How do tech workers' perceptions of corporate structures affect their ethical decision-making?

Introduction

Untested and unregulated autonomous and semi-autonomous data-driven or AI systems are increasingly being used in areas where people have high dependencies and low agency – communications and social networking, social service provision, welfare and healthcare. Efforts to change practices in technology development, by influencing tech workers' design practices and ethical awareness in order to ensure that technologies are developed to maximise equitable benefits and minimise harms, have so far met with limited success.

To inform a significant gap in current understanding, this exploratory study explores the effect of corporate structures on tech workers' ethical decision-making. It locates itself in a sociotechnical model of tech worker practice: technical development and deployment contained within a wider context of legislation, regulation and governance, and the lived experiences of people exposed inequitably to benefits and harms. And proposes that all these factors operate on tech workers and their agency, shaping norms and influencing their attitudes, beliefs and behaviours.

Its focus is on the effect of situational, located corporate structures (those that are constructed and operate within the context of technology companies – such as ways of working or corporate ethics statements), as opposed to extrinsic structures (those that are imposed from outside, like guidelines, frameworks, principles and toolkits created by governments, policymakers, civil society and academia).

By inductively refining specific concepts of corporate structures through a qualitative thematic analysis of current literature, and testing them through an empirical, quantitative survey that explores contemporary technology workers' attitudes, beliefs and reported experiences in the context of corporate environments, this study intends to identify factors that do and don't have an effect on tech workers' ethical decision-making – and evidence that these situational factors are important and overlooked constituent parts of the effective governance of technologies.

As an exploratory study, the aim is to produce social data and mixed-methods research that demonstrates the validity of question and points to further research. The contribution of this study is to demonstrate that there is a benefit to studying tech workers, and that by focusing on technologies – influencing their design and its effect on outcomes – current approaches to supporting ethical decision-making (and studies that analyse them) are missing the additional significance of corporate structures in affecting the development of technologies and tech workers' ethical decision-making.

Background

Emerging technologies present governments and societies with intractable problems that requires multidisciplinary expertise and effective communication throughout complex networks and supply chains. There is a plethora of models – frameworks, principles, toolkits – to support tech workers to consider ethical design practices, and a well-documented need for translation of theoretical principles and high-level guidance into information that is usable in everyday development. In addition, there is a gap in knowledge about how business practices intersect with technology design, as barriers and incentives. Understanding more about how ethics is practised in technology companies is therefore critical.

It's an industry characterised by significant imbalances of power, between big technology companies and small, commerce and regulation, leadership and workers, etc., with considerable incentives for individual actors to prioritise their own interests (companies by

writing their own rules;[1] tech founders by resisting regulation).[2] That picture is changing: AI legislation is proposed in the UK[3] and internationally[4] – and international AI standards are in development.[5]

Tech workers' ethical agency as the developers of potentially beneficial or harmful systems is recognised, but detailed explorations of competing incentives and barriers in tech companies is underexplored in current research. Much of the research into AI ethics practices relates to algorithmic decision-making, focusing on the inputs, outputs and outcomes of data-driven and machine-learning systems. Analysis of interventions like technology ethics, AI ethics, AI for good, tech for good, pro-ethical design and value-sensitive design are well-researched (Powell et al, 2022; Vallor, 2010; Floridi and Cowls, 2019; Morley et al, 2023; Jacobs et al, 2021). Much of the learning that's improving technical processes is through the work of researchers with lived experience of harms, embedded in the technical processes of technology development, and the development of algorithmic assessment, datasheets, red-teaming, model cards, etc. (Buolamwini and Gebru, 2018; Bender et al, 2021; Gebru et al, 2021, Ashurst et al, 2022), and those working on tech worker organising, unionisation and positive, collective actions to secure change in technology practice (Boag et al, 2022; Wong, 2023).

Technology (and particularly AI) development is also significantly represented in the current mainstream press. While these reports form the context for this study, and frame some of the concerns that study seeks to evidence, it also recognises that the people who make technologies are rarely represented in discourse and narratives about technology. And so public perceptions of tech workers are skewed: limited to large corporations and frequently overlooking the majority of people who create technologies. We see them at crisis points, as whistleblowers, exploited marginalised workers and unionisers using collective bargaining to sidestep NDAs, but not as individuals whose everyday practice is to make – and therefore make decisions about – technologies that deliver benefits and harms to people, communities and societies. Also in this narrative, tech workers' negotiating power is diminishing: already damaged reputationally by the techlash,[6] they are experiencing unprecedent job uncertainty.[7]

A second part of the narrative is the deliberate obfuscation by technology company representatives of what technologies do and how they are built (for example, prioritising narratives of future risk over current effects). This is 'an everyday form of the modernist fallacy of the separability of society from technology (Latour, 1993), separating code from harms it enables.' (Widder and Nafus, p. 2). Companies and corporations have a vested interest in keeping the production of technologies mystified – the language of machine learning, existential risk and singularity contribute to this. Being able to conceptualise technologies as built by humans in corporate settings makes them much more open to strategies of regulation and governance.

---

[1] Benkler, Y. (2019). Don't let industry write the rules for AI. *Nature*, 569 (161).
[2] MTP on Twitter / X. ' Former Google CEO Eric Schmidt: companies developing AI should be the ones to establish industry guardrails.' Meet the Press. Accessed at: https://twitter.com/MeetThePress/status/1657778656867909633
[3] The UK Online Safety Bill, Data Protection and Digital Information Bill and forthcoming AI regulation.
[4] The EU AI Act is the first attempt to regulate artificial intelligence, and takes a risk-based approach.
[5] ISO / IEC 42001, 'Information technology – AI – management system'. *BSI (The British Standards Institute).* Accessed at: https://www.bsigroup.com/en-GB/standards/isoiec-42001/
[6] Su, N. M., Lazar, A. and Irani, L. (2021). 'Critical Affects: Tech Work Emotions Amidst the Techlash'. *Proceedings of the ACM on HCI* 5, no. CSCW1 (22 April 2021): 179:1-179:27.
[7] Vanian, J. 'Meta laying off more than 11,000 employees'. *CNBC*, 9 November 2022. Accessed at: cnbc.com/2022/11/09/meta-to-lay-off-more-than-11000-thousand-employees.html

While the current trajectory of technological development appears – and is frequently framed – as a novel problem, there are historical precedents that help to explain where we are. Corporate compliance, privacy practices (Rakova, 2021, p. 7) and professional codes of practice (Frankel, 1984) are all prior examples where the production of high-level principles is followed by a realisation that alignment with corporate practices is a significant challenge. The time elapsed between the development of principles, toolkits, etc. enables analysis of their effect within organisations. To explore this requires a sightline across technology ethics (science and technology studies, and philosophy of technology) and business ethics, which are traditionally siloed disciplines.

Theoretical background

This literature review has three purposes: 1) to locate this research within categories of existing knowledge in academic literature in tech ethics (science and technology studies (STS), and philosophy of technology) and business ethics about extrinsic and situational measures and interventions that affect tech workers' ethical decision-making; 2) to ground the empirical research that forms the basis of this research in tech workers' experiences of ethical decision-making, in order to build a legitimate basis for enquiry through a quantitative survey without using qualitative interview methods to develop concepts and theories; and 3) to inform the methods, concepts and measurements of this study through an analysis of research using related methods in relevant subject areas (see p. 19).

To do this effectively, the review of literature was designed using inductive research analysis techniques, identifying and coding results from studies identified as relevant. Studies were located using keyword searches in Google Scholar, Semantic Scholar and Connected Papers, and by 'snowballing', following references within papers identified as particularly relevant, both to papers cited, and to papers that reference papers of note within this subject area. What the review surfaced was an emerging field of study, without (as yet) clear findings, but with empirical evidence discoverable in related research. The theoretical bases of corporate structures are referenced in tech ethics literature, but still substantially located in business ethics.

The academic literature in this review addresses the following research questions, which guided the inductive process of the literature review, and inform the study hypotheses.

- **RQ1:** What factors do tech workers claim are important to them in relation to ethical decision-making?
- **RQ2:** Do tech workers perceive corporate structures as significant factors in supporting or constraining ethical decision-making?
- **RQ3:** Are perceptions of the effect of corporate structures expressed differently by tech workers in different situations?

The rapid growth of new technological capabilities reflect trajectories and learnings that are recognisable in more mature domains including legal compliance and privacy, which 'appear to have gone through a process that is mirrored in current algorithmic responsibility discussions: publication of high-level principles and values by a variety of actors, the creation of dedicated roles within organizations, and urgent questions about overcoming challenges and achieving "actual" results in practice and how to avoid investing in processes that are costly but do not deliver beyond cosmetic impact.' (Rakova et al, 2021, p. 4)

We can extend Rakova et al's logic to the literature of technology ethics, and trace a pattern from proposals for ethical principles and frameworks, with the intention that these are adopted within technology companies, alongside local legal requirements, to improve development practices and produce better outcomes for affected people and broader society (Veale et al, 2018; Leikas, 2019; Mittelstadt, 2019; Floridi et al, 2020; Raji et al, 2020). This

is followed by analysis of emerging principles and frameworks, and attempts to rationalise the proliferation by synthesising convergences around collections of principles (Jobin et al, 2019; Fjeld et al, 2020; Floridi and Cowls, 2021), recognising the increasing difficulty of mapping concepts and normative content (Morley et al, 2019, Ryan and Stahl, 2021), and some interventions like toolkits (Wong et al, 2023) or red teaming (Perez et al, 2022).

This in turn leads to a realisation of the gap between principles and practice (Mittelstadt, 2019; Morley et al, 2020; Hickock, 2021) with a sociotechnical framing that takes into account competing incentives, power dynamics, and ethics-washing goals fuelling growing cynicism about motivations, incentives and goals. And leads to critiquing the framing of tech ethics as 'vague and toothless' with 'a myopic focus on individual engineers and technology design, and… subsumed into corporate logics and incentives.' (Green, 2021, p. 209). Widder et al (2023, p. 475) observe that many studies have focused on helping tech workers identify issues, and that needs to develop into empowering them to act ethically – including broadening the terms of the discussion beyond AI and big tech, and using self-identified ethical concerns rather than pre-defined principles, like accountability or fairness.

Following Martin et al (2019, p. 310), this study recognises that 'the intersection of business ethics and technology ethics illuminates how our conceptions of work – and working – shape the ethics of new technology'. Evidence in this intersection is hard to pull out, as it sits between STS and tech philosophy knowledge about principles and frameworks for developers, and translated knowledge about corporate structures that is pulled in from organisational studies and business ethics.

The studies analysed for this review are listed in Appendix x. There are 8 relevant quantitative studies and significantly more 'small-N' qualitative studies, based predominantly on interviews with technology workers. While Johnson and Smith (2021, p. 1) propose a need for further research to assess whether preliminary findings from exploratory studies can be scaled, there are important empirical insights and findings that inform understanding of tech worker's perceptions, values, behaviours – and the analysis of these forms the basis of the development of concepts for measurement in this study.

Factors that affect understanding of tech workers' ethical decision-making

To understand factors affecting tech worker's ethical decision-making, we must first define concepts:

Tech workers are conceptualised in practical terms, through a number of intersecting properties: they are people of working age engaged in activities that contribute to the development and deployment of technologies and services (Holstein, 2019; Orr & Davis, 2020) in public- and private-sector organisations. Their roles include technical and project management knowledge and skills, and are distinct from – and defined by differential power and agency in relation to – management, and other organisational departments such as legal, policy or marketing. Tech workers work in companies of all sizes, not just big tech (and there is an explicit need to study those working outside big tech (Widder et al, 2023, p. 475). Lastly, tech workers described as 'developers' in this study are individual people (rather than companies).

Tech workers are not defined by one observable characteristic (unlike – for example – people who vote for a particular political party, or live in a geographic area). Mittelstadt (2019, p. 5) recognises their heterogeneity: coming 'from varied disciplines and professional backgrounds which have incongruous histories, cultures, incentive structures, and moral obligations.' Research into subjectivities of tech workers, identifies them variously through entrepreneurial 'post-neoliberal subjectivity' (Dorschel, 2022) or through their sociotechnical, embodied attunements in relation to techno-affects (Amrute, 2019) – and there is important

work exploring aspects of tech workers' critical agency in disability, critical race and gender studies.

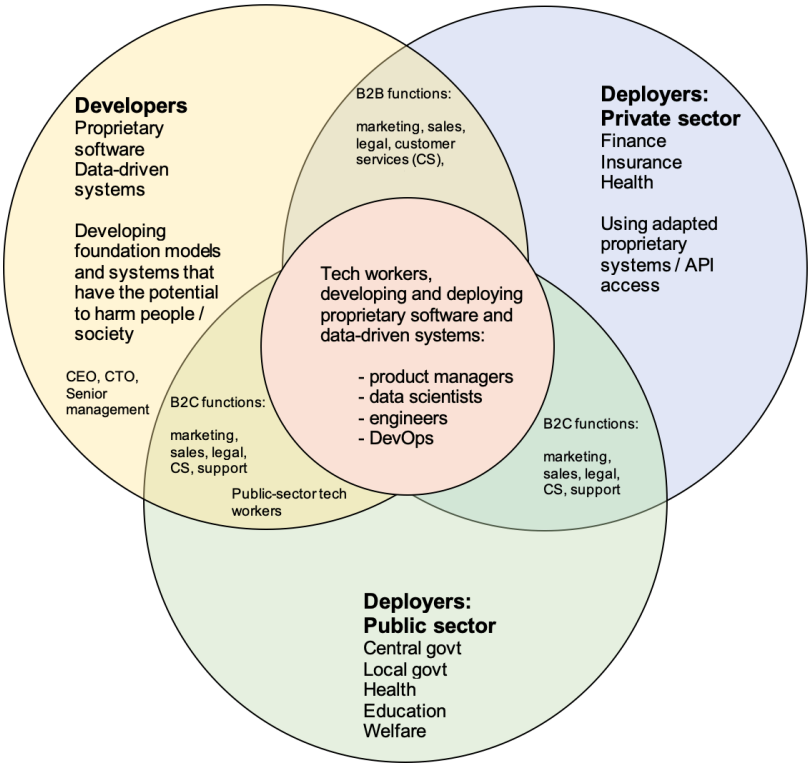

Figure 1: Tech workers: developers and deployers – private and public sectors

Despite perceptions from the press, tech work increasingly takes place not just in big tech but in all kinds of companies and involves many different roles and activities,[8] and tech workers are not just those who engage in collective bargaining, speak out under anonymity,[9] unionise[10] or whistleblow.[11] Rather than exploring these important, norm-violating attitudes and behaviours, this study defines everyday tech workers in order to uncover some of the nuance of their lived experiences, in relation to ethical decision-making.

[8] See CompTIA's state of the tech workforce UK report 2021: there are 2,451, 246 tech workers in the UK . 'State of the Tech Workforce', *CompTIA.* Accessed at: https://www.comptia.org/content/research/uk-tech-industry-and-workforce-trends-2021

[9] Alba, D and Love, J. 'Google's rush to win in AI led to ethical lapses, eemployees say'. *Bloomberg*, 19 April 2023. Accessed at: https://www.bloomberg.com/news/features/2023-04-19/google-bard-ai-chatbot-raises-ethical-concerns-from-employees

[10] Perrigo, B. '150 African Workers for ChatGPT, TikTok and Facebook Vote to Unionize at Landmark Nairobi Meeting'. *Time Magazine*, 1 May 2023. Accessed at: https://time.com/6275995/chatgpt-facebook-african-workers-union/

[11] Brown, S. 'AI workers need whistleblower protection'. MIT Management Sloan School, 6 October 2021. Accessed at: https://mitsloan.mit.edu/ideas-made-to-matter/ex-google-researcher-ai-workers-need-whistleblower-protection

Tech workers include but are not limited to those working in AI – which is an unspecific categorisation of mostly machine-learning technologies that pose particular risks for agency and autonomy. This study recognises that there are many different technologies that pose ethical risks (social media algorithms that are programmed to capture attention, or biometric identification technologies that systematically misrecognise people of colour). Empirical evidence (Vakkuri et al, 2020) finds no significant difference in attitudes between those developing AI products and other technologies.

Tech workers do not operate in isolation, but rather 'act as constrained autonomous agents, deferring judgment to extrinsic bodies while actively translating broad mandates into tangible products that will go out into the world' (Orr and Davis, p. 721). More than that, the deferral of judgement includes situational factors that are outside the purview of technology design (legal, marketing, company policies, management approaches, culture and norms) and have the potential to affect tech workers' ethical decision-making.

Ethical decision-making is distinct from individual moral judgement (the 'ability for the decision maker to decide which course of action is morally correct' Craft, 2011, p. 221). It locates moral judgement within a complex social context of contingent benefits and harms, and operates in a context of social norms and power structures – in this case in a corporate environment). This study distinguishes two categories: ethical beliefs, which are unique to the individual and support (but are not the same as) ethical decision-making – and are not the subject of this study; and ethical decision-making, which is situational (Ford & Richardson, 1994, p. 205).

1. Approaches to tech ethics

Tech ethics is commonly framed through three applied philosophical traditions, which relate abstract ethical conceptions of what guides actions towards others to normative, real-world problems. Deontological (acting in accordance with social rules or norms, based on codified or tacit universalised moral principles), consequentialist (choosing the best or majority (utilitarian) moral outcomes, attending to consequences for affected people) and virtue ethics (making decisions based on internalised moral values, like solidarity, fairness or equity). All attempt to apply moral knowledge to normative considerations arising through the design, development and implementation of technologies (including AI and ML).

The application of ethics to AI development and deployment is highly contested, and practical interventions involve all three approaches (sometimes described as 'principlism' (Stahl et al, 2020, p. 376)). The majority of studies related to tech workers' ethical decision-making propose normative changes to tech workers' design practices in the deontological tradition through the application of high-level principles. Developing toolkits, frameworks and corporate ethics statements assumes a deontological route to the development of technologies that support societal benefits. Recent moves towards impact assessments are based on a consequentialist view – what will be the outcomes for affected people.

Yu et al (2018, p. 5531) propose that consequentialist ethics (or ethics based on the utilitarian analysis of possible outcomes) is most closely related to the 'decision-theoretic frame of mind familiar to today's AI researchers', and that deontological (rule-based) and virtue ethics are less familiar. Yu et al also argue that different approaches have different purposes: that understanding deontological ethics 'can help AI researchers determine which rules are more fundamental and, therefore, should take priority in an ethical decision framework'. While virtue ethics, which centre individual understanding and moral growth, can 'help AI researchers frame ethical discussions in the context of changing social conditions and guide the incorporation of ethics into AI which shape the paths of learning'. Vallor (2010, p. 159) argues for the 'insufficiency of universal moral principles or of consequentialist calculations in the absence of the contextual sensitivity to, and perspicacious ability to read the moral significance of, a given set of particulars and respond

appropriately'. In other words, tech workers' ethical decision-making is impossible without both internalised moral values and contextual knowledge to guide the application of moral principles and rules, accompanied by the motivation to act morally.

These philosophical approaches are dominant, but are other approaches acknowledge the contested nature of ethical decision-making, and its relationship to power and privilege – including decolonial approaches (Mohamed et al, 2020) and the experiences of marginalised peoples (Birhane et al, 2022). When differential effects on distinct people and groups in society are recognised (rather than a narrow, technical capability-based approach to risk), it is evident that technology design, development and deployment are value-laden, and not morally or ethically neutral. Ethical decision-making is inherent to tech workers' practices: 'because values are implicated during technology design, developers make value judgments as part of their corporate roles.' (Martin et al, 2019). Rudschies et al (2021) argue for value pluralism, rather than synthesis into commonalities. As described above, those contested values have been explored in depth by scholars including Gebru, Mitchell, Raji, etc., reflecting Morley's contention, that there is an expectation that minority ethnic workers can act as the 'moral conscience' of the company that they are working for (p. 421).

Aula and Bowles (2023) examine AI ethics through network analysis, and the development of broad framings of ethical practices from rhetorical device to field (e.g data for good, AI for good), and their capacity to shape understanding and practices – because individuals wish to align themselves with accepted norms of practice. Alongside interventions like algorithmic impact assessments and red teaming, there are ethical models developed specifically within the practices of AI development, for example Dignum (2018) proposes ethics 'by design', which addresses ethical decision-making in autonomous systems; 'in design', which positions technical methods and regulation to support ethical outcomes of AI systems; and 'for design'), which involves adherence to certification, standards and codes of conduct. Powell et al (2022) proposed a systematised ethics of 'care and capability'. There are other, cultural measures of ethics, including religious beliefs as a predictor of beneficial behaviours (Spiekermann et al, 2018).

In this study's framing, all these approaches are trying to get to a 'situated' perspective through a variety of methods – turning principles into practice, applying context to moral virtue and locating outcomes-thinking in tech worker positionality. Following Moss, Metcalf & boyd (p. 455) this study proposes a framing of 'ordinary ethics', which 'attends to how everyday practices reveal the moral commitments embedded in actions, in contrast to the tendency to treat ethics as a form of argument or an abstraction'; and follows Orr and Davis' (p. 720) interest in 'ethics as part of practitioners' in-situ processes… tacit definitions [that] speak not to any objective ethical ideal, but reveal how ethics manifest into procedures, decisions, and material goods, and the related responsibilities these manifestations entail.'

2. Approaches to ethical decision-making

This literature review reveals an essential disclarity in the identification and categorisation of corporate and individualised responsibility, which is important for this study. As evidenced, there are many 'aspirational', 'educational' and 'regulatory' principles and codes (Frankel, 1989, p. 110), and many studies that critique them from a 'content-oriented, 'transformation-oriented' and 'outcomes-oriented' perspective (Helin and Sandström, 2007). Because they are not grounded in this granularity of practice, these studies tend to elide corporate and individual responsibility, so that it's unclear who is doing (or intended to be doing) the work of ethical decision-making. Among the proposals of normative principles and frameworks to support technology design to be done better by tech workers, this study has found none that explicitly support pragmatic, on-the-ground ethical decision-making in the context of – and across different components of – companies.

What all but the most recent of these approaches overlook – and what might be significant in their demonstrable lack of success in changing practices – is that ethical decision-making takes place within the context of technology companies. 'Ethical decisions regarding model development and deployment are ultimately made within contexts of organizations that have to align the ethical principles with vested interests such as the organizational culture, mission and goals.' (Krijger, 2021, p. 1429) This has two implications: ethical decision-making in a corporate environment is necessarily competitive rather than cooperative, underpinned by a tension between private and public interests (Mittelstadt, 2019, p. 3); and a risk that an individualised 'responsibility-as-a-virtue' approach means that 'without the support of company senior leadership, or appropriate whistleblowing policies, individual practitioners feeling responsible does nothing other than leaving these individuals vulnerable to retaliation.' (Morley et al, p. 415).

In summary, this study defines tech workers' ethical decision-making as situational, choice-making in the development of technologies, in relation to intrinsic and extrinsic drivers including principles and frameworks, corporate ethics statements, management approach to ethics and job role. Finally, an important distinction for this study is to recognise focus on ethical decision-making as part of the development of an AI system, as opposed to decision-making by – and supported by – AI systems post-deployment.

Corporate structures are situational, organisational norms, practices and role-based hierarchies that operate inside companies and corporations where technology development and deployment takes place (this study does not distinguishing on the basis of organisational constitution – whether organisations are limited liability companies or corporations). These are complex and tractable social structures that shape meaning, culture and operate on tech workers' attitudes and behaviours in a relationship of power dynamics: 'Developers will always be constrained by the institutions that employ them.' (Mittelstadt, 2019).

Corporate structures are distinct from extrinsic legislation, regulation or academic research-led principles and frameworks, because they originate and operate within organisational settings, forming definitional aspects of the environment where tech workers make decisions. These corporate structures are political, and tech workers are socialised into tacit and codified habitual attitudes and behaviours that act powerfully on their agency: 'habits of interaction go along with cultural norms, meanings, and moral valuations that tend to legitimate structures of power.' (Su et al, p. 11) They are activated through imperfect organisational logics: 'challenges include organizational pressures for growth, common software development approaches such as agile working that focus on rapid releases of minimal viable products, and incentives that motivate a focus on revenue within corporate environments'. And they freight systemic imbalances of power: 'practitioners have to grapple with lack of accountability, ill-informed performance trade-offs and misalignment of incentives within decision-making structures that are only reactive to external pressure.' (Rakova et al, 2021, p. 3).

Corporate structures are better defined in business ethics literature (which does not relate specifically to technology company environments), which is referenced and explored in tech ethics studies but not yet the sole subject of enquiry – and principally represented through empirical evidence in qualitative studies (Orr and Davis, 2021; Rakova et al, 2021; Kelley 2022; Widder and Nafus, 2023; Widder et al, 2023).

Factors affecting tech workers' ethical decision-making

Developed through the inductive literature review that supports this research, this study proposes a conceptual framework of extrinsic (outside companies and corporations) and situational (inside companies and corporations). Extrinsic factors include governance and

regulation, and the products and outcomes of tech ethics; and situational factors include the corporate structures that are the subject of this study.

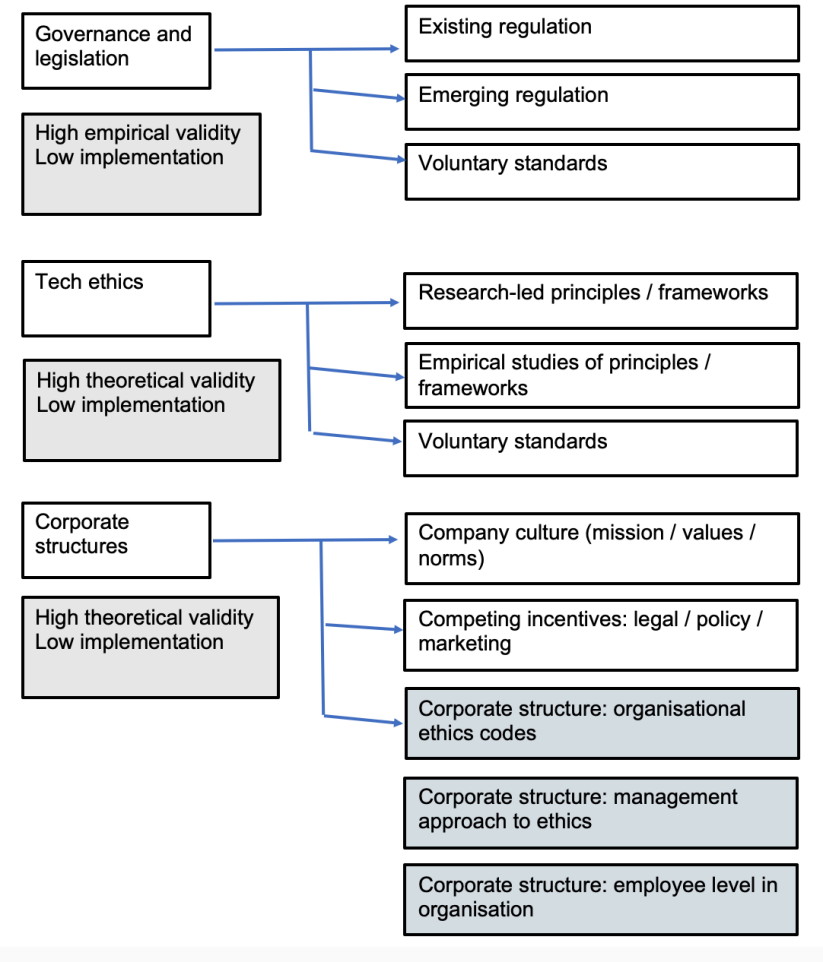

Figure 2: Theoretical basis for concept development

1. Extrinsic factors

This study distinguishes between extrinsic factors, which originate outside companies and corporations, and situational factors, which occur inside companies and corporations, in direct proximity to tech workers' everyday practices. There are many extrinsic factors that aim to affect positively the development of technologies – grouped by Stahl et al (2022, p. 28) into policy-level, guidance mechanisms and corporate governance. Policy-level mechanisms include regulation and governance – existing and emerging regulation – policy initiatives and new institutions. Guidance mechanisms include voluntary standards and many of the products of tech ethics, such as research-led principles and frameworks produced by both advisory and professional associations, and corporate governance covers many of the mechanisms discussed below.

1.1 Regulation and governance

Legislating and regulating for corporate compliance creates a complex dynamic of compliance, evasion and resistance. Corporate decision-making aims to stay within the constraints of what is technically legal under local jurisdictions (and the imperative to do this is one reason why the legal department can affect ethical decision-making). Morley et al's

study (2023**)** evidences that the primary motivation for ethical design currently is law (p. 417). While new, technology-specific legislation lags behind technical development and deployment, tech workers are often at the front line of ethical decision-making and pushing companies towards greater societal benefit (Widder et al, 469–71).

This study is not concerned specifically with the development of regulation, only inasmuch as the current, fragmented patchwork of data protection and human rights laws, and the absence of specific legal instruments for technologies including AI, informs the nature of the environment in which tech workers operate. Empirical evidence provides some relevant insights: Morley et al's (2023) study evidences that practitioners have high awareness of 'pro-ethical design', but associate it with data protection, privacy and security (where legal mechanisms already exist), rather than societal outcomes like solidarity. Company resource is relevant: data protection is resourced because it's mandatory, but other ethical outcomes are less incentivised. Ibanez and Olmeda (2022, p. 1672) find that privacy is a primary concern for practitioners – both from the perspective of compliance with regulations like GDPR and protection of company intellectual property.

Where Stahl et al (2020, p. 374) identify as a weakness the 'disconnect between rigorous academic research on the content and implications of these technologies and the development of governance proposals', this study proposes a disconnect between abstract principles and their implementation in situ. In the proliferation of different approaches to AI regulation (EU – risk based; UK – principles based, it is plausible to extrapolate that this knowledge gap (principles to practice) is reflected through the regulatory ecosystem, impeding the development of effective regulation and leading to different jurisdictional approaches. While recognising the context, this study is interested in (self-)governance of tech workers within organisations, not external regulation. And insofar as the presence and absence of regulation is reflected in practices in organisations – for example, Ali et al (p. 224) find 'a consensus that regulation would undoubtedly incentivize leadership at technology corporations to commit to taking these matters seriously, especially if there were financial penalties for non-compliance'.

1.2 Principles and frameworks

As described above, there has been a proliferation of principles and frameworks, and recent studies conclude that more work is needed in the gap between the proposal of these extrinsic solutions (mostly from academia, professional bodies and international third-sector organisation) and tech workers' own situational experiences and practices within corporate settings. Studies about the limits of these extrinsic principles identify them as overly broad and operating above the level of practice, analysing normative content of ethics guidelines to evidence gaps between principles and practices (Ryan and Stahl, 2021; Vakkuri et al, 2022) and point out the current risk in the proliferation of ethical guidelines – which require consideration of trade-offs as they come into conflict with each other, and are too high-level to translate into meaningful practice (Whittlestone et al, 2019, p. 195). The most recent studies on putting ethical principles and frameworks into practice recommend moving on from formulating high-level principles to translating them into practice, linking AI principles to specific processes aimed at companies and developers (McNamara et al, 2018; Vakkuri and Kemell, 2019; Chivukala et al, 2020; Morley et al, 2020; Orr & Davis, 2020).

The majority of empirical studies are in the philoposophical tradition described above ('principlism'), and accept the premise that emerging guideline development will lead to better technologies. Even studies that purport to be specifically aimed at changing practice, for example Ryan and Stahl (2021, p. 72) synthesise ethical considerations from normative content of ethical guidelines specifically for developers and users, distinguishing aspects of moral responsibility, legal liability, accountability and acting with integrity. Analysis is effectively at a corporate level (conflating AI developers and users into a category of 'AI organisation'), so not with clarity between the individual developers and the company – this

ambiguity is one of the gaps identified by this study. Similarly, Stahl et al (2022) recognise the role of individuals in increasing the complexity of solutions ('organisations work in a political and societal environment and are made up of individuals, so it is important to understand the breadth of interventions and mitigation strategies that are currently discussed', p. 28) while framing organisational responses to AI at the level of corporate governance.

Fjeld et al's study of 36 AI principles documents identify individual and collective 'professional responsibility', which recognises the critical role of those developing and deploying technologies, and identify its presence in 78% of documents in the study's dataset. (p. 5). This framing of responsibility is without a situational, embedded perspective, understanding that 'the behavior of such professionals, perhaps independent of the organizations, systems, and policies that they operate within, may have a direct influence on the ethics and human rights impacts of AI' (p. 56). Whittlestone et al (2019, p. 195) similarly stop at the corporate level in their synthesis of tensions in ethical principles, aimed at helping to translate principles into practice for technology companies, professional bodies, standard-setting bodies, government legislators and researchers.

Some empirical studies distinguish between organisational challenges and factors relating to individual responsibility of software engineers. Khan et al (2022) identify practitioners (those making design decisions about technologies) as having high ethical responsibilities – and therefore the imperative to understand their perceptions of principles and challenges. Spiekermann et al identify that there is a relationship between organisational norms, perceived control and responsibility (p. 2), but that it is 'engineers' individually perceived responsibility that determines most of their engagement with ethics' (p. 10). Chivukala et al (p. 8) find that practitioners identified both individual ethical responsibilities and duties associated with their organisational roles. Morley et al's (2023) study finds low numbers of respondents who have awareness of principles (less than 50% identified principles as important to the company they work for).

This is necessarily an emerging field of study: the proliferation of principles and frameworks has occurred post-2016 (the Partnership on AI's principles,[12] are generally acknowledged to be the first); others of significance include UNESCO and the High-Level Expert Group of the European Commission). Since then, 'Seemingly every organization with a connection to technology policy has authored or endorsed a set of principles for AI' (Fjeld et al, 2020, p. 4). Research mappings identify over 80 separate sets of guidelines in the public domain (Jobin et al., 2019; Morley et al, 2023), and we can begin to map them as situated, operating within the relational contexts of companies and corporations, or extrinsic.

There is an emerging empirical literature on experiences, attitudes and behaviours in relation to ethical principles, looking through the lens of high-level ethical principles and guidelines including transparency, accountability, predictability, responsibility, purpose, privacy, bias minimisation and robustness (Spiekermann et al, 2018; Rothenberger et al, 2019, Vakkuri et al, 2020, Khan et al, 2022). These studies have a variety of aims: Spiekermann et al look explicitly at issues of privacy and security, exploring experiential and instrumental attitudes and beliefs in relation to behaviours of software engineers; Vakkuri et aim to establish the current state of industrial practice, and particularly to determine the impact of high-level ethical guidelines by examining practitioners' perceptions.

Other empirical studies focus on bridging from ethics to technical implementation. Vakkuri et al (2020) demonstrate that, although various ethical guidelines are available, their

[12] Hern, A. 'Partnership on AI' formed by Google, Facebook, Amazon, IBM and Microsoft, *The Guardian*. Accessed at: https://www.theguardian.com/technology/2016/sep/28/go ogle-facebook-amazon-ibm-microsoft-partnership-on-aitech-firms, 2016.

deployment in companies has not had significant impacts on practice, and they need resource investment from companies to translate them into practical guidance for developers. This translation is imperative: 'Leaving ethical issues for the developers to tackle is unlikely to work. With no methods to help them, developers are left to rely on their own capabilities.' (p. 5) The gap between AI ethics research and practice remains an ongoing challenge. Morley et al (2021) ask about what practitioners understand about putting specific principles into practice, motivations, barriers and support required. Zhang et al (2021, p. 2) explore technical experts' attitudes about AI governance in an instrumental, normative way, rather than through an explicitly ethical framework. In Widder et al's (2023) study to explore respondents' ability to identify and act on ethical concerns – and this is perhaps reflective of the emerging recognition that high-level ethics have low salience with technology practitioners – participants identify areas of concern, rather than researchers proposing an ethical framework. They identify proximate concerns about military uses, privacy, surveillance and advertising, ranging in scope from bug fixing to existential risks.

There is some disclarity about who ethical frameworks are for – companies or individual tech workers. Rothenberger et al (2019, p. 8) discusses the issue of responsibility and whether the individual developer, the manufacturer or the end user would be held responsible for system actions. Morley et al (2023**,** p. 415) identify some confusion about responsibility for ethical design, and about the role of management. In their study, 24% of respondents said responsibility should be located with those designing, developing and deploying AI systems, and only 13% of respondents think management should be responsible for ethical design of technologies. Ali et al (p. 224) find that 'the way ethics and responsibility principles are incorporated at many technology corporations puts the onus on individuals to take on work as opposed to relying on organizational processes which may produce more consistent outcomes'. Ethics codes targeted at the individual without an awareness of context present a risk (related also to modularity of processes and seen in research ethics studies); that it is possible to 'do work with due regard for research ethics while also unwittingly collaborating in overall unethical actions'. (Taylor and Dencik, 2020, p. 3)

All these studies are to a greater or lesser extent divorced from the practical realities, incentives and barriers of tech workers operating in corporate environments – either through framing or method. In summary, there's an extensive literature about what ought to happen (normative), some about what is happening (empirical), some that points to localised good practices (such as AIAs, redteaming). There is no literature as yet about current holistic, organisational practice, which suggests those good practices don't yet exist. Moss and Metcalf (2020) propose case studies of good practice and talking about failures (p. 7).

2. Situational factors

Less acknowledged in academic literature is the role of located, situational factors – including what we will define as corporate structures – in ethical decision-making. In the current hiatus between recognition of potential societal harms and clear routes to ethical technology creation, companies and corporations have a distinct role to play in supporting tech workers' ethical decision-making, because: 'if firms ignore the value implications of design, engineers still make moral decisions; they simply do so without an ethical analysis.' Martin et al (p. 309).

Organisations' effects on ethical decision-making – which include corporate structures – are underexplored across the technology ethics literature, which has to date treated tech ethics as a ('techno-solutionist', i.e. looking for answers in technologies rather than sociotechnical context) design problem (Morley et al, p. 411; Greene p. 2129, Metcalf, Moss & boyd, p. 461). Kelley (2022, p. 872) states: 'only a handful of scholars have studied AI ethics in organizations.' Greene et al (2019, p. 2127) concur: 'Business decisions are never positioned as needing the same level of scrutiny as design decisions.' And in business ethics (where mechanisms like codes of practice are well-documented), Martin et al explain:

‘While the ethics of technology is analyzed across disciplines from science and technology studies (STS), engineering, computer science, critical management studies, and law, less attention is paid to the role that firms and managers play in the design, development, and dissemination of technology across communities and within their firm.’ (p. 307).

While the awareness of the gap is not new – Khalil (1993) first highlighted AI ethics as a management concern – empirical studies of technology workers’ attitudes, beliefs and reported behaviours (noting limitations of access that preclude immersive, observational research studies) are only beginning to emerge. Moss and Metcalf’s (2020) ‘ethics owner’ study based on interviews and ethnographic observations identifies tensions in corporate duties versus personal ethical position. Ali et al (2023, p. 218) combine qualitative interviews with observations in industry workshops, and report a culture of power centred around new product launches, workers incentivised by metrics that prioritise product innovation and financial gain, and modularisation associated with Agile development planning that makes it difficult to both access organisational knowledge and build and maintain relationships.

These cultural properties are endemic to technology development and a barrier to research: Spiekermann et al. (2018), for example, found that corporate cultures around privacy had impacts on the possibilities of researching privacy and security design decisions. There is an awareness of how those missing studies could inform future scholarship and decision-making: corporate structures are identified as a ‘dominant form of social relations’ (Widder & Nafus, p. 3) that have a powerful effect on attitudes and behaviours. Additionally, an organisation’s culture has a strong effect on tech workers’ ability to raise ethical concerns, including management attitudes (Widder et al, p. 473). Martin et al (p. 311) note existing studies on work on ethical decision-making by individual developers, but ‘very little work on how pressures from organizations and management matter to those decisions’.

Existing empirical studies produce a rich picture of situational barriers to ethical decision-making. Development practices are value-laden, and companies’ responsibilities for encoding responsibility into technologies is underexplored (Martin, 2019). Tensions between organisational culture and the responsible development and deployment of technologies are located in ‘values, processes, commitments, and governance structures’ (Martin et al, p, 311). Rakova et al (2021, p. 16) note a tendency to rely on individual decision-making, instead of developing and supporting mission- or values-aligned organisational processes. Ali et al (2023, p. 219), in a study of ‘ethics entrepreneurs’ – those with a specific ethical role – describe a ‘decoupling’ of public, corporate statements about ethical behaviours and internal processes and practices. There are tensions between rigorous technical practices and a culture of rapid product development and launch (Winecoff and Watkins, 2021), and this extends to ethical practices to understand risks for people and society (Rakova et al, 2021). The outcome of this can be a reduced form of ethics that aligns with corporate incentives (Ali et al, 2023). Lack of clear processes and responsibilities leads to what Ali et al (2023, p. 223) describe as an ‘individualisation of risk’, which has two effects: putting pressure on individuals and perpetuating a system that enables companies to continue to avoid responsibilities (they are incentivised to do this by the intractability of tensions between corporate drivers).

Taking these observations through the framing proposed by Ford and Richardson (1994, p. 212), we can identify specific, empirically tested, variables for ethical decision-making. These include management actions in the company (p. 212), with management’s impact, based through ‘how it acts and through its granting or withholding organizational rewards and sanctions’ (p. 216). In addition, the existence of a corporate statement on ethical behaviour (as an indicator of management support for ethical behaviour) is found to affect employees’ ethical decision-making (p. 216). And lastly, organisational effects – such as the size of an organisation, or an employee’s level in the organisation.

1. **Corporate structures (corporate ethics statements):**

Corporate structures can be codified, and this includes ethical guidance ('corporate ethics statements') produced internally by companies and corporations, such as IBM's 'Everyday Ethics for Artificial Intelligence';[13] and produced and distributed through technical professional organisations' memberships, including IEEE's Code of Ethics,[14] and the ACM code of ethics.[15] Rather than proposing models of organisational practice, they specify expectations for individuals' behaviour – or members of a professional community (Moss and Metcalf, 2020, p. 34; Mittelstadt, 2019). Kaptein and Schwartz (2008, p. 112) distinguish between micro (company codes), macro (professional, industrial and national codes) and meso (international codes).

As discussed, there is an extensive literature about frameworks and principles, but much less attention is paid to the effect of codes of ethics in technology development. There is extensive business ethics literature on codes of practice. Schwartz's (1999) survey of the relationship between codes of ethics and behaviour suggests existing research is inconclusive (p. 249), but he finds evidence that company codes can change behaviour (p. 260). Helin and Sandstrom's (2007) review of empirical studies on corporate codes of ethics concludes a lack of knowledge 'on how codes work, how they are communicated and how they are transformed inside organizations' (p. 260) and scepticism about the effective transformation of code into practice (p. 253).

Again, this issue is underexplored in relation to technology companies, which are not yet a formal profession, or subject to encoded fiduciary relationships or social contracts (Mittlestadt, 2019, p. 3). Johnson and Smith's 2021 survey of actionable interventions is the only current paper to summarise research in this area (p. 5) and concludes a need for empirical evaluations to determine what works. Empirical perceptions of employees demonstrate the difficulties of developing and implementing meaningful corporate ethics statements: 'Well, it is not formalised at all. It depends a lot on people, not on the process.' And 'I asked for [a code or protocol that includes ethical principles], but we don't have it.' (Ibanez and Olmeda, 2021). Nevertheless Winfield and Jirotka's (2018) Agile framework for ethical governance includes publishing an organisational ethical code of conduct. Codes can reinforce a culture where ethics is performed rather than embedded, so that: 'practitioners feel an aesthetic sensation of ethicality, while actually not attending to the full ethical design complexity present in the design situation.' (Chivukala et al, 2020, p. 2).

There is emerging evidence that codes of ethics produced by organisations extrinsic to technology companies make no difference to corporate practices – even in contained environments (membership communities) where an ethical framework might be appropriately socialised – not landing with practitioners. McNamara et al's (2018) experimental study to understand the influence of the ACM code of ethics in the software engineering decision-making process found no evidence that the ACM code of ethics regulates decision-making activities study. This is a tentative finding that needs further research, because the study was conducted as a lab experiment using student and practitioners, rather than ACM members with a situational investment in the code – a method that has low external validity.

There is, however, relevant literature to support this finding: Ford and Richardson (1994, p. 218) find that 'Industry ethical standards are not related to an individual's ethical beliefs and decision-making behavior.' Helin and Sandstrom (2007) similarly argue that 'placing too

[13] IBM. 'Everyday Ethics for Artificial Intelligence', 2018. Accessed at: https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf
[14] IEEE. 'Code of Ethics', Accessed at: https://www.ieee.org/about/corporate/governance/p7-8.html
[15] ACM, "Code of Ethics and Professional Conduct', Accessed at: https://www.acm.org/about-acm

much faith in the cause-and-effect relationship between the establishment of a code and ‘‘more ethical’’ behavior might be misleading. It might also enforce more of a reductionist view of an implementation process that is highly complex.’ (p. 254). Mittelstadt (2019) argues both that existing codes are too high level to guide actions, and that those produced by companies and developers are political, and motivated by reassuring policy-makers that new regulation is unnecessary (p. 1), as a kind of ‘ethics washing’, as well as distracting from unethical organisations and their ways of working by focusing on individuals’ behaviours

In addition to the codes themselves, the attitude of management towards ethics is a critical factor. Codes are only effective when they are ‘embedded in organisational culture and actively enforced’ (Mittelstadt, 2019. p. 8). Ibáñez and Fernández (2021) identify establishing and enforcing norms of behaviour (written or tacit) as a primary driver of ethical behaviour in SMEs (p. 357). Multiple studies suggest that developing meaningful codes in a complex and emerging area is difficult: empirically, Ibáñez and Olmeda (2022, p. 1687) evidence a ‘bottom-up’ approach, where companies develop systems first and then try to devise appropriate principles and behaviours.

2. Organisational effects: size of company, role in organisation

Corporate structures can be tacit. The ‘ethical culture’ of an organisation is made up of relationships between functions, policies and practices, as well as tacit norms and underlying assumptions (Ibáñez and Fernández, p. 342) Some tacit norms are encoded in structures and ways of working, for example, the dominance of ‘modularity’ as a tech company norm associated with Agile working practices; ways of working in smaller, less resourced companies; and management attitudes to – and support of – ethical practices.

Technology companies use modularity to refer to technical production of code, operational organisation of teams and ‘the broader, inseparable cultural beliefs, epistemologies, and organizational arrangements it mediates and reinforces’ (Hanna and Park, 2020). Modularity describes a fragmented chain of supply and decision-making, in which individual links in the chain are isolated from responsibility and oversight, and is also implicated in the conceptual gap between technology developers and users, including those affected by technologies (McPherson, 2018; Suchman, 2002). A more desirable but less prevalent business model is ‘located accountability’ (Widder and Nafus, 2023). Recognising the realities of corporate structures and designing interventions around them, rather than an idealised version, is critical (Madaio et al, 2020).

This kind of atomised organisational structure is a significant factor in empirical studies of adoption of principles, and affects ethical decision-making. A quarter of those interviewed for Kelley’s study ‘suggested decentralized structures could hinder the effective adoption of AI principles; several challenges with the structure were noted including enforcement, assigning responsibility, distribution, and gathering an AI inventory.’ (p. 884) Organisations wanting to implement ethical approaches to technology development find those goals are often in conflict with legal and economic goals, as well as technical or practical ability (Stahl, p. 32). When companies do institute ethical corporate structures, they may be located in different parts of a company – in legal, policy or research – and siloed from product and development teams (Moss and Metcalf, p. 74). A related risk is that ethics teams are empowered to challenge at the level of individual researchers or product teams, but not ‘institutional culture issues and leadership decisions that create the conditions for unethical behaviour’ (Green 2021, p. 209).

The effect of corporate structures can also be demonstrated through a comparison of tech workers’ attitudes in differently sized companies. Small companies, characterised as independent, owner-managed, comprised of under 50 employees, built on personal

relationships, often cash-limited and with informal mechanisms and processes (Spence, 1999, p. 164) are underexplored in business ethics literature despite accounting for the majority of practice. The effects on ethical decision-making of these different resourcing pressures are reflected in empirical studies: Start-ups are less likely to be equipped to do ethics work (Widder et al, p. 467), and pre-profitability of new companies limits resource for 'a capital-intensive project like "doing ethics"' (Metcalf. Moss & boyd, 2019, p. 465). Small companies are more concerned about business objectives and performance, complying with legislation and reputational risks than about ethical concerns (Ibanez and Olmeda, 2021, p. 1670). Ibáñez and Fernández (2021) identify giving employees a say in critical ethical issues as a primary driver of ethical behaviour in SMEs (p. 357).

The effect of management support for ethical structures, work and behaviours – and, as evidenced above, ethical codes – is also significant. Kelley et al ((2022, p. 71) identify management support as one of 11 components that can positively influence adoption of AI principles. Akbar et al (2023) find that lack of recognised ethics leadership is among the most important barriers identified by respondents. (Widder et al, p. 473) found access to management executives made employees more willing to raise ethical concerns. Additionally, 'suggestions from higher-ups that ethics discussions were a waste of time' were felt to be a significant barrier to ethical decision-making. Ali et al's study of ethics workers (p. 220) finds that 'the main problem they faced, which came with several downstream effects: a lack of support from leadership'. This leadership is multi-level: Agbese et al's (2023) study of middle and higher-level software engineering managers finds 'a fragmented understanding of ethical requirements' characterised by a tension between perceived technical and financial drivers (p. 66). Without adequate support, the ability of tech workers to address ethical issues 'depended on their network of high trust relationships in the organization' (Rakova et al, 2021, p. 16)

In summary, the corporate structures explored in this study are encoded norms (adopted internal policies and corporate ethics statements) and organisational units (legal, policy and marketing departments, and management support). Size of company is taken as a proxy for corporate structures, on the assumption that smaller companies have less resource to invest in ethical support. These inductive, qualitative findings align with the core text on ethical decision-making (Ford and Richardson, 1994, p. 216), which distinguishes situational contexts for organisation effects. That alignment produces the corporate structure concepts that underpin the quantitative element of this study (see Figure 1):

- Management actions, based both through 'how it acts and through its granting or withholding organizational rewards and sanctions'.
- The existence of an ethical code or corporate ethics statement (as an indicator of management support for ethical behaviour) is found to affect employees' ethical decision-making (p. 216).
- The size of an organisation.
- Employee's job role / level in the organisation.

Research design and methodology

This exploratory study developed concepts and measurements inductively, through an extensive literature review, then used a deductive research design to test hypotheses. This reflects the study's ontology and epistemology, which understands that the social world – and observations that can be collected and measured as data – are socially constructed. Specifically this study understands that people and technologies are mutually constituted, and that social processes and practices are materialised in technologies (Orr & Davis, p. 720).

There is a variety of social research methods that could support better understanding of complex societal systems and structures, and the effects they have in the world. Qualitative interviews with tech workers in corporate contexts would provide some evidence to underpin the development of theory, concepts and measurement for this study, informed by an understanding of corporate structures in technology companies, as well as knowledge, beliefs, attitudes and experiences of tech workers. However, given an extensive empirical literature, this is better developed through an analysis of existing quantitative and qualitative studies of tech workers, as well as general studies examining issues of AI ethics.

A quantitative method was selected to enable comparison of different effects on distinct groups within the sample, but with the expectation that the majority of insights will come from descriptive analysis of results, as sample groups will be too small to demonstrate statistical significance, but will provide indications of effects to inform future studies. This research design reflects a broader trend in the social sciences toward control via 'research design, rather than model-based statistical adjustment' (Dunning 2012, p. xvii).

Another method to answer the research question would be qualitative study, proposed as a better method than quantitative studies to understand the relationship between codes and behaviours (Schwartz, 2001, quoted in Kelley, p. 875). Tech companies' practices are notoriously hard to observe: companies tend to create cultures of loyalty and fiercely protect IP through enforcement of NDAs (Ali et al, 2023, p. 220), so methods that use direct observation (ethnography, focus groups and qualitative interviews) are rare and hard to negotiate.

Nevertheless, empirical studies are necessary to connect theory to practice. Empirical studies that represent experiences and perceptions can 'help dispel strongly held beliefs and misconceptions, as well as provide evidence-based suggestions for changes and improvements in research and practice' (Johnson & Smith, p. 1). Acknowledging that it's not possible to collect direct observations, this study uses a number of tactics to ensure the chosen method does provide appropriate evidence to address the research question, including sampling and data collection, and construction of questions.

As described above, this study describes 'tech worker' as those involved directly in the design and production of technical products and services (Orr & Davis, p. 720). Specifically (ref Figure 1), it differentiates tech workers from those in positions of corporate power (directorial and managerial roles), those involved in other parts of the technology development, deployment (legal, policy, marketing within tech companies) and the broader governance ecosystem (academia, professional organisation, government and regulators). Differentials, including power, are important in this area – there is not one 'tech worker', and a strength of survey methods is to identify differences in tech workers' roles.

Recognising that professional sectors operate as 'interrelated parts of a differentiated whole' (Orr & Davies, p. 723), the sample collection was not prescriptive. It used direct social media contact, public social media posts by gatekeepers, posts in professional fora, newsletters and private groups to contact a wide variety of role. In order to ensure respondents fell within

the terms of the required sample, a strongly worded message on the opening page asked people only to respond if they were tech workers, respondents were able to choose specific roles from a dropdown, and to add a role to self-identify themselves as tech workers if their role was not represented. This also enabled a post-collection cleaning of the data sample to isolate the views of corporate tech workers from public-sector workers and director/manager respondents – which allows for some comparison of responses.

Reflecting concerns that tech workers in large – and better resourced – companies may be more aware of ethical debates (Widder et al, p. 475), and that ethical decision-making operates differently in smaller and larger companies (Ford & Richardson, 1994, p. 217), the sampling method deliberately targeted workers in different sizes of companies, to ensure differential views were captured and to avoid collecting only a sample that repeats the dominant discourse of practices in big tech companies (Widder et al, 2023, p. 475).

This is designed as an exploratory study that aims to provide indications for future research questions. As described above (see Figure 2), there are a range of approaches to framing the factors that affect tech workers' ethical decision-making. These have notably different levels of confidence, stages of implementation and extent of exploration. To further understand the factors themselves, we have differentiated extrinsic and situational factors evidenced through the literature, and then identified those with high theoretical validity.

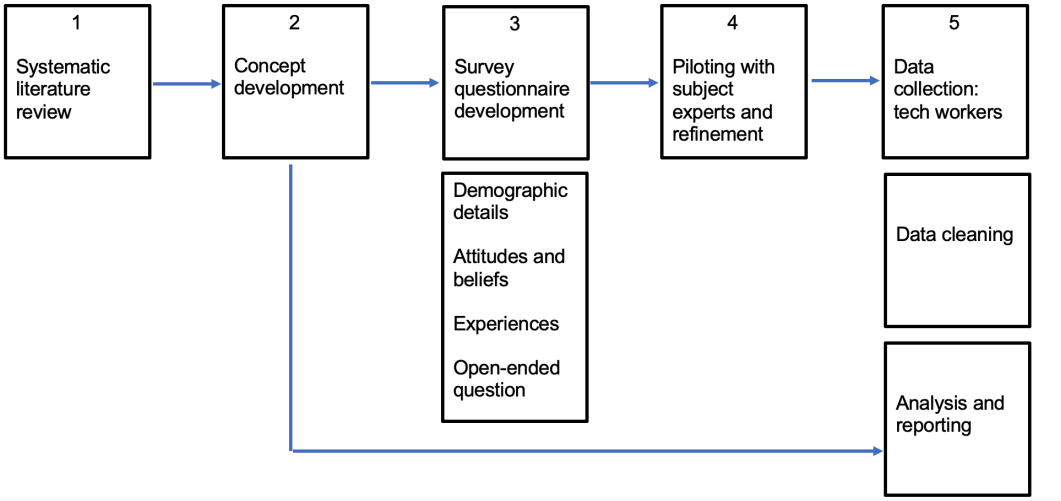

Figure 3: Research design

Learnings from existing studies:

The third purpose of the literature review was to inform the methods of this study. An analysis of the studies that collected data using quantitative methods follows:

Study	Data collection method	Sample	Recruitment	Analysis
Spiekermann et al, 2018	Survey (online)	124 software and systems engineers	Conference list, social network groups and (convenience) software companies	Deductive; descriptive statistics and Chi squared mean value analysis

Rothenberger et al, 2019	Grounded-theory literature review, qualitative interviews with experts, followed by a survey with publics	7 experts, 51 members of broader public.	Convenience sampling using social networks.	Inductive: qualitative and quantitative study data analysed individually; results presented as descriptive statistics and Chi squared mean value analysis
Vakkuri et al, 2020	Face-to-face structured interviews and online survey	249 practitioner respondents in 211 international software companies	Sampling method not discussed.	Inductive: Despite relatively large sample, the study does not propose hypotheses or use statistical analysis, but presents results as descriptive statistics
Zhang et al (2021)	Online survey	524 machine learning (ML) researchers, rather than technology developers	Selected through analysis of publication authorship in the top AI / ML conferences	Inductive: descriptive statistics, Chi squared mean value analysis and regression.
Morley et al (2021)	Mixed-method qualitative approach designed to produce a high-level general description of the experiences of AI practitioners attempting to implement AI ethics practices by combining a survey with semi-structured interviews	54 survey responses; 6 semi-structured interviews	Survey responses, collected through snowballing across social media and relevant mailing lists; semi-structured interviews through purposive sampling	Inductive: descriptive statistics (sample is too small for statistically significant results)
Khan et al (2022)	Survey compares the views of AI practitioners and lawmakers	99 AI practitioners and lawmakers across 20 countries,	Sampling through direct contacts, snowballing and social media	Inductive: descriptive statistics presented as frequency tables
Widder et al (2023)	Survey and qualitative interviews to report on software engineers.	115 international survey respondents and 21 follow-up interviewees across 5 countries) in non-profit, technology	Recruited through social media posts, message boards and channels, in-person at	Inductive: descriptive statistics presented as frequency tables; analysed survey responses and interview responses sequentially

		and broader sectoral organisations	developer meet-up	
Akbar et al (2023)	Exploratory study using a multivocal literature review (including academic and grey literature) and an online survey	156 international responses, anonymous	Recruitment through convenience sampling, snowballing and social media posts	Inductive: descriptive statistics presented as frequency tables
Ibáñez and Fernández, 2021)	Exploratory study, online survey	148 valid completed questionnaires from 65 companies	Recruitment through Global Compact Spain via email, website, social media posts	Deductive: least squares regression, factor analysis and Chi squared

Figure 4: Surveys measurement

Further analysis of research design and methods informs the measurement models for this survey – theoretical and practical approaches that connect high-level social constructs to observable data – as follows:

- **Role analysis:** Many surveys (Akbar et al, Khan et al, Vakkuri et al) include a range of roles in findings, including CEO, CTO, managers and policy. This does not allow isolation of tech workers in the data. To mitigate this, a range of tech worker job roles (data scientist, DevOps, product manager, etc.) are provided for selection by respondents, alongside a free text 'Other' option, which enables respondents to state their own job title. While this requires additional coding, it does enable comparative analysis of tech workers and management. Relatedly, Khan et al report significant differences in views by job sector, encouraging a design that enables stratification. Morley et al report no data about role, size of company, etc. so while the study has other useful comparative measures, there is no scope for comparison across the sample. Vakkuri et al's data collection includes size of company, but not job role. Only Spiekermann et al control for engineers' organisational contexts and hierarchical positions in organisations to inform findings (p. 5).
- **Reported effects:** Spiekermann et al's survey is framed to evidence reported behaviour and perceived responsibility, as well as attitudes and beliefs.
- **Partial completion:** Many studies (Akbar et al; Morley et al) report partial completion of surveys, and some difficulty in reaching response numbers required for analysis. Response rates are generally low: 17% of Zhang et al's sample completed at least part of the survey, which was distributed via email with an incentive for completion.
- **Theory to measurement:** (Akbar et al) report that thematic analysis of coded challenges surfaced from the literature review provide a robust basis for categories relevant to the experiences of practitioners.
- **Social desirability bias:** Many studies (Widder et al) list social desirability bias as a risk. Akbar et al report strongly positive findings, with no challenge marked at less than 65% importance.
- **Validity:** (Akbar et al) The survey design was subject to industrial experts' feedback, to ensure alignment of language.
- **Sample selection:** Widder et al claim to make use of self-selection of participants who self-identify as having ethical concerns, which means that results do not support generalisation. Similarly, Vakkuri et al's respondents self-reported as 'individuals

capable of influencing the development in their companies'. Morley et al's sampling is appropriate for a pilot study but does not reflect ethnic diversity and is too small for analysis. Rothenberger et al's sample numbers were too small to understand the impact of job role, gender, location. Spiekermann et al report a high drop-out rate of sample, and the possibility of self-selection bias meaning that the sample is not representative of broader engineering population. The size and specificity of sample prevents generalisation of results. Zhang et al is the only study to explicitly define the scope condition and critieria for sampling respondents, through a specific filter: researchers who have published at conferences.

- **Qualitative material:** Studies (Khan et al, Widder et al) use qualitative data to good effect, to support findings.

Research design: minimising measurement error

Following the learnings above, some alterations were made to the research design to minimise measurement errors:

Response effects: Some pretesting of question quality was carried out before data collection to support validity (Saris and Gallhofer, p. 165). The draft survey was tested on three subjects: an AI researcher, tech ethicist and tech founder, as well as through Qualtrics' question validity evaluation, which led to changes in language to align more closely with industry terminology, and measurement (reducing the numbers of questions and concepts measured; reducing the numbers of options in rating scale questions, and the points in Likert scale questions from 5 to 3 (Agree, Neither Agree nor Disagree and Disagree)).

Question order effects: These were mitigated through randomising presentation of order in ranking and selection tasks.

Respondent effects: Social desirability pressures are a consideration to studying ethical decision-making of tech workers using self-reported attitudes. Wrapping the questions in a longer survey about ethical decision-making made the study's objectives opaque to participants, reducing the risk of bias through normative pressures. This was minimal enough not to affect their beliefs about participation and did not increase any risk of harm to participants – as per prior ethics approval.

Researcher positionality effects: The power differentials of in-person interviews are not relevant here as data collection took place through an online survey, but the positionality of the researcher is an important factor, and the study would benefit from independent peer review to ensure an appropriate range of studies are included. Considering my positionality prompted a clear mechanism for access to the published study, so that any anticipated benefits from the research are shared back with the people participating in the research.

Analysis: From the experiences of empirical researchers, it seemed likely that the survey wouldn't reach numbers required for statistical analysis. It was therefore designed so that indicative results could be produced through descriptive analysis.

Representation error: This is a digital-only sample, which doesn't provide options for collection of data from digitally marginalised people, but this is appropriate for the sample – who are likely to be technically proficient and connected to devices and systems. Recognising that it may not be possible to collect a representative probability sample for this study within available resources, a convenience sample with transparency over job role will support descriptive findings. While it doesn't address the representation problem, to demonstrate that a survey study is based on non-probability methods, findings can be expressed as indications (Baker et al. 2013).

Establishing concepts

Building on Ford and Richardson's foundational study on corporate ethical decision-making, we operationalise the following concepts developed through the literature review, as dependent variables, to establish their effects:

- Management support for ethical behaviours
- Existence of an ethical code or corporate ethics statement
- Size of company
- Employee role level in organisation (tech worker vs manager)

From the literature review, we develop three hypotheses to test through the empirical data collection and analysis:

Hypotheses

H1 Approaches to ethical decision-making

- Description: Tech workers' perceptions of the importance of rule-based (deontological) approaches to ethical decision-making will not reflect the dominance of this approach in extrinsic development of ethical principles and frameworks.
- Measurement: 1) Positivity to deontological approaches vs virtue or consequentialist; 2) Positivity towards statutory measures (data protection and security) versus extrinsic (ethical principles and toolkits)
- Descriptive statistics: Comparison of positivity towards statutory measures (data protection and security) versus extrinsic
- Statistical test: Mean value analysis – Anova.

H2 Approaches to situational versus extrinsic support for decision-making

- Description: Tech workers will recognise the importance of situational over extrinsic approaches to decision-making (situational = those that are proximate, created within corporate environment, have cultural meaning in the everyday).
- Measurement: positivity to situational approaches vs extrinsic approaches – management support, corporate ethics statements versus extrinsic statutory and voluntary measures (data protection and security, and ethical principles and toolkits)
- Descriptive statistics: Comparison of positivity to situational approaches vs extrinsic approaches.
- Statistical test: Factor analysis and mean value analysis – Anova.

H3 Approaches to corporate structures

- Description: Tech workers will recognise corporate structures as a constraint to ethical decision-making.
- Measurement: Variations in perceptions of corporate structures 1) across job role (management vs tech worker); 2) across size of company; 3) whether they have left a tech role for ethical reasons?
- Descriptive statistics: Frequency tables.
- Statistical test: Factor analysis and mean value analysis – Anova.

Survey design:

The survey is informed by previous UK surveys of technology workers by DataKind[16] and Doteveryone in 2018,[17] not with the aim of collecting longitudinal data, but designing questions that relate well to industry practice. Those surveys find, for example, that more than a quarter (28%) of tech workers in the UK have seen decisions made about a technology that they felt could have negative consequences for people or society. Nearly one in five (18%) of those went on to leave their companies as a result.

To support the concepts, measurements and proposed analysis, the survey is designed as follows:

- A clear indication to respondents on the start page, as to who the subject of the research is, to discourage respondents who do not meet the sample criteria.
- Demographic questions including size of company; age; gender; education.
- Specific demographic questions for size of company and job role, to support analysis of corporate structures and their effects.
- Specific questions about corporate ethics statements (described as 'company codes of practice' to distinguish situational from extrinsic 'principles, frameworks and toolkits').
- Baseline questions that indicate positivity and negativity for different concepts, suitable for descriptive analysis and analysis using T-test / Anova for comparison of groups via means of variables.
- Composite questions with the same response formats (3-point Likert scale), that can be transformed into scales, to support factor analysis (Saris & Gallhofer).
- A range of question formats, to prevent respondent attrition (Saris & Gallhofer).

Data collection:

Data collection: survey data was collected using direct contact via LinkedIn, snowballing and circulation of the survey link to established professional organisations and groups. Test of sending direct InMessages via LinkedIn was not successful. Data collection through gatekeeper public and private groups was preferable, and the issue of respondents being a legitimate part of the sample population was mitigated through self-selection clearly spelled out on front of survey, and encouraging role selection / adding roles if not represented. This enabled transparency of differential roles in the sample, and sortition into groups for comparison.

Recruitment methods so not support external validity, but this study is intended as an exploratory survey, as the literature indicates it is unlikely to get response rate sufficient for statistically significant results, but nevertheless intends to use some statistical analysis. Therefore data collection is designed to prioritise for certainty about job roles of respondents, and groupings that will maximise sample numbers, e.g. company size; gender.

The survey was delivered through Qualtrics, and was open from 30 August to 22 September 2023.

[16] Datakind. (2019) 'Datakind Comminity Survey'. Accessed at: https://www.datakind.org/blog/datakind-uk-community-shares-views-on-ethics
[17] Doteveryone. (2018) 'People, Power, Technology: the Tech Workers' View'. Accessed at: https://doteveryone.org.uk/report/workersview/

Results

Data analysis:

The survey is designed in three parts: Q1–11 ask demographic questions; Q12–20 explore attitudes and beliefs, and Q21–25 explore experiences.

The dataset consists of 192 responses in total, of which 95 completed all questions and a further 25 completed more than 50%. Data was cleaned to remove 19 respondents who completed only 41% of the survey, 53% who completed only 4%, 6 responses with job roles outside the required sample (policy, researcher, etc.) and a further 12 responses identified as test data were discounted from the analysis, leaving a total sample of 102 responses in two categories of role – tech workers (86: public and private sectors) and managers (16) – reported below.

The survey was designed for statistical analysis using Anova / T-test to ascertain whether differences in means of groups (size of company; job role) are statistically significant and representative of systematic effects. However the sample sizes are too small (less than 50) to meet the assumptions (that data is normally distributed, and that the central limit theorem is met) for the proposed tests.

Following is descriptive analysis to propose indicators, to inform future studies:

Demographics:

Tech workers (n: 86)

- **Gender:** 38 identified as female, 43 as male, 4 as non-binary and 1 preferred not to say.
- **Ethnicity:** 5 respondents identified as Asian, 3 as Black African or Caribbean, 2 as Hispanic / Latinx, 4 as more than one ethic grouping, 69 as white, 2 self-described as Jewish and Hmong Southeast Asian, and 4 preferred not to say.
- **Age:** 2 respondents were 18–24 years old, 17 were 25–34 years old, 41 were 35–44 years old, 23 were 45–54 years old, 3 were 55–64 years old.
- **Size of company:** 41 worked in a company with 150+ employees, 11 in companies with 50–50 employees, 7 in SMEs of 10–50 employees, 9 in start-ups of up to 10 employees, 15 in public-sector organisations, 1 in a large non-profit, 1 in a charity and 1 freelancer.
- **Location:** 69 respondents' companies were located in Europe, 16 in North America and 1 in South America.
- **Experience:** 6 respondents had been in a tech worker role for less than 2 years, 11 for 2–5 years, 24 for 6–10 years, 28 for 11–20 years and 17 more than 21 years.
- **Education:** Nearly all respondents (85) had some form of university education, and 50 had a postgraduate qualification (MSc, PhD, etc.)
- **Job role:** 56 respondents (65% of the sample) identified with one of the job roles suggested: 22 (20%) identified as data scientists and 18 (17%) as product managers. 30 others in tech worker roles (35%) self-identified as UX designer, product designer, user researcher, content designer, full stack designer and privacy architect.

- **RQ1:** What factors do tech workers claim are important to them in relation to ethical decision-making?

Descriptive statistics:

There is a strong **tendency towards articulating socially beneficial personal motivations**: 82 respondents (95%) agreed it was important to them to consider outcomes for people and society; and 81 (94%) agreed that they believed they had a personal, moral responsibility to produce products that do good rather than harm. Similarly, when asked to rate activities in order of priority for their own ethical decision-making, participants rated **socially beneficial motivations as fairly high priorities**: ensuring all users can benefit equally (41 or 48% agreed it as a priority, while 23 or 26% disagreed), and anticipating and preventing negative outcomes for people and society (48 or 56% agreed, 13 or 15% disagreed).

There is a strong **articulated lack of alignment with corporate incentives**: 53 (61%) disagreed that they were more motivated by meeting team or company goals than by outcomes for people and society; and 73 (85%) disagreed that there's too much focus on ethics, and we need to get on with building technologies to make the world better. Similarly, when asked to rate activities in order of priority for their own ethical decision-making, participants rated **corporate incentives as low priorities**: developing an innovative product or service (9 or 10% agreed it as a priority, while 45 or 52% disagreed), and company revenue and growth (9 or 10% agreed, while 61 or 71% disagreed).

On **personal responsibility for ethical decision-making**, respondents' own sense of right and wrong was rated a high priority (41 or 48% agreed, while 15 or 18% disagreed). When asked whether they trust the people they work with to think about ethical concerns and consequences, so that they don't have to 59 (62%) disagreed, 23 (27%) neither agreed nor disagreed and only 3 (4%) agreed.

Discussion:

The tech workers in this sample perceive themselves to be highly motivated by factors aligned with consequentialist (outcomes-based) and virtue ethics (personal, moral responsibility). They are less motivated by deontological (rule-based) approaches, with only a small difference in perception of their importance, whether they were extrinsic or situated in corporate practices. One respondent said: 'Virtue is better than rules.' Overall, they reported low importance of both ethical principles and corporate ethics statements.

Their own motivations are less aligned with corporate incentives, including innovation and company revenue, than with socially beneficial outcomes. A comparison of perceptions of priorities of tech workers and companies shows some indications of difference (see Figure 5) shows an indication of a significant difference between tech workers' perceptions of their own and companies' prioritisation of corporate incentives, which may indicate a structural tension.

In relation to the hypothesis, that tech workers' perceptions of the importance of rule-based (deontological) approaches to ethical decision-making will not reflect the dominance of this approach in extrinsic development of ethical principles and frameworks, there is an indication that warrants further exploration.

- **RQ2:** Do tech workers perceive corporate structures as significant factors in supporting or constraining ethical decision-making?

Descriptive statistics:

When asked to rate activities in order for their own ethical decision-making **prioritisation of compulsory compliance mechanisms is high** (data privacy and security (52 or 60% agreed it was important). Though respondents thought **legal compliance alone was insufficient**: 76 (88%) disagreed that they didn't need to think too hard about things, as long as what their company does is legal. This is reflected in **positive attitudes towards regulatory oversight**: 59 respondents (67%) disagreed with the statement that government, regulators and industry standard-setters didn't need to get inside companies to see what's going on.

Prioritisation of voluntary mechanisms to support ethical decision-making is significantly lower: established, situational mechanisms (following internal company codes and policies) are given low priority (16 or 17% agreed; and 42 or 49% disagreed), as are established extrinsic mechanisms (using ethical principles, frameworks or toolkits), where (26 or 28% agreed; and 42 or 49% disagreed).

When asked about beliefs about corporate culture, 76 (88%) **felt strongly about corporate responsibility**, agreeing that companies have a responsibility to ensure that their technologies do not have negative consequences; there was **a balanced view on corporate vs societal incentives,** and whether companies can be prevented from prioritising profit, funders or shareholders over outcomes for people and society: 30 (35%) agreed, 35 (41%) disagreed and 17 (20%) neither agreed nor disagreed.

On perceptions of **companies' awareness of their own ethical capacity**, 52 (60%) said they didn't think companies are well equipped to recognise when technologies have potential to harm people and society, and on perceptions of **support for tech workers' ethical capacity**, 67 (78%) disagreed that tech workers in companies were already well supported to answer ethical questions.

Discussion:

Tech workers in this sample prioritise extrinsic compliance mechanisms that already have a legal basis highly, demonstrating that these are perceived as a significant factor in their ethical decision-making.

Priorisation of voluntary mechanisms (ethical principles and corporate ethics frameworks) are both low, with a slightly higher agreement that ethical principles are a priority. Following Vakkuri et al, who identify a low response rate to be indicative of low knowledge, we can propose that these have low salience with tech workers, so are not a factor in ethical decision-making. One respondent said: '"Ethical frameworks, principles or toolkits" kept coming up in the survey, and I'm not sure I've encountered those or know what particularly they refer to.'

Tech workers perceive companies to be inadequately equipped to recognise potential harms of technologies, and tech workers in companies to be insufficiently supported to make ethical decisions. A comparison of perceptions of priorities of tech workers and companies shows some indications of difference (see Figure 5) in tech workers' perceptions of corporate structures as supportive or constraining: tech workers are more motivated to use ethical principles, frameworks and toolkits, and less motivated to follow company codes and policies than companies. One important attribute of codes is how they encode attitudes of management, but there is no compelling evidence about the effect of management support.

In relation to the hypothesis: that tech workers will recognise the importance of situational over extrinsic approaches to decision-making, there is insufficient evidence in this sample. We do not have sufficient sample to assess perceptions of differential groups. However, the effects of legal mechanisms, tech workers' perceptions of the inadequacy of current corporate conditions are indicators for further research.

Factor	Tech workers		Companies		Descriptive analysis
	Agreed	Disagreed	Agreed	Disagreed	Tech workers perceive themselves as…
Developing an innovative product or service	9 (10%)	45 (52%)	32 (37%)	19 (22%)	Less motivated by developing innovative products or services than companies.
Ensuring all users can benefit equally	41 (48%)	23 (26%)	16 (19%)	32 (37%)	Significantly more motivated by ensuring all users can benefit equally than companies
Company revenue and growth	9 (10%)	61 (71%)	36 (42%)	27 (31%)	Significantly less motivated by company revenue and growth than companies.
Using ethical principles, frameworks or toolkits	26 (28%)	42 (49%)	11 (13%)	59 (69%)	More motivated to use ethical principles, frameworks and toolkits than companies.
Own sense of right and wrong	41 (48%)	15 (18%)	34 (40%)	21 (24%)	More motivated by their sense of right and wrong than companies.
Following company codes and policies	16 (17%)	42 (49%)	16 (27%)	14 (16%)	Less motivated to follow company codes and policies than companies.
Anticipating and preventing negative outcomes for people and society	48 (56%)	13 (15%)	48 (56%)	19 (22%)	Equally motivated by preventing negative outcomes for people and society.

Fig 5: Rate these activities in order of priority when you / your company are faced with ethical decision-making

- **RQ3:** Are perceptions of the effect of corporate structures expressed differently by tech workers in different situations.

The intention was that this analysis would compare perceptions of the effect of corporate structures across the sample, analysed by differences in job role, size of company and whether tech workers have recently left an ethical role for ethical reasons. There are no

direct descriptive statistics to evidence here – as described above, the sample size makes this statistical factor and mean value analysis through impossible – so instead we will take indications from frequency tables.

Job role: Sample numbers in differential job roles are too small to be analysed meaningfully.

Job mobility: There is evidence of job mobility in the sample, however only 3 had left job roles through ethical choice: 25 of 68 respondents had left a role in the last year; 16 as a career choice, 3 as an ethical choice (plus one who said '"Career choice" is the primary reason I left but there was also an ethical component which made the decision a great deal easier'), and 1 through redundancy.

Size of company: frequency tables

The literature review suggests that tech workers in small companies will feel less effect of ethical infrastructure than those in larger companies. Again, the sample numbers in each category are too small for meaningful analysis. There is a small indication that contradicts the literature analysis – showing a higher percentage of disagreement in small companies that ethics is not something that's practised.

Q17_6: Ethics is not something that's practised, it's just a set of principles that people generally overlook				
	Total	Agree	Neither agree nor disagree	Disagree
Total Count	78.0	20.0	22.0	36.0
Missing Count	0.0	0.0	0.0	0.0
Small company or start-up < 10 employees	11.0	2.0	0.0	9.0
Small-to-medium enterprise 10 – 50 employees	11.0	4.0	4.0	3.0
Medium-sized company 50 – 150 employees	8.0	0.0	4.0	4.0
Large company or corporation 150 + employees	32.0	9.0	11.0	12.0
Public-sector organisation	13.0	4.0	3.0	6.0
Other	3.0	1.0	0.0	2.0

Fig 6: Ethics is not something that's practised, it's just a set of principles that people generally overlook

There is also a more balanced picture of agreement and disagreement that ethical guidance is available in small companies than larger companies. Again, this contradicts the literature view analysis that small companies have low support for ethical practice, and is an indicator for future research.

Q17_5: TThere is clear guidance available about how to develop and deploy technologies ethically				
	Total	Agree	Neither agree nor disagree	Disagree
Total Count	73.0	16.0	17.0	40.0
Missing Count	0.0	0.0	0.0	0.0
Small company or start-up < 10 employees	10.0	4.0	3.0	3.0
Small-to-medium enterprise 10 – 50 employees	10.0	0.0	3.0	7.0
Medium-sized company 50 – 150 employees	8.0	0.0	2.0	6.0
Large company or corporation 150 + employees	32.0	9.0	6.0	17.0
Public-sector organisation	12.0	3.0	3.0	6.0
Other	1.0	0.0	0.0	1.0

Fig 7: There is clear guidance available about how to develop and deploy technologies ethically, analysed by size of company

In relation to the hypothesis, that tech workers will recognise corporate structures as a constraint to ethical decision-making, small sample sizes mean that there is limited analysis possible. There may be some differentials in small companies, which – in this study – are proposed as a proxy for the absence of corporate structures, but these are inconclusive.

Limitations

As described in the research design analysis, efforts were made to reduce limitations, but some remain inherent to the method, sample and nature of the study.

Sample bias: Because of the data collection method, there is a risk of self-selection bias in the sample – that respondents were aligned with the intentions of the research questions. Data collection for this type of exploratory survey is challenging, and different methods have a variety of trade-offs. Selection of tech workers: tried to design a data collection method that enabled randomisation of the sample as well as transparency of roles, so that tech workers could be identified.

The risk of social desirability bias is also present, as – despite attempts to mitigate in the question design – respondents may have felt social pressure to give 'ethical' answers. There is also a risk that attempts to construct questions to make it hard for respondents to decode the right / wrong answer to the measures that are significant to this study made the survey difficult to answer. The concepts to be measured are complex, it's not clear that respondents would understand the distinctions between principles and corporate ethics statements, for example – and the extrinsic / situational framing that is central to this exploratory research. Briefing survey respondents with key definitions would have ensured a common understanding of concepts, but arguably reduced research validity.

While anonymisation of results was necessary to assure respondents that there was no risk to them from completing the study, this had an effect on data collection – as it was not possible to identify and nudge partially completed survey responses. This accounts for much of the attrition (192 responses to 108). The complexity of concepts and question structure might also account for some of the high attrition rates – though this is a trade-off against greater validity of data. Similarly, the research design means that the significant dropout rate is indicative of self-selection that strengthens the data.

Survey methods: There is a methodological limitation to results of a quantitative study: that the results evidence what people think but not why. This is mitigated somewhat by the addition of a free-text field for participants to add anything they feel they weren't asked and want to say, but can only be addressed substantially by follow-up research.

Internal validity: A sample coding of the literature review was verified by a researcher colleague. To support validity, the study results should have been pre-registered.

External validity: The convenience collection method and numbers of responses mean that this study is not representative and can't be generalised to the larger population of tech workers. Extensive documentation of research methods for the qualitative study support reproducibility.

Construct validity: The method for constructing concepts is described above. One potential threat to this study's construct validity is the gap in the literature over individual versus corporate responsibility – where studies claim to analyse responsibility on an individual basis (for example, by examining high-level principles' effects in corporate environments), but their actual analysis is at the corporate level.

More testing of concepts and concept measures with representative subjects would have

supported the validity of the research – to demonstrate that respondents understood the terms effectively, but this was not possible given the study resource. In a related empirical study, Morley et al (2021)'s use of interviews alongside a questionnaire reveal disclarity about meaning of AI ethics and ethical principles (p. 421), which means survey questions may not have been understood. Rothenberger et al (2019) provided supporting definitions of guidelines, to ensure common understanding. Vakkuri et al (2020) report that 39% of respondents skipped or answered 'don't know' to the liability question, indicating low understanding, which they interpret as low ethical priority (p. 3).

Some refinements to constructs were overlooked and could be improved on:

- The addition of organisational statements of principles to the coding of extrinsic principles and frameworks and corporate ethics statements. Moss and Metcalf (2020, p. 30) disambiguate the symbolic purpose of organisational statements of principle, and they are worthy of exploration, as a statement of managements' ethical intention.
- In addition, the conclusions of this study are based on literature discovered through snowballing from existing studies and keyword searches. Despite these measures, the study does not claim to be a complete survey of the literature.

Conclusion validity: There is a self-evident limitation relating to the size and composition of the sample. Although sufficient for this study's analysis, it is not representative of the broad community of tech workers. The findings are considered indicative, and judgement should be exercised when extrapolating the results.

Real-world effects: Consideration of the effects of technology – and particularly AI – on society has had high news salience during the research period. Despite peaks in interest include the launch of ChatGPT, the resignation of Geoffrey Hinton from Google and an explosion of concern about existential risk, leading to a UK-led global summit on AI Safety in November 2023, there has been nothing significant enough to have a demonstrable effect during the data collection period.

Ethical approval

This study has ethical approval through the University of London Birkbeck College ethics committee under reference BBKPOL2022/23-15. Participants were not incentivised or paid to be part of the study.

There is an identified risk of harm to participants divulging information about corporate practices, and a potential disincentive to participate in the study. This was mitigated by anonymising results so that any identifying, categorical information including size of company, role, etc. would not enable re-identification of respondents (Veale et al, 2018; Holstein, 2019; Orr and Davis, 2020).

Conclusion

As discussed, this study is a detailed and technical analysis of one particular – and overlooked – aspect of tech workers' experiences: the effect of corporate structures on their ethical decisionmaking. It adds to the literature about interventions to support tech workers' ethical practices, intended as an 'and', to support embedding those practices by observing and removing barriers produced by corporate structures.

In conclusion, this paper proposed that current proposals for measures and interventions to change the ways technology companies operate are founded on an incomplete understanding of the effects of voluntary controls (corporate ethics statements, frameworks,

principles), which leads to a contested picture of measures in the regulatory environment (standards, regulatory institutions and laws). It attempts to evidence that the absence of analysis that addresses holistically – across tech and business ethics literature – corporate structures' effect on tech workers' decision-making is inhibiting comprehensive understanding and approaches across technology studies, policy and governance. And that – whether produced or endorsed by academia, corporations, professional associations or public institutions, these principles are not politically neutral – companies have an investment in soft governance and self-regulation – and the avoidance of 'hard' regulatory measures that constrain their practices.

Much of the literature around tech worker's ethical decision making is located in philosophical approaches to ethics – deontological, consequentialist and virtue ethics. It is the contention of this paper that none of these approaches is individually useful to interrogating what happens when an individual is operating in an environment where their virtuous attitudes, behaviours or decision-making are valued or devalued, or incentivised or de-incentivised, by the structures and norms encoded in the prevailing corporate culture – and that more attention should be paid to their everyday experiences in corporate environments.

With Martin et al (p. 309), this study contends that: 'It follows from these two observations—technology firms operate with nascent external oversight and designers are making value-laden decisions as part of their work in firms—that the most direct means of addressing ethical challenges in new technology is through management decisions within technology firms.' As an exploratory study, the aim of the research was to produce social data and mixed-methods research that demonstrates the validity of the question and points to further research.

Bibliography

[For empirical texts analysed in study, see Appendices]

Agbese, M., et al. ‘Implementing AI Ethics: Making Sense of the Ethical Requirements’. In *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering*, 62–71. EASE ’23. New York, NY, USA: ACM 2023.

Amrute, S.. ‘Of Techno-Ethics and Techno-Affects’. *Feminist Review* 123, no. 1 (1 November 2019): 56–73.

Ashurst, C., et al. ‘Disentangling the Components of Ethical Research in Machine Learning’. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 2057–68. FAccT ’22. New York, NY, USA: ACM, 2022.

Aula, V., and Bowles, J. ‘Stepping Back from Data and AI for Good – Current Trends and Ways Forward’. *Big Data & Society* 10, no. 1 (January 2023): 205395172311739.

Babri, M., Davidson, B., & Helin, S. (2019). An updated inquiry into the study of corporate codes of ethics: 2005–2016. Journal of Business Ethics. https://doi.org/10.1007/s10551-019-04192-x

Bender, E. M., et al. ‘On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?’. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. FAccT ’21. New York, NY, USA: ACM, 2021.

Birhane, A., et al. ‘The Forgotten Margins of AI Ethics’. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 948–58, 2022.

Boag, W. et al. ‘Tech Worker Organizing for Power and Accountability’. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 452–63. Seoul Republic of Korea: ACM, 2022.

Bryman et al. *Social Research Methods (6th edition).* Oxford University Press, 2021

Buolamwini, J., and Gebru, T. ‘Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification’. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. PMLR, 2018.

Craft, J. L. ‘A Review of the Empirical Ethical Decision-Making Literature: 2004–2011’. *Journal of Business Ethics* 117, no. 2 (October 2013): 221–59.

Dignum, V. et al.: “Ethics by design: Necessity or curse?,” in Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society 60–66 (2018)

Dorschel, R.. ‘A New Middle-Class Fraction with a Distinct Subjectivity: Tech Workers and the Transformation of the Entrepreneurial Self’. *The Sociological Review* 70, no. 6 (November 2022): 1302–20.

Dunning T. 2012. Natural Experiments in the Social Sciences: A Design-Based Approach. Cambridge, UK: Cambridge Univ. Press

Fjeld, J. *et al.* (2020) ‘Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI’. Rochester, NY.

Floridi, L. *et al.* (2020) ‘How to Design AI for Social Good: Seven Essential Factors’, *Science and Engineering Ethics*, 26(3), pp. 1771–1796.

Floridi, L. and Cowls, J. (2019) ‘A Unified Framework of Five Principles for AI in Society’, *Harvard Data Science Review*, 1(1).

Ford, R. C, and Richardson, W. D. ‘Ethical Decision Making: A Review of the Empirical Literature’, n.d.

Frankel, M. S. ‘Professional Codes: Why, How, and with What Impact?’ *Journal of Business Ethics* 8, no. 2–3 (1989): 109–15.

Gebru, T., et al. ‘Datasheets for Datasets’. arXiv, 1 December 2021.

Green, B. ‘The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice’. *Journal of Social Computing* 2, no. 3 (September 2021): 209–25.

Greene, D., et al. *Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning*, 2019.

Helin, S., & Sandström, J. (2007). An inquiry into the study of corporate codes of ethics. Journal of Business Ethics, 75(3), 253–271.

Hickok, M. ‘Lessons Learned from AI Ethics Principles for Future Actions’. *AI and Ethics* 1, no. 1 (1 February 2021): 41–47.

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. Journal of Service Research, 21(2), 155–172.

Jacobs, M.. et al. ‘Value Sensitive Design and Power in Socio-Technical Ecosystems’. *Internet Policy Review* 10, no. 3 (30 September 2021)..

Jobin, A., et al. ‘The Global Landscape of AI Ethics Guidelines’. *Nature Machine Intelligence* 1, no. 9 (September 2019): 389–99.

Johnson, B., and Smith, J. ‘Towards Ethical Data-Driven Software: Filling the Gaps in Ethics Research & Practice’. In *2021 IEEE/ACM 2nd International Workshop on SEthic)*, 18–25. Madrid, Spain: IEEE, 2021.

Khalil, O. E. M. (1993). Artificial decision-making and artificial ethics: A management concern. Journal of Business Ethics, 12(4), 313–321.

Kaplan & Haenlein, 2019. Rulers of the world, unite! The challenges and opportunities of artificial intelligence. Business Horizons.

Kaptein, M., & Schwartz, M. S. (2008). The effectiveness of business codes: A critical examination of existing studies and the development of an integrated research model. Journal of Business Ethics, 77, 111–127.

Krijger, J. ‘Enter the Metrics: Critical Theory and Organizational Operationalization of AI Ethics’. *AI & Society,* 37, no. 4 (1 December 2022): 1427–37.

Leikas, J., Koivisto, R. and Gotcheva, N. (2019) ‘Ethical Framework for Designing Autonomous Intelligent Systems’, *Journal of Open Innovation: Technology, Market, and Complexity*, 5(1), p. 18.

Martin, K. E. (2019). Ethical implications and accountability of algorithms. Journal of Business Ethics, 160, 835–950.

Martin et al., 2019 - Martin, K. E., Shilton, K., & Smith, J. (2019). Business and the ethical implications of technology: Introduction to the symposium. Journal of Business Ethics, 160, 307–317.

Mittelstadt, B. (2019) 'Principles alone cannot guarantee ethical AI', *Nature Machine Intelligence*, 1(11), pp. 501–507.

Mohamed, S., Png, M. T. and Isaac, W. 'Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence'. *Philosophy & Technology* 33, no. 4 (December 2020): 659–84.

Moss, Metcalf, J. & boyd, d.'Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics', n.d.

Orr, W., and Davis, J. L. 'Attributions of Ethical Responsibility by Artificial Intelligence Practitioners'. *Information, Communication & Society* 23, no. 5 (15 April 2020): 719–35.

Perez, E., et al.. 'Red Teaming Language Models with Language Models'. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 3419–48. Abu Dhabi, United Arab Emirates: ACL, 2022.

Powell, A. B, et al. 'Addressing Ethical Gaps in "Technology for Good": Foregrounding Care and Capabilities'. *Big Data & Society* 9, no. 2 (July 2022):

Raji, I., et al. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33–44.

Rudschies, C. et al. 'Value Pluralism in the AI Ethics Debate – Different Actors, Different Priorities'. *The International Review of Information Ethics* 29 (30 March 2021).

Ryan, M, and Stahl, B. C. 'Artificial Intelligence Ethics Guidelines for Developers and Users: Clarifying Their Content and Normative Implications'. *Journal of Information, Communication and Ethics in Society* 19, no. 1 (1 January 2020): 61–86.

Schwartz, M. (1999), 'The Nature of the Relationship Between Corporate Codes of Ethics and Behaviour'. International Association for Business and Society

Spiekermann, S. et al. 'Inside the Organization: Why Privacy and Security Engineering Is a Challenge for Engineers'. SSRN Scholarly Paper. Rochester, NY, 20 August 2018.

Su, N. M., et al. 'Critical Affects: Tech Work Emotions Amidst the Techlash'. *Proceedings of the ACM on Human-Computer Interaction* 5, no. CSCW1 (22 April 2021): 179:1-179:27.

Taylor, L. and Dencik, L. (2020) 'Constructing Commercial Data Ethics', *Technology and Regulation*, pp. 1-10.

Vakkuri, K. & Kemell, K. "The current state of industrial practice in artificial intelligence ethics," IEEE Software, vol. 37, no. 4, pp. 50–57, 2020.

Vallor, S. (2010) ‘Social networking technology and the virtues’, *Ethics and Information Technology*, 12(2), pp. 157–170.

Veale, M., et al. ‘Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making’. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. Montreal QC Canada: ACM, 2018.

Whittlestone, J., et al. ‘The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions’. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 195–200. AIES ’19. New York, NY, USA: ACM, 2019.

Winfield, A. F. and M. Jirotka, “Ethical governance is essential to building trust in robotics and artificial intelligence systems,” Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 376, no. 2133, p. 20180085, 2018.

Yu, H. et al. (2018) ‘Building Ethics into Artificial Intelligence’, in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. *Twenty-Seventh International Joint Conference on Artificial Intelligence {IJCAI-18}*, Stockholm, Sweden: IJCAIO, pp. 5527–5533.

Appendix 1: Empirical studies analysed in this study

Numbers of studies analysed:

- 18 studies address corporate structures in relation to technology development and deployment. None does this as a primary research question, all are subsidiary questions in relation to a related area of enquiry.
- 19 analytical studies address aspects of tech workers' ethical responsibilities.
- 22 studies take an analytical view of barriers or challenges to dominant views of measures and interventions.
- 8 quantitative surveys, and 1 'small n' qualitative questionnaire that is analysed descriptively.
- 18 qualitative studies using interviews, focus groups and case studies.

Quantitative studies:

This is an emerging field of study, and there are currently eight principal relevant studies in technology and business ethics, where researchers have set out to understand attitudes, opinions and behaviours of technology workers in a range of settings and locations, using survey methods (Spiekermann et al, 2018; Rothenberger et al, 2019; Vakkuri et al, 2020; Zhang et al, 2021; Morley et al, 2021; Khan et al, 2022; Widder et al 2023; Akbar et al, 2023); and one study in business ethics (Ibáñez and Fernández, 2021).

Akbar, Muhammad Azeem, Arif Ali Khan, Sajjad Mahmood, Saima Rafi, and Selina Demi. 'Trustworthy Artificial Intelligence: A Decision-Making Taxonomy of Potential Challenges'. *Software: Practice and Experience* n/a, no. n/a. Accessed 10 September 2023. https://doi.org/10.1002/spe.3216.

Khan, Arif Ali, Muhammad Azeem Akbar, Mahdi Fahmideh, Peng Liang, Muhammad Waseem, Aakash Ahmad, Mahmood Niazi, and Pekka Abrahamsson. 'AI Ethics: An Empirical Study on the Views of Practitioners and Lawmakers'. arXiv, 15 November 2022. http://arxiv.org/abs/2207.01493.

Morley, Jessica, Libby Kinsey, Anat Elhalal, Francesca Garcia, Marta Ziosi, and Luciano Floridi. 'Operationalising AI Ethics: Barriers, Enablers and next Steps'. *AI & SOCIETY* 38, no. 1 (February 2023): 411–23. https://doi.org/10.1007/s00146-021-01308-8.

Rothenberger, Lea, Benjamin Fabian, and Elmar Arunov. 'RELEVANCE OF ETHICAL GUIDELINES FOR ARTIFICIAL INTELLIGENCE – A SURVEY AND EVALUATION'. *Research-in-Progress Papers*, 15 May 2019. https://aisel.aisnet.org/ecis2019_rip/26.

Spiekermann, Sarah, Jana Korunovska, and Marc Langheinrich. 'Inside the Organization: Why Privacy and Security Engineering Is a Challenge for Engineers'. SSRN Scholarly Paper. Rochester, NY, 20 August 2018. https://papers.ssrn.com/abstract=3598930.

Vakkuri, Ville, Kai-Kristian Kemell, Joni Kultanen, and Pekka Abrahamsson. 'The Current State of Industrial Practice in Artificial Intelligence Ethics'. *IEEE Software* 37,

no. 4 (July 2020): 50–57. https://doi.org/10.1109/MS.2020.2985621.

Widder, David Gray, Derrick Zhen, Laura Dabbish, and James Herbsleb. 'It's about Power: What Ethical Concerns Do Software Engineers Have, and What Do They (Feel They Can) Do about Them?' In *2023 ACM Conference on Fairness, Accountability, and Transparency*, 467–79. Chicago IL USA: ACM, 2023. https://doi.org/10.1145/3593013.3594012.

Zhang, Baobao, Markus Anderljung, Lauren Kahn, Noemi Dreksler, Michael C. Horowitz, and Allan Dafoe. 'Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers'. arXiv, 5 May 2021. http://arxiv.org/abs/2105.02117.

Qualitative studies:

Ali, Sanna J., Angèle Christin, Andrew Smart, and Riitta Katila. 'Walking the Walk of AI Ethics: Organizational Challenges and the Individualization of Risk among Ethics Entrepreneurs'. In *2023 ACM Conference on Fairness, Accountability, and Transparency*, 217–26. Chicago IL USA: ACM, 2023. https://doi.org/10.1145/3593013.3593990.

Chivukula, Shruthi Sai, Chris Rhys Watkins, Rhea Manocha, Jingle Chen, and Colin M. Gray. 'Dimensions of UX Practice That Shape Ethical Awareness'. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020. https://doi.org/10.1145/3313831.3376459.

Groves, Lara, Aidan Peppin, Andrew Strait, and Jenny Brennan. 'Going Public: The Role of Public Participation Approaches in Commercial AI Labs'. arXiv, 16 June 2023. http://arxiv.org/abs/2306.09871.

Holstein, Kenneth, Jennifer Wortman Vaughan, Hal Daumé, Miro Dudik, and Hanna Wallach. '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. Glasgow Scotland Uk: ACM, 2019. https://doi.org/10.1145/3290605.3300830.

Ibáñez, Javier Camacho, and Mónica Villas Olmeda. 'Operationalising AI Ethics: How Are Companies Bridging the Gap between Practice and Principles? An Exploratory Study'. *AI & SOCIETY* 37, no. 4 (December 2022): 1663–87. https://doi.org/10.1007/s00146-021-01267-0.

Kelley, Stephanie. 'Employee Perceptions of the Effective Adoption of AI Principles'. *Journal of Business Ethics* 178, no. 4 (1 July 2022): 871–93. https://doi.org/10.1007/s10551-022-05051-y.

Lu, Qinghua, Liming Zhu, Xiwei Xu, Jon Whittle, David Douglas, and Conrad Sanderson. 'Software Engineering for Responsible AI: An Empirical Study and Operationalised Patterns'. In *2022 IEEE/ACM 44th International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, 241–42, 2022.

https://doi.org/10.1109/ICSE-SEIP55303.2022.9793864.

Madaio, Michael, Lisa Egede, Hariharan Subramonyam, Jennifer Wortman Vaughan, and Hanna Wallach. 'Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support'. *Proceedings of the ACM on Human-Computer Interaction* 6, no. CSCW1 (7 April 2022): 52:1-52:26. https://doi.org/10.1145/3512899.

McNamara, Andrew, Justin Smith, and Emerson Murphy-Hill. 'Does ACM's Code of Ethics Change Ethical Decision Making in Software Development?' In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 729–33. ESEC/FSE 2018. New York, NY, USA: Association for Computing Machinery, 2018. https://doi.org/10.1145/3236024.3264833.

Orr, Will, and Jenny L. Davis. 'Attributions of Ethical Responsibility by Artificial Intelligence Practitioners'. *Information, Communication & Society* 23, no. 5 (15 April 2020): 719–35. https://doi.org/10.1080/1369118X.2020.1713842.

Rakova, Bogdana, Jingying Yang, Henriette Cramer, and Rumman Chowdhury. 'Where Responsible AI Meets Reality: Practitioner Perspectives on Enablers for Shifting Organizational Practices'. *Proceedings of the ACM on Human-Computer Interaction* 5, no. CSCW1 (13 April 2021): 1–23. https://doi.org/10.1145/3449081.

Sanderson, Conrad, David Douglas, Qinghua Lu, Emma Schleiger, Jon Whittle, Justine Lacey, Glenn Newnham, Stefan Hajkowicz, Cathy Robinson, and David Hansen. 'AI Ethics Principles in Practice: Perspectives of Designers and Developers'. *IEEE Transactions on Technology and Society* 4, no. 2 (June 2023): 171–87. https://doi.org/10.1109/TTS.2023.3257303.

Stahl, Bernd Carsten, Josephina Antoniou, Mark Ryan, Kevin Macnish, and Tilimbe Jiya. 'Organisational Responses to the Ethical Issues of Artificial Intelligence'. *AI & SOCIETY* 37, no. 1 (1 March 2022): 23–37. https://doi.org/10.1007/s00146-021-01148-6.

Taylor, Linnet, and Lina Dencik. 'Constructing Commercial Data Ethics'. *Technology and Regulation*, 10 April 2020, 1-10 Pages. https://doi.org/10.26116/TECHREG.2020.001.

Vakkuri, Ville, Kai-Kristian Kemell, Joni Kultanen, Mikko Siponen, and Pekka Abrahamsson. 'Ethically Aligned Design of Autonomous Systems: Industry Viewpoint and an Empirical Study' 27, no. 1 (2022).

Veale, Michael, Max Van Kleek, and Reuben Binns. 'Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making'. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14. Montreal QC Canada: ACM, 2018. https://doi.org/10.1145/3173574.3174014.

Widder, David Gray, and Dawn Nafus. 'Dislocated Accountabilities in the AI Supply Chain: Modularity and Developers' Notions of Responsibility'. *Big Data & Society* 10,

no. 1 (January 2023): 20539517231177620. https://doi.org/10.1177/20539517231177620.

Winecoff, Amy A., and Elizabeth Anne Watkins. 'Artificial Concepts of Artificial Intelligence: Institutional Compliance and Resistance in AI Startups'. arXiv.org, 2 March 2022. https://doi.org/10.1145/3514094.3534138.

Wong, Richmond Y. 'Tactics of Soft Resistance in User Experience Professionals' Values Work'. *Proceedings of the ACM on Human-Computer Interaction* 5, no. CSCW2 (18 October 2021): 355:1-355:28. https://doi.org/10.1145/3479499.

Appendix 2: Survey

Thank you for taking this survey, which investigates attitudes, beliefs and behaviours of tech workers. There are 3 sections, and the whole survey will take 10–15 minutes to complete.

About this study: This study aims to explore ethical decision-making by tech workers in corporate settings, using a quantitative survey. There are many 'small n' qualitative studies exploring extrinsic approaches to ethics (principles, frameworks, toolkits). This research uses quantitative methods to explore intrinsic incentives, motivations, internalised norms and potential barriers across a larger population.

For the purposes of this study, ethical decision-making is defined as critical thinking about a technology's potential effects on people or society, with a motivation to support benefits and mitigate harms. This kind of ethical decision-making (whether or not it results in a positive outcome) looks beyond corporate dynamics or incentives to consider the needs of affected people and broader society (compared with, for example, a technical capability-based approach to risk).

About you: If you've been sent a link to this survey, that's because you've been identified through your job role, or recommended by someone who knows you. If you don't work (or have not recently worked) in a company or public-sector organisation that develops or deploys technologies, please do not start the survey.

About your data: Survey responses collected through Qualtrics are hosted by third-party data centres that are SSAE 16 SOC II certified, and data is stored in line with GDPR requirements. Only the survey owners (researcher and supervisor) can view the responses. The project does not collect, store or process data that could be used to identify people: no names of people or companies are collected, and job role categories are at a high level that would not permit identification of individuals.

Participants can withdraw their data at any time by emailing oeareeve@gmail.com, in which case that data will be deleted from the sample. In addition participants can voluntarily supply an email address, to be stored separately from the study and deleted once the research is complete, in order to be sent a copy of the final research.

This study has ethical approval from Birkbeck College, University of London, reference BBKPOL2022/23-15. Participants may disclose information about their professional lives that is not in the public domain, but published results will be anonymised and there will be no identification of people or companies, so participating in the research does not present a risk to survey respondents.

About me: My name is Octavia Reeve, and I am an early-career researcher and Associate Director at the Ada Lovelace Institute, London. This survey is my personal research, and the results will be explored initially in my MSc Social Research dissertation, which is a preliminary study for a PhD. I expect to publish the results and analysis.

Thank you in advance for your time, expertise and generosity in sharing this information with me. Any questions, please email me at oeareeve@gmail.com.

Section 1: This section asks general demographic questions about you.

Q3 Which business area do you work in?

Agriculture (1)
Construction (2)
Climate tech (3)
EdTech (4)
General-purpose AI (5)
FinTech (6)
Health (7)
Retail & consumer (8)
Security & defence (9)
Social media (10)
Other (11) _____

Q4 Which of the following best describes the technology company you work for?

Small company or start-up < 10 employees (1)
Small-to-medium enterprise 10 – 50 employees (2)
Medium-sized company 50 – 150 employees (3)
Large company or corporation 150 + employees (4)
Public-sector organisation (5)
Other (6) _____

Q5 In which continent is the company you work(ed) for located?

Africa (2)
Asia (1)
Australia (6)
Europe (5)
North America (3)
South America (4)

Q6 Which of these most closely describes your job role?

Software engineer (1)
Data engineer (2)
Data scientist / ML engineer (3)
DevOps / site reliability engineeer (4)
Product manager (5)
Other (6) _____

Q7 How long have you worked in a technology role?

Less than 2 years (1)
2–5 years (2)
6–10 years (3)
11–20 years (4)
More than 21 years (5)

Q8 How old are you?

18–24 years old (1)
25–34 years old (2)
35–44 years old (3)
45–54 years old (4)
55–64 years old (5)
65+ years old (6)

Q9 What best describes your ethnic origin?

Arabic (Palestinian, Egyptian, Syrian, Saudi, any other Arab background) (1)
Asian (Indian, Pakistani, Bangladeshi, Chinese, any other Asian background) (2)

Black / African / Caribbean (3)
Hispanic / Latinx (4)
Mixed – two or more ethnic groups (5)
White (6)
Prefer to self-describe (7) _____
Prefer not to say (8)

Q10 How do you describe your gender?

Female (1)
Male (2)
Non-binary / third gender (3)
Prefer to self-describe (4) _____
Prefer not to say (5)

Q11 What is the highest level of education you have completed?

Completed aged 16 exams / assessment (1)
Completed aged 18 exams / assessment (2)
Post-school vocational qualification (3)
Some university education without qualification (4)
Undergraduate university degree (BA, BSc, etc.) (5)
Postgraduate university degree (MA, MSc, PhD, etc.) (6)
Other (7) _____

Section 2: This section asks about your attitudes and beliefs, and what matters to you in relation to ethical decision-making.

Q13 To what extent do you agree or disagree with these statements?

	Agree (1)	Neither agree nor disagree (2)	Disagree (3)	Don't know (4)
It's important to me to consider possible outcomes for people and society when designing or deploying technologies (1)	○	○	○	○
I am more motivated by meeting goals for my team or company than by outcomes for people or society (2)	○	○	○	○
I believe I have a personal, moral responsibility to produce products that do good rather than harm (3)	○	○	○	○
I think technological innovation is essential to solving big, societal challenges like climate change or the energy crisis (4)	○	○	○	○
I find the amount of information about ethical guidelines, frameworks and toolkits overwhelming (5)	○	○	○	○
It's important to me to do the right thing, and I trust my own judgement to be the best guide of what that is (6)	○	○	○	○
I feel a personal responsibility to stay informed about ethics so I can be alert to potential outcomes for people and society when I make decisions (7)	○	○	○	○
I think there's too much focus on ethics, and we need to get on with building technologies to make the world better (8)	○	○	○	○

Q14 Rate these activities (drag and drop) in order of priority when you're faced with ethical decision-making?

_____ Data privacy and security (1)
_____ Product design and user experience (2)
_____ Developing an innovative product or service (3)
_____ Ensuring all users and affected people can benefit equally (4)
_____ Company revenue and growth (5)
_____ Using ethical principles, frameworks or toolkits (7)
_____ My own sense of what's right and wrong (8)
_____ Following internal company codes and policies (9)
_____ Anticipating and preventing negative outcomes for people and society (10)

Q15 To what extent do you agree or disagree with these statements?

	Agree (1)	Neither agree nor disagree (2)	Disagree (3)	Don't know (4)
I think companies have a responsibility to society, to ensure that their technologies do not have negative consequences (1)	○	○	○	○
I trust the people I work with to think about ethical concerns and consequences, so that I don't have to (2)	○	○	○	○
As long as what my company does is legal, I don't need to think too hard about things that might happen (3)	○	○	○	○
I find academic work is useful for big-picture questions, but I struggle to apply it to what I do at work (4)	○	○	○	○
I don't think you can stop companies prioritising profit, funders or shareholders over outcomes for people and society (5)	○	○	○	○
I don't think companies are well equipped to recognise when the technologies they're developing have potential to harm people or society (6)	○	○	○	○
I don't think government, regulators or industry standard-setters need to get inside companies and see what's going on (7)	○	○	○	○
I think tech workers in companies are already well-equipped and supported to answer ethical questions (8)	○	○	○	○

Section 3: This section asks about your experiences working in a company or corporation that develops or deploys technologies

Q17 Thinking about your workplace, to what extent do you agree or disagree with each of the following statements?

	Agree (1)	Neither agree nor disagree (2)	Disagree (3)	Don't know (4)
Outcomes for people and society are not a consideration in whether a technology is developed or deployed (1)	○	○	○	○
There is a clear process for raising ethical concerns, and it's known to produce positive outcomes when people use it (2)	○	○	○	○
There is a person or team that's responsible for assessing and mitigating the societal outcomes of technology products or services (3)	○	○	○	○
Whether a technology is fair, transparent or accountable is not a factor in whether it's developed or deployed (4)	○	○	○	○
There is clear guidance available about how to develop and deploy technologies ethically (5)	○	○	○	○
Ethics is not something that's practised, it's just a set of principles that people generally overlook (6)	○	○	○	○
People who raise ethical concerns are treated with respect, involved in decision-making and included in the company culture (7)	○	○	○	○
It is possible to raise ethical concerns, but legal or marketing have stronger voices as to which technologies are produced (8)	○	○	○	○

Q18 Rate these activities (drag and drop) in order of priority for your organisation when faced with ethical decision-making?

_____ Data privacy and security (1)
_____ Product design and user experience (2)
_____ Developing an innovative product or service (3)
_____ All users and affected people can benefit equally (4)
_____ Company revenue and growth (5)
_____ Using ethical frameworks or toolkits (7)
_____ The organisation's sense of what's right and wrong (8)
_____ Following internal company codes and policies (9)
_____ Anticipating negative outcomes for people and society (10)

Q19 Rate the significance of these potential barriers faced by tech workers wanting to assess the outcomes of their work on people and society

	Not significant — Significant — Very significant
The pace of product development and deployment	
Lack of education or training	
Lack of organisational guidance about where ethics fits into development or deployment processes	
Revenue and growth targets or incentives	
Lack of authority over decision-making	
Lack of interest from management colleagues	
Lack of incentive or recognition for the importance of ethics work	
A sense that ethical thinking isn't aligned with organisational values or goals	
Lack of resources, e.g. dedicated job role, policies or processes	

Q20 Which of the following has potential to support technology workers to develop and deploy technologies that work for people and society?

	Agree (1)	Neither agree nor disagree (2)	Disagree (3)	Don't know (4)
Government legislation and effective regulation (1)	○	○	○	○
Industry standards and accreditation by a professional association (2)	○	○	○	○
Company mission, values and policies that include ethical position (3)	○	○	○	○
Commercial incentives for ethical development or deployment (4)	○	○	○	○
Support from company leadership (5)	○	○	○	○
Having clear ethical responsibilities in roles and teams (6)	○	○	○	○
Ethical frameworks, principles or toolkits (7)	○	○	○	○
Education and training in ethics and corporate decision-making (8)	○	○	○	○
Employee activism and labour movements (9)	○	○	○	○
Involving affected people in decision-making (11)	○	○	○	○

Q21 Have you experienced decision-making about a technology in your workplace that you felt could have negative outcomes for people or society?

Yes (1)
No (2)
Don't know (3)

Q22 What did you do as a result of that experience? (Select as many answers as apply to you)

Raised concerns with my team, or a member of my team (1)
Raised concerns with a manager or leader (2)
Raised concerns with an internal ethics body (3)
Raised concerns via a bug / ticket / feedback form (4)
Reported concerns to an external person or institution (5)
Considered leaving the company (6)
Did nothing (7)
Other (8) _____

Q23 What were the outcomes of that episode for you?

I was not able to make my team or leadership hear my concerns (1)
My concerns about the product / service / development or deployment process were heard and changes were made (2)
My concerns were heard but other concerns or incentives were prioritised (3)
Raising concerns changed the outcomes of the product for people and society (4)
Raising concerns adversely affected my progression in the company (5)
I left the company (6)
I kept my views to myself and took no action, but moved on (7)
I kept my views to myself, took no action and wished I had (8)

Q24 Have you left a tech role in the last year?

No (1)
Yes (2)

Q25 Why did you leave your most recent tech role?

Career choice, e.g. change company / new role (1)
Short-term contract ended, e.g. freelance / consultant / contractor (2)
Redundancy / lay-off (3)
Ethical choice, e.g. didn't want to work on a project / disparity with company mission or values (4)
Other (5) _____

Q26 Is there anything else you'd like to add?_____

Q27 Do you want to be sent a copy of the final research?

No – in which case thank you very much for your responses (1)
Yes – and I'm happy to add my email address here: (2)
_____
Yes, and I will email oeareeve@gmail.com to request a copy (3)