**Who is an AI Ethicist? An Empirical Study of Expertise, Skills, and Profiles to Build a Competency Framework**

Mariangela Zoe Cocchiaro[1,2], Jessica Morley[5], Claudio Novelli[1], Enrico Panai[3], Alessio Tartaro[4], Luciano Floridi[5,1,]

[1] Department of Legal Studies, University of Bologna, Via Zamboni, 27/29, 40126, Bologna, Italy

[2]Institute of Philosophy, Jagiellonian University, Gołębia 24, 31-007 Krakow, Poland

[3] Linguistic Computing, University Cattolica del Sacro Cuore di Milano, Largo Fra Agostino Gemelli, 1, 20123 Milano, Italy

[4] Department of Humanities and Social Sciences, University of Sassari, Via Roma, 151, 07100 Sassari, Italy

[5] Digital Ethics Center, Yale University, 85 Trumbull Street, New Haven, CT 06511, US

Email for correspondence: enricopanai@gmail.com

**Abstract**

Over the last decade the figure of the AI Ethicist has seen significant growth in the ICT market. However, only a few studies have taken an interest in this professional profile, and they have yet to provide a normative discussion of its expertise and skills. The goal of this article is to initiate such discussion. We argue that AI Ethicists should be experts and use a heuristic to identify them. Then, we focus on their specific kind of moral expertise, drawing on a parallel with the expertise of Ethics Consultants in clinical settings and on the bioethics literature on the topic. Finally, we highlight the differences between Health Care Ethics Consultants and AI Ethicists and derive the expertise and skills of the latter from the roles that AI Ethicists should have in an organisation.

**Keywords**

AI Ethicist, ethics of AI, expertise, practical moral expertise, skills

## 1. Introduction

The lack of expertise within a domain is generally regarded as a significant risk factor. Most risk management frameworks emphasise the pivotal role that domain experts play in offering crucial

guidance (ISACA, 2012; ISO, 2018; ISO/IEC, 2023). *A fortiori,* this is true in the field of Artificial Intelligence, where "identifying and managing AI risks and potential impacts […] requires a broad set of perspectives" (NIST, 2023).

As AI systems become increasingly integral to our society, the scrutiny of their ethical implications grows, highlighted by incidents like biased algorithms causing significant public concern.[1] AI risk management frameworks now require integrating ethical expertise (Tartaro et al., 2024). To address this need, a new professional role has emerged: the AI Ethicist (Rismani & Moon, 2023a). Some AI Ethicists have become leading figures in the public discussion around ethical concerns, and there is a rising demand for them within organisations, as highlighted by the number of job advertisements in the field (Rismani & Moon, 2023b).

Despite the growing interest in this subject, only a few studies have analysed this professional figure. Some scholars highlight that AI Ethicists often need a mandate, support, and resources from the very organisations that create ethical issues (Sætra et al., 2022b). Others have highlighted how AI Ethicists often face a dilemma: communicating about an identified problem might diminish the likelihood of addressing that issue successfully (Sætra et al., 2022a). Finally, some question the necessity of AI Ethicists altogether (Hagendorff, 2023).

However, the skills and knowledge required to assess AI Ethicists' expertise and role within organisations remain underexplored. This gap leads to companies filling these positions based on their specific needs and perspectives, often relying on existing employees or aspiring AI ethicists. For some, the role involves developing technical tools and workflows for monitoring AI ethics issues such as bias and explainability, requiring skills in software development and system safety design. Others view AI Ethicists as responsible for developing and evaluating ML models against fairness metrics, necessitating advanced analytical and programming skills (Rismani & Moon, 2023b). Often, AI Ethicists are seen as legal consultants focusing on compliance with higher-level norms, potentially leading to legal advisors' absorption of this role in the future.

In short, there is a multitude of heterogenous roles – as many as seven, as recently pointed out by Rismani and Moon (2023) – that AI Ethicists currently (are expected to) take on: (1) technical

[1] Examples by now "classic" include the misuse of data for political purposes as revealed by the Cambridge Analytica scandal (Wilson, 2019), the SyRi system in the Netherlands being dismissed for biased profiling (Algorithm Watch, 2020), and Amazon's AI recruitment tool discriminating against women (Dastin, 2022).

researcher (e.g., computer scientist), (2) data scientist (3) engineer, (4) non-technical researcher (e.g., social scientist), (5) policy analyst, (6) manager, (7) director.

The disparate roles ascribed to AI Ethicists reflect a lack of consensus on their professional identity and roles, leading to at least three problems. First, it obscures their primary mission of bridging the gap between ethical frameworks and real-world AI applications. The absence of an accreditation program exacerbates this, making it difficult to implement responsible AI practices (Schiff et al., 2021).[2] Second, the disconnect between AI ethics and the job market can undermine the field's credibility. Many new technology regulations use flexible, risk-based legal frameworks and adopt a principle-based approach due to extensive negotiations, broad civil society participation, and the need for consensus.[3] Yet, without professionals to implement these principles, this approach might be seen as ineffective (Chamberlain, 2023; Novelli et al., 2024). As we will discuss later, job vacancies for AI Ethicists show a market need for applying these principles in the real world. Finally, the varied roles hinder the emergence of a coherent professional identity, essential for establishing norms, standards, and best practices.

This paper addresses this topic by combining empirical and normative approaches. Empirically, it examines the emerging role of AI Ethicists by analysing LinkedIn data to track the growth and distribution of professionals in this field across Europe. This data-driven analysis provides insights into the current landscape and trends. Normatively, the paper argues for the essential expertise and moral authority that AI Ethicists should possess, drawing parallels to Health Care Ethics Consultants – henceforth, HC Ethicists (Brummett, 2023; Watson & Guidry-Grimes, 2018). It advocates for a structured competency framework to define and standardise the role of AI Ethicists, emphasising their critical function in bridging ethical principles with practical application in AI development and deployment.

The article is structured into four more sections. Section two outlines the methodology used to identify the expertise of AI Ethicists, drawing on the well-established role of HC Ethicists. Section three explores a threefold definition of moral expertise, focusing on practical moral expertise. We discuss the benefits of academic moral expertise and argue that performative moral expertise is crucial for AI Ethicists. We examine practical moral expertise, suggesting that the facilitator role best suits

[2] More generally, this also represents an additional barrier to the operationalisation of AI ethics, since many AI practitioners would find it useful to consult AI ethicists, but few of them have this kind of consultation available (Morley et al., 2023).

[3] When specific rules are hard to agree on, consensus is often achieved at the level of broader or intermediate principles that balance competing needs, such as innovation and rights protection. This enables the regulatory cost-benefit analysis to adapt more effectively to real-world demands.

Ethics Consultants. Section four develops a competency framework for AI Ethicists inspired by the medical sector. Section five concludes the paper.

## 2. Methodology: Empirical and Normative

AI Ethics support is an underexplored phenomenon as a recently emerged consultancy service. We lack sufficient data yet to describe and understand this job market, and it is unclear whether this new professional figure has come to stay, even if there is currently a significant demand for AI Ethicists.

To understand the growth of individuals active in this field, we collected data through LinkedIn searches in April 2023. We used the keywords "AI ethics" and "digital ethics" to identify relevant profiles, which were then geographically filtered to focus on European countries. The presence of the AI Act dictated the geographical choice and, hence, a more complex interaction between legal and ethical expertise about AI. To improve accuracy, we eliminated duplicate profiles found under both keywords and profiles where the keywords were not mentioned in the user's headline, personal summary, or licenses and certifications sections. We conducted the same search a year later, in April 2024, to analyse historical trends in the field. The results of the April 2024 search are visualised in Figure 1, while Table 1 provides a comparative analysis between April 2023 and April 2024. Table 2 provides a detailed breakdown of the variation for each keyword over the period.

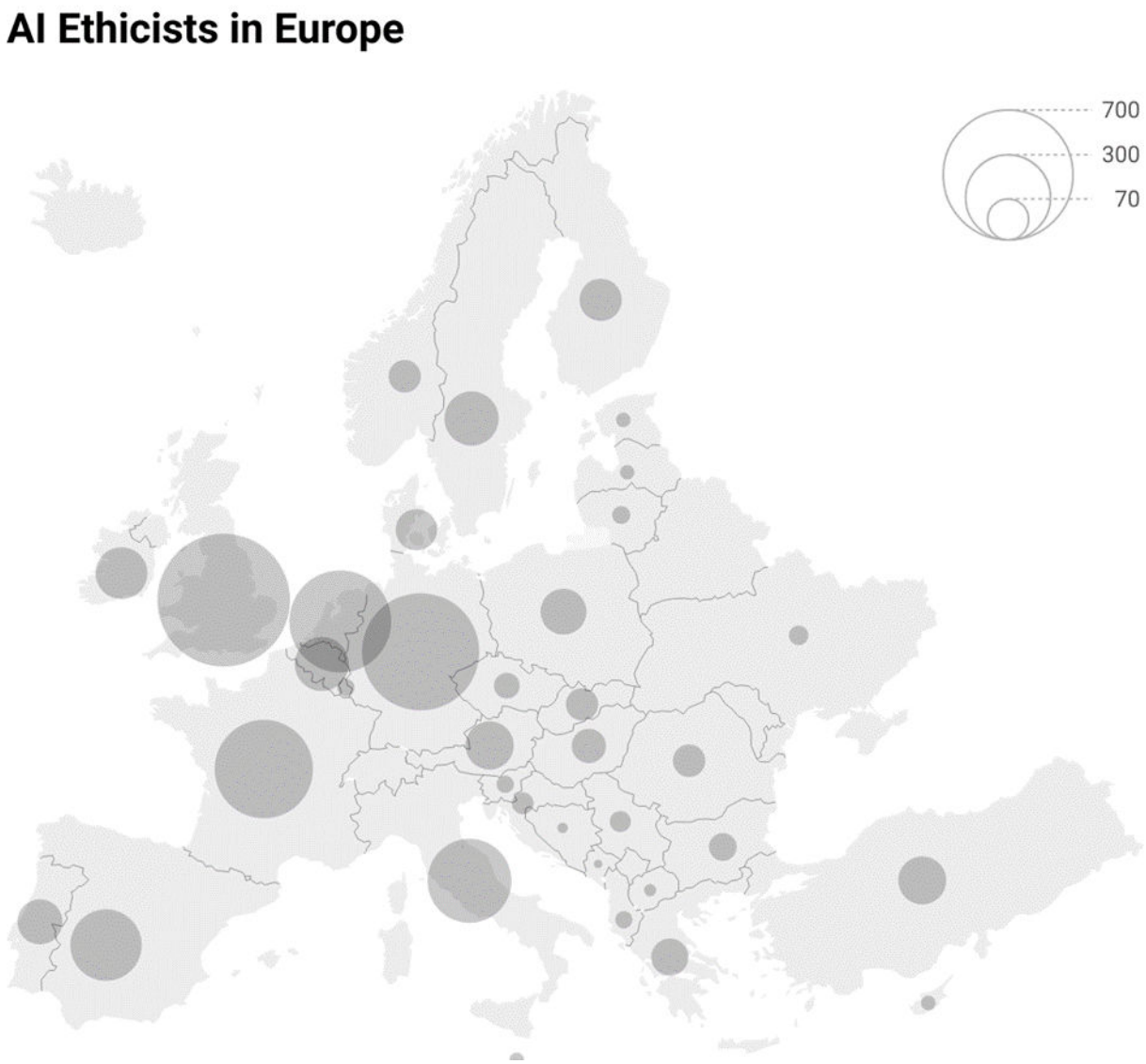


*Figure 1 - Distribution of AI Ethicists on LinkedIn in Europe in April 2024*

| Country | Apr. 2023 | Apr. 2024 |
|---|---|---|
| Albania | 2 | 8 |
| Austria | 71 | 87 |
| Belgium | 118 | 111 |
| Bosnia and Herzegovina | 0 | 2 |
| Bulgaria | 22 | 25 |
| Croatia | 13 | 14 |
| Cyprus | 4 | 5 |
| Czech Rep. | 17 | 21 |
| Denmark | 84 | 62 |
| Estonia | 7 | 5 |
| Finland | 123 | 65 |
| France | 351 | 393 |
| Germany | 649 | 566 |
| Greece | 30 | 49 |
| Hungary | 32 | 41 |
| Ireland | 116 | 102 |
| Island | 0 | 0 |
| Italy | 58 | 283 |
| Latvia | 3 | 5 |
| Lithuania | 6 | 9 |
| Luxembourg | 9 | 8 |
| Malta | 4 | 5 |
| Moldova | 0 | 0 |
| Montenegro | 0 | 1 |
| Netherlands | 414 | 424 |
| North Macedonia | 0 | 3 |
| Norway | 18 | 36 |
| Poland | 58 | 78 |
| Portugal | 46 | 75 |
| Romania | 25 | 36 |
| Serbia | 7 | 13 |
| Slovakia | 18 | 36 |
| Slovenia | 6 | 8 |
| Spain | 171 | 202 |
| Sweden | 102 | 112 |
| Turkey | 31 | 86 |
| Ukraine | 5 | 11 |
| United Kingdom | 687 | 731 |

*Table 1 - List of AI Ethicists in Europe in April 2023 and April 2024*

| KEYWORDS | Apr. 2023 | Apr. 2024 | Variation |
|---|---|---|---|
| "AI ethics" | 2342 | 2794 | +19% |
| "digital ethics" | 965 | 924 | -4% |
| Total | 3307 | 3718 | +12% |

*Table 2 – Variation for the keywords "AI ethics" and "digital ethics".*

While these findings offer valuable insights, they are subject to some limitations. First, our search methodology may underestimate the field's size, not capturing some of the people working in the field. Since we rely on the keywords "AI ethics" and "digital ethics" to identify relevant profiles, we may exclude individuals who work on ethically relevant projects but do not use these specific terms. This misses professionals like lawyers and engineers who address ethical issues in AI (e.g., mitigating bias in recruitment algorithms) but would not identify themselves as "AI ethicists". We are uncertain whether LinkedIn's algorithm searches only for exact phrases or performs some form of inference

based on the query. Consequently, we may potentially miss profiles that contain related phrases, such as "I conduct research on the ethical implications of digital technologies - including AI," but do not explicitly mention "AI ethics" or "digital ethics."

Second, our exclusive reliance on LinkedIn data might have excluded professionals who are not active on this social media platform. Third, using English search terms could have introduced a linguistic bias, potentially favouring profiles from English-speaking countries or geared towards an international audience. Fourth, these findings should be considered as estimates rather than precise figures, as the data may capture individuals who are not dedicated only to AI ethics as their profession. Consequently, there is a limited likelihood that this study could be repeated and produce the same results. Despite these limitations, the results point towards a burgeoning community of professionals actively involved in AI ethics — whom we designate as AI ethicists.

We saw above the results for "self-description" on LinkedIn: how many people consider themselves AI Ethicists. Turning to the job market, in June 2023, we conducted a preliminary online search to ascertain the existence of such a market and began exploring it. We looked for online postings for the position by using the keywords "AI Ethicist jobs", "AI Ethics jobs", and "Responsible AI jobs". We restricted the search to Google for Jobs, LinkedIn and Indeed platforms for only five English-speaking countries: the UK, the US, Canada, Australia, and New Zealand. We used a VPN service to simulate location presence in each country of interest. Specific job titles, such as data scientist, reporter, professor, and research scientist, were deliberately excluded to focus on roles directly relevant to AI ethics. This strategy ensured a targeted analysis of job market demands for AI Ethicists. We found that the companies advertising relevant positions were multinational corporations and start-ups. Our search ultimately identified 26 active job listings for the position of AI Ethicists. A follow-up search in April 2024 revealed a growth in these job positions. While Canada, Australia, and New Zealand showed no results, we found a significant increase in the US (22 listings) and UK (13 listings). Again, it is essential to acknowledge that these results are estimates, but they point towards an emerging market for AI ethics professionals.

In the job postings reviewed, AI Ethicists were assigned several core responsibilities, including creating Ethical AI solutions, conducting research and analysis on AI Ethics, and overseeing AI projects' monitoring, evaluation, and compliance. As our preliminary research showed, the description of the AI Ethicist resembles that of a consultant tasked with offering advice, guidance, and expert knowledge to enhance clients' ethical decision-making capabilities within the AI domain. But what should such expert knowledge include?

To answer this normative question, we drew on the best practices and extensive multi-decade experience of professional medical ethicists in the US. Among all fields of applied ethics, bioethics appears to have the closest affinity with AI ethics in dealing ecologically with new forms of agents, patients, and environments (Floridi, 2013; Floridi & Cowls, 2019).

When healthcare ethics consultancy first appeared in the US, it was an unregulated field, and so was the specific role of the Ethics Consultant within an organisation. Then, its presence in American hospitals began to increase, and with it arose the need for standards and a principled discussion of the type of expertise possessed by this emerging figure. This led the American Society for Bioethics and Humanities (ASBH) to present its Core Competence Framework (ASBH Task Force, 1998; ASBH Task Force, 2011), a comprehensive framework of the required expertise for HC Ethicists in hospitals. The ASBH's framework for HC Ethicists comprises three categories of skills (assessment, process, interpersonal) and knowledge areas categorised into basic and advanced levels.

To develop a similar framework, we analysed the selected job offers, identified the tasks, and inferred the necessary skills and knowledge to perform them. This bottom-up approach first allowed us to identify and categorise core skills and supporting subskills per ASBH's skill categorisation and then to specify the knowledge required for an AI Ethicist.

## 3. The practical moral expertise of AI Ethicists

AI Ethics serves as the guiding framework for moral behaviour in developing, applying, and revising AI technologies. In general, an AI Ethicist's role is akin to that of a moral reasoner, that is, someone engaged in a particular form of practical reasoning that focuses on right and wrong. The role requires applying moral principles and established values to specific situations to identify the most ethical course of action (Richardson, 2003). To achieve this, AI Ethicists draw upon conceptual analysis and moral norms, principles, and reasoning, to ensure that AI aligns with human rights, values, and interests.

An important aspect of AI Ethicists' role is improving legal compliance, which often serves to operationalise general moral values into specific legal standards. For example, the concept of reparative justice is concretely manifested in the legal regulations governing tort liability. However, as a multidisciplinary branch of applied ethics (like Medical Ethics, Bioethics, Environmental Ethics, etc.), AI Ethics involves more than just legal compliance, considering what ought and ought not to be done over and above existing regulations (soft ethics in (Floridi, 2018)). This means doing more than what is strictly required by the law, or less than is allowed. In short, the AI Ethicist is an individual

who has the *moral expertise* to provide correct answers to questions that fall under the domain of AI Ethics.[4]

According to a substantial body of literature on bioethics, there are at least three distinct, though compatible, kinds of moral expertise: *academic* moral expertise, *performative* moral expertise, and *practical* moral expertise (Brummett, 2023; Watson & Guidry-Grimes, 2018). Our next goal is to explore the potential reusability of specific elements of this framework within the domain of AI Ethicist's expertise. Before doing so, it may be worth clarifying two points.

First, any moral expertise might prompt organisations to offload all ethical decisions onto the AI Ethicist, leading to moral disengagement. In the latter scenario, a "wicked environment" could potentially make ethicists targets of any blame for undesirable outcomes. As will be more apparent in the following sections, we are not arguing for this delegation of responsibilities; actually, we are criticising it. Rather than deciding, AI Ethicists give advice and are thereby responsible for the given advice, while decision-makers remain fully responsible for the actions implemented.

Second, the role and profile of an AI Ethicist are a matter of current debate but should not be confused with those of an "evangelist". As our online research suggests, this role is sometimes ascribed to the AI Ethicist (Microsoft, 2023). We recommend against this approach for at least two reasons:

i. Moral expertise goes beyond simply classifying things as right or wrong. It entails a nuanced and contextualised approach that considers the potential consequences of various specific actions to ensure decisions align with established ethical frameworks, whether legal or socially constructed. This analysis requires a deep understanding of legal rules and principles, including how courts have interpreted them (case law) and the ability to weigh the potential benefits for one group against the potential harms to another. Additionally, it necessitates the capacity to identify and prioritise solutions that minimise overall suffering. If AI ethicists are expected to act as evangelists for the employing organisation and its technology, their ability to assess and challenge their organisation's practices critically may be compromised. For instance, they may feel compelled to promote and support their organisation's products and services, despite emerging ethical, societal, and environmental concerns. This would make them *de facto* members of the organisation's marketing team. Consequently, AI ethicists' independence, credibility, and professional stance might be compromised, and they may be induced into a form of conceptual

[4] We are operating under the assumption that moral statements are apt for correctness and incorrectness. For more details on what we mean by it, see: (Floridi, 2011).

conservatism, persisting in their beliefs despite contrary evidence (Nissani, 1990); and most critical

ii. AI ethicists acting as evangelists face a fundamental conflict of interest. On the one hand, their mission is to identify and prevent or mitigate the negative ethical and social impacts of AI technologies. This often requires transparency with internal and external stakeholders regarding potential issues. On the other hand, their role as evangelist for the technology and the organisation might discourage (or even prevent) them from taking positions that could negatively impact their employer, e.g., economically or reputationally. In this scenario, serving one interest would likely compromise the other. As a result, AI ethicists' activities can be disrupted, compromising their integrity and credibility.

Let us turn now to the kinds of moral expertise that an AI ethicist may be required to have by drawing on the distinction among the three, possibly overlapping, kinds of expertise identified in bioethics (Brummett, 2023; Watson & Guidry-Grimes, 2018).

### 1. 3.1. Academic moral expertise

Academic moral expertise is most commonly manifested in scholars with advanced degrees in moral philosophy who actively engage in academic life. These individuals explore, among others, questions such as whether moral truths exist, whether these truths are accessible to us, the nature of moral beliefs, whether some moral value is strong or weak normatively, and differences between thin and thick moral terms. As such, they can explain the intricacies and consequences of different normative moral theories, sketch out ongoing disputes surrounding the latter, and help others comprehend the significance of these discussions (Brummett, 2023; Watson & Guidry-Grimes, 2018).

However, in their discussions, academic scholars often abstract from the hardships of daily life and keep some variables fixed. With an example from bioethics: in the abortion debate, moral philosophers frequently talk about the nature of (legal) personhood, individual rights, and the types of moral standing. All these factors are relevant to practical moral decisions about abortion. Still, they certainly do not (need to) exhaust the range of variables that an ethical consultant must consider in clinical cases. When considering abortion in complex cases (i.e. late termination of pregnancy, the mother is not mentally competent to make a decision or opposes abortion even though it threatens her life), the panel of decision-makers of which the HC Consultant might be a member must take into account issues such as decision-making capacity, religious beliefs, health conditions, cultural values,

and non-identity (Hope & McMillan, 2012). In all clinical cases referred to HC Ethicists, nothing can be held fixed or suspended as a moral philosopher might do in the classroom.

As a result, while possibly necessary, academic moral expertise is only one of many kinds of moral expertise required for the HC Ethicist. Just like scholars in the field, AI Ethicists need to be able to answer questions about the terms, claims, and arguments in AI Ethics. But, just like HC Ethicists, academic expertise cannot be the only kind of moral expertise required.

### 2. 3.2. Performative moral expertise

Performative moral expertise refers to the ability to make good moral decisions for oneself, i.e., doing the right thing in specific circumstances (Burch & The Hegeler Institute, 1974; Hulsey & Hampson, 2014). Like academic moral expertise, some degree of performative moral expertise is necessary for the role of an ethicist. After all, clients or employers can hardly place their trust in an ethicist who demonstrates an inability to make sound ethical decisions independently. Performative moral expertise can boost the ethicist's reputation, increasing their appeal to potential employers. This should not be confused with two other issues: a requirement for consistency and uniqueness. On the one hand, an ethicist whose behaviour does not align with their advice might be akin to an overweight cardiologist. Although they recognise that obesity is a risk factor for heart disease, their circumstances do not correlate with their professional competency. On the other hand, it has been shown that an ethicist's moral decision-making capacity is comparable to anyone else's (Hagendorff, 2023). This indicates that an AI Ethicist should have at least the same performative moral expertise as an average individual.

### 3. 3.3. Practical moral expertise and the role of Ethics Consultants

Practical moral expertise is the ability to provide justified moral recommendations (i.e., statements about what is ethically permissible, preferable, obligatory, or prohibited). It refers to the application of ethical principles and judgments effectively in real-world scenarios, bridging theory with action to resolve moral problems. In medical ethics, the mere possibility of its existence has sparked numerous philosophical discussions to the point of dubbing it the "basilisk"[5] of bioethics (Brummett, 2023; Watson & Guidry-Grimes, 2020, 2018). One of the most widely investigated questions on the matter came up when Ethics Consultants started to become increasingly common in American hospitals, and it amounts to asking whether Ethics Consultants *can* give justified moral recommendations that might

[5] A basilisk is a legendary serpent monster that can kill with its fatal gaze.

inform the decision-making process.[6] In this context and depending on the organisation, the compound 'Ethics Consultants' can refer to the responsibilities and tasks assigned to an individual and the functions and activities performed by a group of people working collectively.

This debate has also significantly informed the core competencies for the Health Care Ethics Consultation report published by the American Society for Bioethics and Humanities in 1998 and 2011 (ASBH Task Force, 1998; ASBH Task Force, 2011). Both editions of the report distinguish among three different roles that can be ascribed to HC Ethicists: the roles stem from three different approaches to the profession, and, in turn, these approaches roughly follow from three families of positions in the debate.

The first is the "authoritarian approach". Heavily informed by the positive views, it posits that the Ethics Consultant *can* give justified moral recommendations. In this approach, the role of the consultant is as the primary – if not the only – moral decision maker whose expertise in ethical analysis might grant them a sort of moral "hegemony".

The second is the "pure consensus approach". Informed by negative views, this approach posits that the Ethics Consultant *cannot* give justified moral recommendations. Under this approach, their sole role is to reach an agreement among involved parties (ASBH Task Force, 1998; ASBH Task Force, 2011).

The third is the "ethics facilitation approach". Informed by positive views, ethicists apply established bioethical consensus to define the spectrum of ethically acceptable options for a particular case. This approach involves using a collective agreement or understanding within bioethics to guide decision-making processes. The role of the consultant is reduced to a facilitator who helps elucidate the issues, aids effective communication, and integrates different stakeholders' perspectives. Ethicists may listen to the moral complaints of all the relevant stakeholders and help them identify their values or commitments. Additionally, their knowledge of ethical arguments may help them point out unrecognised implications of the stakeholders' views. However, in line with negative views, ethicists do not tell stakeholders what to do in cases of genuine uncertainty (Brummett, 2023).

According to the ASBH, the ethics facilitation approach is the most appropriate for Health Care Ethics Consultations. This method is shaped by its operational context and is distinguished by two defining characteristics: "i) identifying and analysing the nature of value uncertainty, and ii) facilitating the building of a principled ethical resolution" (ASBH Task Force, 2011). Eventually, it may be

[6] The question continues to fuel debate, with the family of so-called 'positive views' affirming that ethicists can give justified moral recommendations, while the family of so-called 'negative views' deny it (Brummett & Salter, 2019).

assumed that, akin to HC Ethicists, AI Ethicists are also envisioned to undertake the role of facilitators in the panel of decision-makers of which they are members. However, in the next section, we argue that AI Ethicists should not be just facilitators, and that this is one of the main differences between the two professional figures.

### 4. 3.4. Beyond the facilitator: AI Ethicist as a researcher and educator

Several distinctions between HC Ethicists and AI Ethicists suggest that the facilitator role may not be fully adequate. First, the rapid evolution and increasing complexity within the AI sector, compared to the decades of experience HC Ethicists hold in a highly regulated medical field, places AI Ethicists in the position of evaluating new technologies in an area where regulation often lags. This situation leads them to navigate the legislative grey, fillable or empty areas created by the systems' swift development and complex nature, as outlined by the soft ethics approach (Floridi, 2018). A direct consequence of this first point is that organisations still need to mature in integrating technologies or the ethical issues associated with them. While HC Ethicists have had decades to establish their roles within hospitals, AI Ethicists find themselves in an unstable work environment where their roles are not yet fully recognised or sufficiently valued.

There are also more subtle differences between the HC Ethicists and AI Ethicists' work. Although both deal with complex multi-agent systems and diverse stakeholders, the nature and scale of the ethical challenges they face differ. HC Ethicists often work within established frameworks and protocols, even when dealing with complex scenarios such as organ donation or resource allocation. In contrast, AI Ethicists frequently encounter novel ethical dilemmas arising from emerging technologies, where established guidelines may be lacking or insufficient. Moreover, the potential impact of AI systems can be more far-reaching and less predictable than many healthcare interventions. While healthcare decisions typically affect individuals or specific groups, deploying AI systems can have scalable effects that impact entire societies or even the entire informational ecosystem or infosphere (Floridi, 2013). This broader scope of influence adds complexity to the ethical considerations AI Ethicists must navigate.

Based on these observations, it may be necessary to enhance the facilitator model by ensuring that the AI Ethicist undertakes at least two additional roles: researcher and educator. AI Ethicists should be researchers if they aim to cope with AI's rapid evolution and increasing complexity. Engaging actively in research enables them to remain abreast of innovative trends, deepen their

insights, and contribute to advancing AI ethics. AI Ethicists should also be educators to advance companies' maturity levels. For example, their educational expertise can foster the development of instructional materials and the facilitation of workshops for employees, developers, managers, and directors. Through education, AI Ethicists can cultivate a shared ethical language and increase the ethical awareness of the stakeholders (Morley et al., 2021).

## 4. Knowledge and skills of the AI Ethicist

Three elements underpin expertise: skills, knowledge, and attitude. In this article, we focus on skills and knowledge in alignment with the ASBH's framework, and we exclude attitude because it is a criterion that is currently difficult to measure. In a more standardised way, attitude could be understood as adherence to professional practices or established rules. For example, it could help circumvent conflicts of interest when promoting the ethical integrity of an AI system. However, further studies regarding the relevance of attitude as a criterion of expertise are needed.

To clarify the knowledge and skills required for an AI Ethicist acting as facilitator, researcher, and educator, we will first link specific tasks to each role. Then, we will identify the skills necessary to perform these tasks effectively. Finally, we will isolate ten critical areas of knowledge that inform and guide the actions of AI Ethicists.

### 5. 4.1. The tasks of the AI Ethicist

AI ethicists may be expected to perform many tasks, often leading to unrealistic expectations. To narrow the scope of responsibilities for AI ethicists, we identified a core set of tasks, outlined in Table 3. Drawing from job listings and the literature, we applied two criteria to identify these tasks. First, we only considered tasks consistent with the three key roles of AI ethicists identified in the previous section: facilitator, researcher, and educator. Accordingly, we excluded tasks typically associated with other roles, such as manager or director, as these are not necessary part of the remit of AI ethicists. Second, we focused on tasks that can be performed specifically under an individual's expertise in AI ethics. Consequently, we excluded some tasks for which specialised expertise of AI ethics is not strictly necessary. Through this process, we have identified the following as the main core tasks for AI ethicists, with a further distinction between two facilitator's roles:

- Fh = facilitator who helps identify and analyse the nature of value uncertainty,
- Fo = facilitator who oversees the building of a principled ethical resolution,
- R = researcher, and

- E = educator.

| | |
|---|---|
| Fh-1 | Conducting ethical assessments by evaluating AI technologies, applications, and practices to identify potential ethical and societal concerns and risks. |
| Fh-2 | Managing trade-offs by identifying, analysing, and making decisions about competing priorities and values throughout the lifecycle of AI technologies, prioritising and weighing the pros and cons of different options. |
| Fh-3 | Conducting policy analysis by reviewing, evaluating, and providing recommendations on AI-related guidelines, standards, and regulations. |
| Fo-1 | Leading ethical mediation by facilitating discussions and resolving conflicts on ethical matters in AI development, deployment, and usage. |
| R-1 | Developing a Code of Ethics by creating a set of guiding principles and practices that govern the ethical development, deployment, and use of AI technologies. |
| R-2 | Developing, implementing, and maintaining a Code of Data Ethics by creating a set of guidelines and principles that govern the generation, collecting, recording, curation, processing, dissemination, sharing, and use of data. |
| R-3 | Producing peer-reviewed publications, white papers, memos and/or participating in research projects or attending conferences. |
| E-1 | Training the staff by designing, developing, and delivering educational programs, workshops, or courses that raise awareness and improve understanding of ethical and societal issues in AI. |
| E-2 | Conveying AI ethics-related information, insights, and recommendations to a wide range of internal and external stakeholders. |

*Table 3 - Identified tasks for AI Ethicists*

Based on the listed tasks, we identified skills and knowledge areas through a two-fold approach. Some of the skills and knowledge areas required of the AI Ethicist were inferred directly. Consider, for example, the evaluation of AI technologies that the AI ethicist is supposed to conduct as per the task Fh-2. A necessary condition for the successful completion of this task, among others, will be the knowledge of the AI system to be assessed, as per K-5 in Table 5 below. Other skills and knowledge areas required of the AI Ethicist were identified indirectly by drawing on the HC Ethicist's core skills and knowledge areas listed by the ASBH. Consider, for example, the ability to "identify which

individuals (e.g., patient, healthcare professionals, family) need to be involved in a consultation" – one of the HC Consultants' core skills (ASBH Task Force, 2011). *Mutatis mutandis*, this skill seems necessary for AI Ethicists as well, particularly concerning task Fo-1: how could AI Ethicists facilitate discussions and resolve conflicts on ethical matters without the ability to identify the relevant stakeholders in the first place?

Thanks to such a two-fold approach, we outlined a core set of essential knowledge and skills. However, it is important to acknowledge that individual tasks may necessitate combining multiple knowledge and skills. Conversely, the same knowledge and skills may be applicable across various tasks. Consequently, a one-to-one mapping of tasks to specific sets of knowledge and skills can be overly cumbersome and unnecessarily redundant. To address this, we present the identified skills and knowledge independently of the tasks while providing a thematic categorisation for better organisation.

### 6. 4.2. The skills that underpin the AI Ethicist's expertise

The investigation of the tasks led us to infer a variety of 'Core Skills Requirements,' detailed in Table 4, along with supporting subskills essential for underpinning the former. We adopted the same classification into three distinct categories of skills as proposed by the ASBH framework:

- (A) "Ethical assessment and analysis skills", e.g., when an AI Ethicist has the skill to review a standard
- (P) "Process skills", e.g., when an AI Ethicist has the skill to identify a category of people at risk
- (I) "Interpersonal skills", e.g., when an AI Ethicist has the skill to deliver educational programs to staff (regardless of their lack of interest in the subject matter)

| | |
|---|---|
| A-1 | Identify the nature of the value uncertainty or conflict for all possible cases of consultation |
| A-2. | Access relevant ethics literature, policies, guidelines, and standards |
| P-1 | Determine the stakeholders who are affected and evaluate the risk magnitude |
| P-2 | Facilitate the implementation of the chosen option |
| P-3 | Communicate and collaborate effectively with all the individuals, departments or divisions involved in the ethical consultation |
| P-4 | Facilitate formal meetings |
| P-5 | Document and communicate the consultation activities |

| | |
|---|---|
| P-6 | Identify system issues |
| I-1 | Listen well and communicate interest, respect, and empathy to the affected stakeholders |
| I-2 | Educate the affected stakeholders regarding the ethical concerns from which the need for the consultation arises |
| I-3 | Train the company's staff to identify ethical instances |
| I-4 | Represent the affected stakeholders |
| I-5 | Avoid radicalisation in the ethical debate |
| I-6 | Create an environment where ethical reasoning can flourish |

*Table 4 - Core Skills Requirements for AI Ethicists*

The study delineated a spectrum of secondary but crucial skills vital for reinforcing the foundational skills.

For (A) "Ethical assessment and analysis skills", the ability to discern and gather relevant data; distinguish the ethical dimension of the consultation from other overlapping dimensions, such as the legal one; clearly articulate the ethical concerns and the most important ethical questions; identify various assumptions of the involved parties; identify relevant beliefs and values of the involved parties; identify their own relevant moral values and intuitions and how these might influence the process or analysis; access relevant ethical knowledge; clarify and explain ethical concepts; critically evaluate and use knowledge of AI ethics, law, policy, and professional codes, if any; apply relevant ethical considerations; identify and justify a range of acceptable options and their implications; evaluate evidence and arguments for and against different options; research (among others) peer-reviewed academic journals and books using online databases and/or libraries; recognise and acknowledge personal limitations and possible areas of conflict between personal views and one's role as a consultant.

For (P) "Process skills", the ability to identify the stakeholders and possibly involve them in the consultation; use the structures and resources of the company to facilitate the implementation of the chosen option; document consultation clearly and thoroughly; inform the relevant parties of the results of the consultation promptly; communicate and collaborate effectively with other responsible individuals, departments, or divisions in the company; identify underlying system issues and bring them to the attention of the company's individuals or divisions who can address them; facilitate formal meetings by starting with introductions and goal setting, maintaining focus, establishing timelines for action items, and recognising the need for follow-up meetings:

For (I) "Interpersonal skills", the ability to articulate complex ideas and recommendations effectively; actively listen and communicate empathy and respect to the affected stakeholders; educate involved parties regarding the ethical dimension of the consultation; represent the moral views of involved parties; design, develop and teach well-structured, engaging, and effective training materials and curricula tailored to the needs of the target audience. Both core skills and subskills are instrumental in articulating and showcasing the expertise of AI Ethicists.

### 7. 4.3. The knowledge areas that underpin the AI Ethicist's expertise

AI ethicists need the relevant knowledge to develop and enhance their skills and educate, inform or guide their activities. We identified ten relevant areas for professional AI Ethicists. Table 5 summarises the essential knowledge for AI ethicists.

| | |
|---|---|
| K-1 | Moral reasoning and ethical theory |
| | ● Consequentialist and non-consequentialist approaches (e.g., utilitarian, deontological approaches, natural law, communitarian, and rights theories).<br>● Virtue and feminist approaches.<br>● Principle-based reasoning and case-based approaches.<br>● Related theories of justice.<br>● Non-Western theories (Ubuntu, Buddhism, etc.). |
| K-2 | Common issues and concepts from AI Ethics |
| | ● Familiarity with applied ethics (such as business ethics, ecology, medical ethics and so on).<br>● Familiarity with ethical frameworks, guidelines, and principles in AI, such as beneficence, non-maleficence, autonomy, justice and explicability (Floridi & Cowls, 2019). |
| K-3 | Companies and business's structure and organisation |
| | ● Wide understanding of the internal structure, processes, systems, and dynamics of companies and businesses operating in the private and public sectors. |
| K-4 | Local organisation (the one advised by the AI Ethicist) |
| | ● Terms of reference. |

| | |
|---|---|
| | • Structure, including departmental, organisational, governance and committee structure.<br>• Decision-making processes or framework.<br>• Range of services.<br>• AI Ethics' resources include how the AI Ethics work is financed and the working relationship between the AI Ethics service and other departments, particularly legal counsel, risk management, and development.<br>• Knowledge of how to locate specific types of information. |
| K-5 | AI Systems |
| | • Wide understanding of AI+ML technology's current state and future directions: Theory of ML (such as causality and ethical algorithms) OR of mathematics on social dynamics, behavioural economics, and game theory<br>• Technical awareness of AI/ML technologies (such as the ability to read code rather than write it).<br>• Good understanding of other advanced digital technologies such as IoT, DLT, and Immersive.<br>• Understanding of Language Models – e.g., LLMs – and multi-modal models.<br>• Understanding of global markets and the impact of AI worldwide.<br>• Familiarity with statistical measures of fairness and their relationship with socio-technical concerns. |
| K-6 | Employer's policies |
| | • Informed consent. |
| K-7 | Beliefs and perspectives of the stakeholders |
| | • Understanding of societal and cultural contexts and values.<br>• Familiarity with stakeholders' needs, values, and priorities.<br>• Familiarity with stakeholders' important beliefs and perspectives.<br>• Resource persons for understanding and interpreting cultural communities. |
| K-8 | Relevant codes of ethics, professional conduct, and best practices |
| | • Existing codes of ethics and policies from relevant professional organisations (e.g. game developers, software developers, and so on), if any. |

| | |
|---|---|
| | • Employer's code of professional conduct (if available).<br>• Industry best practices in data management, privacy, and security. |
| K-9 | Relevant AI and Data Laws |
| | • Data protection laws such as GDPR, The Data Protection Act and so on.<br>• Privacy standards.<br>• Relevant domestic and global regulation and policy developments such as ISO 31000 on risk.<br>• AI standards, regulations, and guidelines from all over the world.<br>• Policy-making process (e.g., EU laws governance and enforcement). |
| K-10 | Pedagogy |
| | • Familiarity with learning theories.<br>• Familiarity with various teaching methods. |

*Table 5 - Area and subareas of knowledge for AI Ethicists*

## 5. Conclusion: AI Ethics as a profession. What's next?

While knowledge, skills, and tasks are the core of any profession, other elements are crucial for its establishment and recognition in the market.

As previously mentioned, credibility and independence are essential for the growth and development of the profession of the AI ethicist. In this regard, avoiding conflicts of interest is crucial. Several critical voices have highlighted how AI ethics is already vulnerable to industry capture (Saltelli et al., 2022). This risk is particularly acute in practical and professional applications. To protect AI ethicists, organisations should not reduce their role to mere virtue signalling or ethical washing. AI ethicists must be free to express critical viewpoints and not be constrained to defend their employer's interests.

When AI ethicists are free from conflicts of interest, other factors could further strengthen their position. While organisations ultimately make value-driven decisions, AI ethicists might have some limited liability for errors or negligence in factual or normative assessments. However, given their exposure to such civil liability risks, AI ethicists, particularly those in consulting or advisory roles, might also benefit from some form of protection. For instance, a company could sue an AI Ethicist if it believes it has suffered financial or other damages due to consulting or advisory services. In this context, forms of civil liability insurance could help alleviate this type of pressure.

Finally, developing professional associations and certifications can foster the formation of a professional identity and establish quality criteria that enhance the credibility of these professionals in the market. Successful examples exist in fields closely related to AI ethics. For instance, associations like the International Association of Privacy Professionals (IAPP) have played an essential role in the emergence and establishment of the Data Protection Officer. Similar developments could strongly incentivise the growth of the AI ethicist profession.

While there has been a wealth of discussion on principles for trustworthy and responsible AI in both scholarly and policy circles, along with numerous suggestions for governance structures and processes to achieve these objectives, the focus on the actual capabilities and expertise of individuals tasked with navigating AI ethics in professional environments has been notably lacking. This oversight leads to a scenario where, paraphrasing Kant, principles and processes without skilled professionals are empty, while skilled professionals without those principles and processes are blind. In this article, we address this discrepancy. We showed that AI Ethicists should be experts or, more precisely, moral experts, drawing on a parallel with the professional figure of the Health Care Ethics Consultant in bioethics. We then argued that they should be not only facilitators, like the latter, but also researchers and educators. Finally, we outlined the skills and knowledge that together are necessary to answer questions in the domain of AI Ethics correctly and should underpin the expertise of AI Ethicists.

**Acknowledgements**

MZC's work was supported by the grant 2019/34/E/HS1/00044, 'Epistemic inaccuracy and foundational issues in formal epistemology', financed by the National Science Centre (Poland) and the University of Bologna.

**Declaration**

On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Bibliography**


Algorithm Watch. (2020). Automating Society Report 2020. *Automating Society Report 2020*. https://automatingsociety.algorithmwatch.org

ASBH Task Force. (2011). *Core Competencies for Healthcare Ethics Consultation* (Second edition). American Society for Bioethics and Humanities.

Brummett, A. (2023). Burying the Basilisk of Bioethics: What Can Be Resolved, Dissolved, and Refocused in the Ethics Expertise Debate. *Bioethics*, *37*(6), 515–522.

Brummett, A., & Salter, E. K. (2019). Taxonomizing Views of Clinical Ethics Expertise. *The American Journal of Bioethics*, *19*(11), 50–61.

Burch, R. W. & The Hegeler Institute. (1974). Are There Moral Experts?: *Monist*, *58*(4), 646–658.

Chamberlain, J. (2023). The Risk-Based Approach of the European Union's Proposed Artificial Intelligence Regulation: Some Comments from a Tort Law Perspective. *European Journal of Risk Regulation*, *14*(1), 1–13. https://doi.org/10.1017/err.2022.38

Dastin, J. (2022). Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women *. In *Ethics of Data and Analytics*. Auerbach Publications.

Floridi, L. (2011). *Semantic Information and the Correctness Theory of Truth*. *74*, 147–175.

Floridi, L. (2013). *The Ethics of Information*. Oxford University Press.

Floridi, L. (2018). Soft ethics, the governance of the digital and the General Data Protection Regulation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133), 20180081. https://doi.org/10.1098/rsta.2018.0081

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

Hagendorff, T. (2023). AI ethics and its pitfalls: Not living up to its own standards? *AI and Ethics*, *3*(1), 329–336. https://doi.org/10.1007/s43681-022-00173-5

Hope, T., & McMillan, J. (2012). Physicians' duties and the non-identity problem. *The American Journal of Bioethics: AJOB*, *12*(8), 21–29.

Hulsey, T. L., & Hampson, P. J. (2014). Moral expertise. *New Ideas in Psychology*, *34*, 1–11.

ISO. (2018). *ISO 31000. Risk management—Guidelines*. https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en

ISO/IEC. (forthcoming). *ISO/IEC DIS 23894—Information technology- Artificial intelligence—Guidance on risk management*. https://www.iso.org/standard/77304.html

Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mökander, J., & Floridi, L. (2021). Ethics as a Service: A Pragmatic Operationalisation of AI Ethics. *Minds and Machines*, *31*(2), 239–256. https://doi.org/10.1007/s11023-021-09563-w

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., & Floridi, L. (2023). Operationalising AI ethics: Barriers, enablers and next steps. *AI & SOCIETY*, *38*(1), 411–423. https://doi.org/10.1007/s00146-021-01308-8

Novelli, C., Casolari, F., Rotolo, A., Taddeo, M., & Floridi, L. (2024). AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act. *Digital Society*, *3*(1), 13. https://doi.org/10.1007/s44206-024-00095-1

Richardson, H. S. (2003). *Moral Reasoning*. https://plato.stanford.edu/Entries/reasoning-moral/

Rismani, S., & Moon, Aj. (2023a). *What does it mean to be a responsible AI practitioner: An ontology of roles and skills* (arXiv:2205.03946). arXiv. https://doi.org/10.48550/arXiv.2205.03946

Rismani, S., & Moon, Aj. (2023b). What does it mean to be a responsible AI practitioner: An ontology of roles and skills. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 584–595.

Sætra, H. S., Coeckelbergh, M., & Danaher, J. (2022a). The AI ethicist's dilemma: Fighting Big Tech by supporting Big Tech. *AI and Ethics*, *2*(1), 15–27. https://doi.org/10.1007/s43681-021-00123-7

Sætra, H. S., Coeckelbergh, M., & Danaher, J. (2022b). The AI Ethicist's Dirty Hands Problem. *Communications of the ACM*, *66*(1), 39–41. Scopus. https://doi.org/10.1145/3529751

Saltelli, A., Dankel, D. J., Di Fiore, M., Holland, N., & Pigeon, M. (2022). Science, the endless frontier of regulatory capture. *Futures*, *135*, 102860. https://doi.org/10.1016/j.futures.2021.102860

Schiff, D., Rakova, B., Ayesh, A., Fanti, A., & Lennon, M. (2021). Explaining the Principles to Practices Gap in AI. *IEEE Technology and Society Magazine*, *40*(2), 81–94. https://doi.org/10.1109/MTS.2021.3056286

Tabassi, E. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1; p. NIST AI 100-1). National Institute of Standards and Technology (U.S.). https://doi.org/10.6028/NIST.AI.100-1

Tartaro, A., Panai, E., & Cocchiaro, M. Z. (2024). AI risk assessment using ethical dimensions. *AI Ethics*, *4*, 105–112. https://doi.org/10.1007/s43681-023-00401-6

Watson, J. C., & Guidry-Grimes. (2020, January 1). *Moral Expertise in Clinical Ethics*.

Watson, J. C., & Guidry-Grimes, L. K. (2018). Introduction. In J. C. Watson & L. K. Guidry-Grimes (Eds.), *Moral Expertise: New Essays from Theoretical and Clinical Bioethics* (pp. 1–33). Springer International Publishing. https://doi.org/10.1007/978-3-319-92759-6_1

Wilson, R. (2019). Cambridge Analytica, Facebook, and Influence Operations: A Case Study and Anticipatory Ethical Analysis - ProQuest. *Cambridge Analytica, Facebook, and Influence Operations: A Case Study and Anticipatory Ethical Analysis*. European Conference on Cyber

Warfare and Security.

https://www.proquest.com/openview/b2de0d9e0b088a0711696220cfa58d82/1?pq