unclear over whether actions had been taken following the sessions or even whether the workshops had succeeded in establishing what these follow up actions should look like. Further, across participants there was a lack of *formal* evaluation of how far use of the tool had led to change.

However, some participants described different ways that the use of the tool had influenced projects. These included:

- Preventing projects going ahead - demonstrating why a certain project is too risky (P18, P22)
- Allowing projects to go ahead - allowing participants to recognise that there are means to mitigate their concerns (P1, P2, P6, P11, P12)
- Adding elements to a risk register as a project progressed (P8)
- Generating an assessment of the project to inform senior decision makers (P21)
- Generating a written set of challenges that are then referenced throughout project decision-making (P20, P21)
- Building a set of recommendations from a consultant about how to progress with a project (P1, P3, P5).

Developing these actions does fall under the remit of the Canvas, which includes questions such as "Are you routinely building in thoughts, ideas and considerations of people affected by your project?" and "How will data ethics issues be monitored, discussed and actioned?" However, as one participant put it: *"it tells you what is a potential risk or harm [but] it doesn't link to the enforcement work"* (P11). Ensuring that action is taken requires deep organisational expertise, and decision-making authority. As such, engagement needs to reach those who have such authority.

Applying Data Ethics Practices for AI Applications The tool was used across diverse application domains (see Table 1). Across these domains, participants discussed whether aspects of the Data Ethics Canvas were 'relevant' for their specific project - often suggesting that this led them to 'pick and choose' those that they saw as more relevant: *"a lot of the work we do is around [a specific form of] environmental data. So all of the stuff involving personal privacy isn't relevant"* (P16). At times, as already discussed, participants also spoke about adding further context to help participants relate the tool to language or case studies they are familiar with.

However, the applicability of the tool appeared to generate particular challenges for participants who were looking to use the tool for AI use cases (including P4, P8, P10, P11 P12, P19, P20, P21). Participants typically identified

their uses of data in AI applications as being within the broad remit of the Canvas, and commonly discussed how their roles had transitioned from primarily overseeing data analysis or stewardship to more emphasis on AI applications within the same role (e.g., P12, P19, P21). Multiple participants expressed uncertainty over how far the tool would cover their data-related concerns within AI contexts. This included highlighting specific data-oriented ethical concerns that were missing from the tool: *"there's nothing around environment, for example [...] which is a big ethical concern at the moment, particularly around AI models"* (P8).

Some participants (e.g., P10, P16, P19) recognised that they *"would also need some other other tools in in the during the AI life cycle"* (P10). However, a focus specifically on *data* ethics is also important for organisations stewarding datasets that may be used within AI applications, or for those completing *"the constant activities people need to be doing to have a good pipeline of high quality accurate data that can inform the AI models that people want to design and build"* (P16). This raises questions over how far existing AI ethics tools support data-centric considerations, including data stewardship for AI applications.

Discussion of Theme 2 Findings While participants were largely positive about their engagements with the tool, they described variable outcomes, in terms of the depth of ethical assessment and how far this led to follow-up actions. Their outcomes appeared to be highly dependent on factors including the time teams spent with the tool, stakeholders involved, and the level of adaptation of the tool and approach to its use, as well as other organisational factors. Wider discourse on the role of data and AI ethics tools highlights pressures placed on specific organisational actors for carrying the burden of driving responsible practice (Widder 2024; Gambelin 2021). Our findings corroborate this, and demonstrate the requirement for decision-makers to have authority and expertise to support effective engagement.

We support the proposal made by (Qiang, Rhim, and Moon 2024), that **clarifying expectations of different tools is of key importance for their ongoing design and use**. This should include highlighting the role of mediators and experts in tool use, and outlining expected outcomes from different types of interaction with the tool, to support users in allocating appropriate resources. Setting the expectation to produce clear, evidenced guidance to accompany tools may also go some way towards preventing their use in ineffectual efforts towards responsible practice, where such tools can act as 'false solutions' (Munn 2023).

The findings also demonstrate that **there is a need for tools to be adapted or adaptable to contextual demands and to keep up with changes in the data and AI landscape**. While individuals often put in great efforts to make adaptations themselves, exploring design options that better support updates and adaptations would help to formalise this process. Further, we identify ambiguities in how 'data ethics' and 'AI ethics' converge with one another in the practical settings that participants encounter. Those overseeing the use and sharing of data increasingly need to be aware of possible uses in AI applications, as these are increasingly common. On the other side, those primarily responsible for AI deployment are increasingly challenged to address their responsibilities to incorporate appropriate, high quality and well-governed datasets. Widespread concerns that data practices are often overlooked in the context of AI development and deployment (Aroyo et al. 2022; Sambasivan et al. 2021) give cause for further consideration of how far guidance over data ethics and governance is prioritised within AI ethics or RAI tools, and how specific data oriented tools are needed to complement these.

4.3 Theme 3: Organisation Level Uses - Improving Data Ethics Maturity

A notion echoed by many participants was that data ethics needs to be integrated into standard organisational decision-making. For example, participant 2 argued: *"For organisations to sustainably think about data ethics, it needs to really be built into the way that you work and the way that you operate, and it needs to sit within broader functions like data governance, for example, because otherwise it gets lost or it will end up being de-prioritised or de-funded"* (P2).

Beyond the use of the Canvas at a project level (Theme 2), many used the tool to generate improved data ethics literacy or skills - and as a means to inspire or provide an argument for stronger data governance initiatives. Multiple participants alluded to or directly described these organisational uses of the tool as attempts to improve the maturity of the organisation in their data ethics practices.

Some participants (P15, P16) spoke about this maturity as a spectrum, with introductory engagement at one side, which this tool was better designed to support, and a strong suite of governance controls at the other. Other participants from the Open Data Institute referenced their Data Ethics Maturity Model (ODI 2022), highlighting that the organisation incorporates the tool within this model: *"The data ethics maturity exercise [...] [is] about looking at organisationally how you are fit for applying data ethics, and the Data Ethics Canvas is simply one of the tools you can plug into that organisational process"* (P1).

Improving Organisational Literacy Some engaged with the tool for capacity building, to support future (potential) engagements with the tool, or to generate a culture change.

Multiple participants spoke about engaging with the Canvas as a means to understand how it can be used in future work (e.g., P10, P8), with two participants mentioning their informal efforts to build an *"arsenal"* (P16, P21) of different tools that can help to build robust data governance, including the Data Ethics Canvas.

Participants spoke about using the tool in combination with a wider range of other complementary tools, some of which are also delivered by the ODI and offered different benefits, for example 'ecosystem mapping', 'maturity assessment', 'risk assessment', 'consequence scanning', governance boards, and more. These each offer different benefits in terms of offering more detailed guidance, determining which stakeholders you need to involve, more emphasis on governance, and project auditing, among others. For example, some (P10, P12, P19) mentioned the role of ecosystem

mapping, using the ODI's Data Ecosystem Mapping tool, to first identify project stakeholders, before engaging these individuals in a data ethics workshop using the Canvas.

Many participants engaged with the tool as a learning device for understanding the data ethics 'lens'. For example, one participant spoke about using newspaper articles within training based on the tool, to help partipants understand how ethical problems arise. They used the tool to demonstrate, about one government use of an algorithm, that *"where they were going wrong is really in the transparency area because their process sounds like it's good"* (P4). Some participants emphasised its role in differentiating between the need for legal compliance and 'ethics' (P7, P21), with participant 11 saying that they found that there could be misconceptions over how far working within the law would prevent harms.

These types of literacy-oriented practices are appropriate for earlier stages of maturity. By contrast, participant 16 described how the emphasis within their role was increasingly focused on creating formalised risk assessment, controls and procedure, and suggesting that they are at a higher level of maturity. As such, they argued, they were no longer in need of emphasising the *"fun discovery kind of activities"* (referring to the Data Ethics Canvas as an example), although they indeed still used them as a light-touch reference point.

Inspiring Data Governance Approaches Two participants used the tool as inspiration for their own adapted frameworks that included mandated follow up stages. For example, *"we created a workshop that uses the individual canvases that we've created and that the product manager can use them to check on how have people thought through all of this stuff"* (P20). Another participant (P22) used the tool in a workshop format to demonstrate to senior members of the organisation the value of introducing a data ethics lens, and more formalised data ethics activities, to complement existing governance interventions. This then inspired a range of mandated 'ethics' activities across the team, with stringent requirements to act upon identified concerns.

Participants also mentioned how these types of embedded governance approaches rested on the *existing* resources and motivations of the organisation: *"it was always a case of, public sector and nonprofits were more engaged, less likely to embed. If you got a highly engaged private organisation, they were more likely to embed"* (P5).

Discussion of Theme 3 Findings Concerns about the ability of the tool to support long-term change to organisational decision-making were in some cases tackled through building use of the tool into organisational governance processes. Such an approach could be seen as analogous to the use of impact assessment inscribed into policy-making, like the requirements for use of Data Protection Impact Assessments under the General Data Protection Regulation (GDPR). **Building ethics assessment into mandated organisational processes offers more scope for sustainable practices**, and tool designers can explore avenues for guiding organisations to do so. Further, requirements to engage with ethical deliberation increases the importance of addressing the core limitations of tools to support this.

Identifying which ethics interventions should be used and when is increasingly the subject of research attention (Longpre et al. 2024), and **there is scope for development of further robust evidence on how different tools support different forms of change** (Berman, Goyal, and Madaio 2024), including methodologies similar to this study. From the findings in this study, it is evident that factors including existing organisational 'maturity' in data governance, motivation for responsible practice, the involvement of 'experts' or champions mediating tool use, and extent of resources available all influence the nature of engagement with the tool. Such factors may therefore be considered when deliberating best practice for data ethics assessment and governance practices.

5 Recommendations

Our findings offer insights on how organisations have used ODI's Data Ethics Canvas in practice. Overall, we have identified organisations' use of the tool as a means for brokering and aligning different stakeholders on their ethical responsibilities with data. We further outlined three ways that this manifested: its use for legitimising and creating an entry point for discussions about data ethics (Theme 1); its use in consensus-building about ethical concerns and feeding into project recommendations and decisions, to varying degrees (Theme 2); and its use for influencing the standard practices within organisations through literacy and governance (Theme 3). However, we also discussed concerns associated with tool use, including: its potential to over-simplify complex ethical challenges; the reliance of successful ethical assessment and follow-up on organisational factors and efforts of specific individuals; and the need for clearer pathways to sustainable change from tool use. We share key recommendations for translating these insights to research and practice.

Recommendation 1: Tool designers should hack intervention design using critiques of existing tools and leveraging features for engagement This research has demonstrated opportunities in the positive response to the tool. Some participants reported using the tool to motivate and align team members on data ethics challenges, due to its simple and versatile contents and form. Enthusiasm can be used as a lever, and tool design considerations should aim to address *engagement* of users. Such opportunities should not be underestimated, but need to be balanced with a critical approach to intervention development.

Academic literature offers useful perspectives on the limitations of data and AI ethics frameworks and other similar tools that must be accounted for in best practices for tool design. For example, authors observe (Kawakami, Wilkinson, and Chouldechova 2024) that these frameworks and other similar tools emphasise organisational perceptions of ethical practice over perspectives of end users and other stakeholders. More work should be done on assessment of best practices to bring meaningful stakeholder engagement into typical data decision-making contexts (Kallina, Bohné, and Singh 2025). Our findings identify additional limitations including the need for frameworks to better navigate the converging fields of data and AI ethics.

Recommendation 2: Tool designers should focus on guiding realistic expectations of tool use We highlight that

the goals described for the Data Ethics Canvas, in the guidance accompanying the tool, are too broad. We agree with Qiang, Rhim, and Moon (2024) that creating more specific expectations of outcomes from interactions with tools is essential, and further argue that this should form the basis of improved guidance accompanying the tools. This is for two key reasons: to guide pathways for meaningful use and evaluation of existing interventions; and to prevent such tools distracting from systemic changes that are needed. For example, guidance could include specific insights on 'success metrics' and how to assess these, formats for engagement with the tool and resources required for these, and examples of how organisations have generated sustainable change.

Additionally, we outline a need to explicitly reframe ethics artefacts as limited tools in need of adaptation and contextualisation, conferring responsibility back onto users and designers to build on the frameworks continuously. Such guidance can address warnings about misinterpretations of such tools as more comprehensive 'solutions' to complex challenges in data ethics and governance (Sadek et al. 2024; Wang and Blok 2025). Exploring design features that make it easier for users to adapt tools to their contextual demands may be of use for supporting this re-framing.

Encouraging simultaneous use of diverse tools may also prevent dangers in creating narrowed perspectives and competing schools of thought on the scope of 'ethical practice'. While some did interpret the Data Ethics Canvas as a 'basic' tool, to be adapted to their own organisational context, valid concerns remain about the ability of single organisations to set even the basic terms of ethical consideration.

Recommendation 3: Researchers should proactively aim to support the development of better guidance through building robust evidence of tools' impacts Within the emerging field on evaluation of tools for responsible data and AI practices, authors highlight that identifying specific objectives for tool use can underpin experimental studies to understand their effectiveness (Berman, Goyal, and Madaio 2024). These studies can also explore impacts across different types of users and contexts. Such research should seek to contribute towards best practices for intervention development and use (Kazim and Koshiyama 2021). This present study has offered specific objectives for users of the Data Ethics Canvas, including the use of the tool in motivating or engaging unmotivated stakeholders, generating formalised 'ethical assessments' that are incorporated into official project recommendations and governance processes, and the creation of repeatable or mandated requirements for ethical consideration. Further research should continue to investigate the real world impact of such data and AI ethics interventions. This should seek to include further qualitative and in-depth investigation to understand the nuances of practice, in addition to wider effectiveness studies. Future work should also include qualitative assessment of how far ethical deliberation among project teams matches the concerns of wider stakeholder and advocacy groups, which will further add value in understanding the impact of tool use and the demand for stakeholder deliberation (Kawakami, Wilkinson, and Chouldechova 2024; Zilka et al. 2023).

6 Limitations

Two of the authors of this study are employed part-time at the non-profit organisation deploying the Data Ethics Canvas. However, they have not been directly involved in the design or deployment of the tool examined in this study, and their roles at the organisation were relevant in reducing barriers to data collection, as described in section 3.2. The selected analytical approach (reflexive thematic analysis) reflected the authors' recognition of their position in relation to the interviewees and the data collected, where participants likely have pre-existing trust of the organisation, but may also withhold some level of critique of the tool. Specifically, the interpretive nature of the analysis allowed the authors to critically consider the context in which reflections were shared and the resultant findings, broadly interpreting the participants as 'champions' of the tool.

Findings regarding its usefulness in motivating participants were in this study reported second hand by single study participants from each organisation. Thus, future research to investigate perspectives of (initially) less motivated team members may offer a greater insight into its utility for motivating and aligning different types of stakeholders.

Further, the study explores use of a single tool, and is therefore intended to offer exploratory findings to inform future research approaches and tool design considerations, as opposed to robust generalisable findings to be extrapolated directly across other tools and approaches. The participant sample is a low number of the wider population of tool users, and is intended to represent diversity among types of organisations by including participants from organisations of different sizes, from different sectors, and covering diverse types of interactions with the tool. As such, the approach aimed to develop a rich understanding of diverse ways the tool can be used and its potential, through in depth interview findings, rather than a representative perspective on the tool. This would be difficult to achieve due to the open nature of the tool, where adoption is not easily traceable.

Finally, study findings were based on recall by study participants, and emphasises the *use* of the tool over its *impact*. A longitudinal study approach may offer more robust findings on the experiences and practices of users, as well as for understanding effectiveness.

7 Conclusions

In this study, we have explored the use of the Data Ethics Canvas over the past four years, to demonstrate how it is being applied in real world organisational settings. We have found that the tool's versatility forms the basis of its wide use across diverse contexts, and its role in brokering perspectives on ethical challenges. However, there is a need to better guide interactions with the tool, and for redesign of this and similar tools to navigate trade-offs between engagement and guidance of users. Our findings highlight the value of rich, qualitative insights for understanding their impact on decision-making. We join calls for further research investigating the real-world impact of data and AI ethics tools, which we argue can feed into robust evidence to support their design and use.

Acknowledgments

This work was supported by the UK Research and Innovation (UKRI) Centre for Doctoral Training in Safe and Trusted Artifcial Intelligence [grant number EP/S023356/1[4]]. We would like to thank the Open Data Institute for their support of this study, and the participants who generously contributed their time.

We would also like to thank Ben Snaith, Mackenzie Jorgensen and Sinem Görücü for their valuable feedback on this research.

References

Aroyo, L.; Lease, M.; Paritosh, P.; and Schaekermann, M. 2022. Data excellence for ai: why should you care? *Interactions*, 29(2): 66–69.

Berman, G.; Goyal, N.; and Madaio, M. 2024. A Scoping Study of Evaluation Practices for Responsible AI Tools: Steps Towards Effectiveness Evaluations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–24. Honolulu HI USA: ACM. ISBN 979-8-4007-0330-0.

Bogucka, E.; Constantinides, M.; Šćepanović, S.; and Quercia, D. 2024. AI Design: A Responsible Artificial Intelligence Framework for Prefilling Impact Assessment Reports. *IEEE Internet Computing*, 28(5): 37–45.

Braun, V.; and Clarke, V. 2021. *Thematic analysis: A practical guide*. SAGE publications Ltd.

Co-op. 2019. Data ethics canvas: helping us make good data decisions from the start.

Collett, C. 2024. The hustle: How struggling to access elites for qualitative interviews alters research and the researcher. *Qualitative Inquiry*, 30(7): 555–567.

Constantinides, M.; Bogucka, E.; Quercia, D.; Kallio, S.; and Tahaei, M. 2024. RAI Guidelines: Method for Generating Responsible AI Guidelines Grounded in Regulations and Usable by (Non-)Technical Roles. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW2): 1–28.

Corrêa, N. K.; Santos, J. W.; Galvão, C.; Pasetti, M.; Schiavon, D.; Naqvi, F.; Hossain, R.; and Oliveira, N. D. 2024. Crossing the principle–practice gap in AI ethics with ethical problem-solving. *AI and Ethics*.

Crawford, K. 2021. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence (2021).pdf*. Yale University Press.

DSIT. 2023. Ethics, Transparency and Accountability Framework for Automated Decision-Making.

Edquist, A.; Liz, G.; Griffiths, S.; and Rowshankish, K. 2022. Ethics, Transparency and Accountability Framework for Automated Decision-Making.

Figueras, C.; Verhagen, H.; and Pargman, T. C. 2022. Exploring tensions in Responsible AI in practice. An interview study on AI practices in and for Swedish public organizations. *Scandinavian Journal of Information Systems*, 34(2).

[4]https://safeandtrustedai.org/

Franzke, A. S.; Muis, I.; and Schäfer, M. T. 2021. Data Ethics Decision Aid (DEDA): a dialogical framework for ethical inquiry of AI and data projects in the Netherlands. *Ethics and Information Technology*, 23(3): 551–567.

Gambelin, O. 2021. Brave: what it means to be an AI Ethicist. *AI and Ethics*, 1(1): 87–91.

GDS. 2020. Data Ethics Framework.

GDS. 2025. Developing frameworks and tools to support responsible data and AI use across the public sector.

Google. 2023. Secure AI Framework.

Greene, D.; Hoffmann, A. L.; and Stark, L. 2019. Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. *HICSS*.

Häußler, H. 2021. The underlying values of data ethics frameworks: a critical analysis of discourses and power structures. *Libri*, 71(4): 307–319.

Hirsch, D.; Bartley, T.; Chandrasekaran, A.; Norris, D.; Parthasarathy, S.; and Turner, P. N. 2024. *Business Data Ethics: Emerging Models for Governing AI and Advanced Analytics*. SpringerBriefs in Law. Cham: Springer International Publishing. ISBN 978-3-031-21490-5 978-3-031-21491-2.

Hollanek, T. 2024. The ethico-politics of design toolkits: responsible AI tools, from big tech guidelines to feminist ideation cards. *AI and Ethics*, 1–10.

Holstein, K.; Vaughan, J. W.; Daumé III, H.; Dudík, M.; and Wallach, H. 2019. Improving fairness in machine learning systems: What do industry practitioners need? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–16. ArXiv:1812.05239 [cs].

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

Kallina, E.; Bohné, T.; and Singh, J. 2025. Stakeholder Participation for Responsible AI Development: Disconnects Between Guidance and Current Practice. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, 1060–1079.

Kawakami, A.; Wilkinson, D.; and Chouldechova, A. 2024. Do responsible AI artifacts advance stakeholder goals? four key barriers perceived by legal and civil stakeholders. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, volume 7, 670–682.

Kazim, E.; and Koshiyama, A. 2021. The interrelation between data and AI ethics in the context of impact assessments. *AI and Ethics*, 1(3): 219–225.

Khan, A. 2025. *Data Ethics in the Age of AI*. Walter de Gruyter GmbH & Co KG.

Khan, A. A.; Akbar, M. A.; Fahmideh, M.; Liang, P.; Waseem, M.; Ahmad, A.; Niazi, M.; and Abrahamsson, P. 2023. AI ethics: an empirical study on the views of practitioners and lawmakers. *IEEE Transactions on Computational Social Systems*, 10(6): 2971–2984.

Leonelli, S. 2016. Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083): 20160122.

Longpre, S.; Biderman, S.; Albalak, A.; Schoelkopf, H.; McDuff, D.; Kapoor, S.; Klyman, K.; Lo, K.; Ilharco, G.; Rauh, M.; Skowron, A.; Vidgen, B.; Commons, M.; Ai, C.; Weidinger, L.; Narayanan, A.; Sanh, V.; Adelani, D.; Liang, P.; Bommasani, R.; Henderson, P.; Luccioni, S.; Jernite, Y.; and Soldaini, L. 2024. The Responsible Foundation Model Development Cheatsheet: A Review of Tools & Resources. *Transactions in Machine Learning Research*.

LOTI. 2022. The Brent Council Data Ethics Approach.

Lu, Q.; Zhu, L.; Xu, X.; Whittle, J.; and Xing, Z. 2022. Towards a Roadmap on Software Engineering for Responsible AI. ArXiv:2203.08594 [cs].

Met. 2021. Metropolitan Police Service Quarter 4 2021-24 Business Plan progress report.

Metcalf, J. 2019. Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics. *Social Research: An International Quarterly*, 86(2): 449–476.

Micheli, M.; Hupont, I.; Delipetrev, B.; and Soler-Garrido, J. 2023. The landscape of data and AI documentation approaches in the European policy context. *Ethics and Information Technology*, 25(4): 56.

Morley, J.; Floridi, L.; Kinsey, L.; and Elhalal, A. 2020. From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, 26(4): 2141–2168.

Munn, L. 2023. The uselessness of AI ethics. *AI and Ethics*, 3(3): 869–877.

Nethope. 2020. Artificial Intelligence (AI) Ethics for Nonprofits - Toolkit.

NSW Government. 2024. NSW Artificial Intelligence Assessment Framework.

ODI. 2021. Helping Barclays build trust through data ethics.

ODI. 2022. Data Ethics Maturity Model: benchmarking your approach to data ethics.

ODI. 2024. Embedding data ethics at Global Fishing Watch.

Olsen, D. R. 2007. Evaluating user interface systems research. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, 251–258. Newport Rhode Island USA: ACM. ISBN 978-1-59593-679-0.

Parker, C.; Scott, S.; and Geddes, A. 2019. Snowball sampling. *SAGE research methods foundations*.

Prem, E. 2023. From ethical AI frameworks to tools: a review of approaches. *AI and Ethics*, 3: 699–716.

Qiang, V.; Rhim, J.; and Moon, A. 2024. No such thing as one-size-fits-all in AI ethics frameworks: a comparative case study. *AI & SOCIETY*, 39(4): 1975–1994.

Sadek, M.; Kallina, E.; Bohné, T.; Mougenot, C.; Calvo, R. A.; and Cave, S. 2024. Challenges of responsible AI in practice: scoping review and recommended actions. *AI & SOCIETY*.

Samarawickrama, D. M. 2022. AI Governance and Ethics Framework for Sustainable AI and Sustainability.

Sambasivan, N.; Kapania, S.; Highfill, H.; Akrong, D.; Paritosh, P.; and Aroyo, L. M. 2021. "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. Yokohama Japan: ACM. ISBN 978-1-4503-8096-6.

Scheuerman, M. K. 2024. In the walled garden: Challenges and opportunities for research on the practices of the AI tech industry. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 456–466.

Schiff, D.; Rakova, B.; Ayesh, A.; Fanti, A.; and Lennon, M. 2021. Explaining the principles to practices gap in AI. *IEEE Technology and Society Magazine*, 40(2): 81–94.

Schiff, D. S.; Kelley, S.; and Camacho Ibáñez, J. 2024. The emergence of artificial intelligence ethics auditing. *Big Data & Society*, 11(4): 20539517241299732.

Solanki, P.; Grundy, J.; and Hussain, W. 2023. Operationalising ethics in artificial intelligence for healthcare: a framework for AI developers. *AI and Ethics*, 3(1): 223–240.

Sorell, T. 2024. AI-related data ethics oversight in UK policing. *Policing: A Journal of Policy and Practice*, 18: paae016.

Tirado, A. M.; Mulholland, P.; and Fernandez, M. 2024. Towards an Operational Responsible AI Framework for Learning Analytics in Higher Education. *arXiv preprint arXiv:2410.05827*.

Vakkuri, V.; Kemell, K.-K.; and Abrahamsson, P. 2019. Implementing Ethics in AI: Initial Results of an Industrial Multiple Case Study. In Franch, X.; Männistö, T.; and Martínez-Fernández, S., eds., *Product-Focused Software Process Improvement*, volume 11915, 331–338. Cham: Springer International Publishing. ISBN 978-3-030-35332-2 978-3-030-35333-9. Series Title: Lecture Notes in Computer Science.

Varon, J.; and Peña, P. 2021. Building a Feminist toolkit to question A.I. systems.

Wang, H.; and Blok, V. 2025. Why putting artificial intelligence ethics into practice is not enough: Towards a multi-level framework. *Big Data & Society*, 12(2): 20539517251340620.

Widder, D. G. 2024. Epistemic power in AI ethics labor: Legitimizing located complaints. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 1295–1304.

Yildirim, N.; Pushkarna, M.; Goyal, N.; Wattenberg, M.; and Viégas, F. 2023. Investigating how practitioners use human-ai guidelines: A case study on the people+ ai guidebook. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–13.

Zilka, M.; Ashurst, C.; Chambers, L.; Goodmann, E. P.; Ugwudike, P.; and Oswald, M. 2023. Exploring Police Perspectives on Algorithmic Transparency: A Qualitative Analysis of Police Interviews in the UK. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–19.