How SMEs Ought to Operationalize AI Risk Assessments Under the AI Act

Fabienne Ufert [1] and Zachary J. Goldberg [2]

[1] *Trilateral Research Ltd, Marine Point, 2nd Floor, Belview Port, Waterford, X91 W0XW, Ireland*
[2] *Trilateral Research Ltd, Marine Point, 2nd Floor, Belview Port, Waterford, X91 W0XW, Ireland*

Abstract

Although the current draft of the AI Act contains explicit provisions and general exclusions that create significant innovation opportunities for SMEs, these companies will still need to comply with most provisions of the Act. After surveying these opportunities, this position paper recommends concrete steps how SMEs can overcome the challenges associated with the operationalization of the AI Act through work processes that evidence compliance and aim to create responsible and trustworthy AI. These steps include risk and impact assessments methods and best practice approaches to embed them in company procedures.

Keywords

AI Act, SME compliance, innovation-friendly, risk assessment, impact assessment, operationalization of AI Act

1. Introduction

Already at the time of adoption in 2021, the AI Act proposal triggered concerns of "overburdening" companies with excessive regulation among various actors [1]. Regulators voiced the desire to learn from the "mistakes of the GDPR" and aim for a very "innovation-friendly" AI regulation instead. Especially SMEs with moderate capacities might struggle to understand and consistently follow growing regulatory requirements. First, this short position paper sheds light on the provisions of the AI Act that are particularly innovation friendly for SMEs, followed by an overview of general AI Act exceptions that create innovation opportunities for SMEs. Nonetheless, SMEs will be required to create a compliance governance structure to comply with most of the AI Act's provisions. Therefore, Section 4 explains the value of, as well as recommends, concrete work processes for SMEs to set up, or integrate into their already existing processes, such as a compliance governance structure. These recommendations are based on best practice methods in place at a selected SME based in UK and Ireland that creates ethical AI products and services. The main objective of this guidance is to help SMEs to operationalize the AI Act and use it to their advantage.

2. How innovation-friendly is the AI Act for SMEs?

As a first step to operationalizing the AI Act, SMEs should understand how the Act addresses them specifically. The following paragraphs shed light on the aspects of the AI Act Council General Approach of 6 December 2022 [2], that aim at making the Act less burdensome and innovation-friendly for SMEs. The more recent amendments to the AI Act proposed by the European Parliament (EP) in May 2023 [3] are also taken into consideration in this paper, should they fundamentally change the contents of the provisions of the AI Act of 6 December 2022 that are subsequently mentioned.

HHAI-WS 2023: Workshops at the Second International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 26-27, 2023, Munich, Germany
EMAIL: fabienne.ufert@trilateralresearch.com (A. 1); zachary.goldberg@trilateralresearch.com (A. 2)
ORCID: 0000-0002-5323-9141 (A. 1); 0000-0001-9369-1098 (A. 2)


In comparison to the initial proposal, the current version of the AI Act now includes obligations for general-purpose AI systems (e.g., ChatGPT) which may be used as high-risk AI systems or as components of high-risk AI systems [2, Art. 4b]. However, micro, small and medium-sized enterprises are exempted from the requirements and obligations for such general-purpose AI systems [2, Art. 55a(3)]. SMEs would enjoy nearly unchecked opportunities when it comes to the development and deployment of general-purpose AI systems. Furthermore, SMEs are subject to less strict requirements regarding the technical documentation of high-risk AI systems. While other organizations that do not fall within the SME category have to comply, at a minimum, with the elements set out in Annex IV of the AI Act for their technical documentation, SMEs may submit documentation meeting the same objectives [2, Art. 11(1)] and have thus a slightly lesser administrative burden in terms of the format of documentation.

Title V of the AI Act specifically focuses on 'Measures in Support of Innovation', including AI regulatory sandboxes [2, Art. 53] and testing of high-risk AI systems in real world conditions outside AI regulatory sandboxes [2, Art. 54a]. The industry and regulators alike have identified the sandboxes as crucial for innovation and have suggested that there should be reduced or even no liability for solutions put through the sandboxes at the Member State level, thus resulting in a more lenient approach overall [4]. This suggestion is aimed in particular to assist SMEs and start-ups, who are considered to be the main innovators in this field. The main value of regulatory sandboxes lies with the possibility to test new ideas that are at a breakthrough but still face regulatory uncertainty, or test compliance with already existing rules and mitigate outstanding risks before market uptake. Serving as a good example, the British Information Commissioner's Office (ICO) adopts a multi-agency approach for its Regulatory Sandbox where regulators and experts work with industry through providing clear and quick guidance and bridging potential trust gaps between regulators and providers of AI tools [5]. However, regulatory sandboxes also come with a risk of being misused or abused and thus need the appropriate legal framework to succeed [6]. The AI Act aims at providing such a framework, including eligibility and selection criteria for participation in the AI regulatory sandbox, procedures for the application, participation, monitoring and exiting from and terminations of the sandbox, as well as the terms and conditions applicable to the participants [2, Art. 53]. The specifics will be determined with or after the entry into force of the AI Act.

Title V also includes special support measures for operators, in particular SMEs, including start-ups [2, Art. 55]. These measures refer to priority access to the AI regulatory sandboxes for SMEs, dedicated communication channels between SMEs and Member States, as well as specific awareness raising and training about the application of the AI Act, tailored to the needs of SMEs, provided at the Member State level and/or by the European Commission. Under the AI Act, the European Artificial Intelligence Board (AI Board) will be tasked to establish a sub-group composed of various stakeholders, including SMEs, to advise the AI Board on issues related to the Act's implementation [2, Art. 56(3)]. In addition to the AI Act's innovation-friendly provisions that are directly beneficial to SMEs, the Act also includes a few general exceptions that are potentially innovation-relevant and beneficial for SMEs.

3. What other innovation-relevant AI Act exceptions are potentially beneficial for SMEs?

While the AI Act is not supposed to become a regulatory monster but, instead, be innovation friendly as the examples in the previous section demonstrate, the AI Act provides for a few other prominent exceptions that could further become relevant for AI innovation. Research-focused and/or tech-developing SMEs should be aware of these exceptions that will not be regulated, or only in a limited manner, as these exempted AI systems can be immediately excluded from the risk assessment obligations under the Act. As such, a lesser burden and more freedom with regard to their development will be attached to them – within the bounds of the AI Act and without prejudice to other existing legislation that may fill these gaps.

First, in its current state, the AI Act will not apply to AI systems, including their output, that are specifically developed and put into service for the sole purpose of scientific research and development, as well as any research and development activity regarding AI systems [2, Art. 2(6)-(7)]. Secondly, although a very niche and controlled sector, the AI Act will not apply to AI systems for activities

concerning military, defense or national security [2, Art. 2(3)]. Thirdly, general-purpose AI systems where the provider has explicitly excluded all high-risk uses in the instructions of use or information accompanying the general-purpose AI system will be exempted from any risk assessment under the AI Act [2, Art. 4c]. Finally, some AI systems for law enforcement purposes are no longer included in the list of high-risk AI systems in the current version of the Act and would thus be exempted from the risk assessment requirements for high-risk AI systems. The concerned systems include, for example, AI systems for the detection of deepfakes and AI systems for crime analytics regarding natural persons, the latter allowing law enforcement agencies to search complex related and unrelated large data sets available in different data sources or in different data formats to identify unknown patterns or discover hidden relationships.

Having pointed out the aspects of the AI Act that make it more innovation friendly or at least speak to less burdensome innovation potential for SMEs, SMEs will still need to comply with most requirements of the Act. The following section talks about how.

4. SME compliance under the AI Act

Having identified the special conditions and innovation support that SMEs will enjoy under the AI Act, as well as exemptions under the Act that may be more easily exploited for innovation, SMEs will still need to obey the usual compliance requirements under the AI Act most of the time. Already, SMEs should think about how to operationalize the AI risk assessment procedure that will be required by the Act. These considerations include: how to embed the risk assessment procedure in organizational processes and structures or outsource them, how to conduct efficient checks to understand whether the SME benefits from exemption or less burdensome procedures regarding their line of business and the innovation opportunities that come with them, and even how best to collaborate and potentially influence Member State authorities and the forthcoming AI Board when it comes to issuing harmonized guidance on the implementation of, and compliance with, the AI Act.

Regarding the latter, SMEs' involvement in the AI Board's sub-group mentioned in Section 2 above seems to be a goal-oriented opportunity. Prior to the adoption of the AI Act and the subsequent establishment of the AI Board, SMEs can forego the Board and issue recommendations on matters related to the implementation of the AI Act and, especially, the operationalization of risk assessments under the compliance procedures of the Act. The creation of internally developed conformity assessment methodologies and, for example, corresponding templates for self-assessment, will help SMEs to position themselves early and firmly with their AI products as soon as the Act enters into force.

4.1. Operationalizing the AI Act: Work processes and evidence of compliance

One recommended approach to operationalizing the AI Act is the involvement in standards development (e.g., with a company standards expert or team) in response to the European Commission's standards request to the European standards organizations in support of safe and trustworthy AI. This request involves the standards bodies CEN, CENELEC and ETSI and asks them to create standards around requirements for design and development of high-risk AI systems, AI provider's quality management systems, conformity assessments and auditing of AI systems, robustness specifications or where appropriate, to adopt standards in these areas developed by ISO and IEC.

In addition to following and adhering to standards, a recommended approach to operationalizing the principles of the AI Act includes conducting, and embedding the processes corresponding to, an impact assessment. An impact assessment helps the company to clearly define its ethical framework and ensures compliance with current and future regulations like the AI Act. An impact assessment must be an ongoing and iterative process because ethical concerns arise and require attention at different stages of the AI lifecycle: before development, during development and after deployment [7]. An impact assessment includes a focus on the risks that a particular technology poses to individual and societal wellbeing, individual and group rights, and to a sustainable environment. However, the objectives of impact assessments can go beyond the identification and mitigation of risk [8]. Whereas risk assessments tend to be motivated by the desire to reduce the possibility of harms or wrongs occurring,

the aim of impact assessments is to evaluate the full range of impact that a technology could have. This wider perspective can include benefit to individuals, society or the environment as well as the protection and promotion of the fundamental rights of individuals and groups. The AI Act itself adopts a risk-based approach and requires the establishment of a risk management system [2, Art. 9] but, for the described reasons, a broader impact assessment can be a source of competitive advantage for private companies and is thus a more desirable approach for SMEs.

The literature focused on justifying the use of impact assessments before a new technology is developed or deployed is well-documented [9]. As a result, it is not the intention of the authors to repeat these justifications at present. The purpose of this section is to provide a blueprint for operationalizing in an SME the processes corresponding to an impact assessment in order to facilitate compliance with the principles of the AI Act, particularly the requirements for high-risk AI systems in Title III, Chapter 2 of the Act. With a particular focus on human rights, these requirements include: data protection, documentation and traceability, provision of information, transparency and explainability, diversity, non-discrimination and fairness, human oversight, technical robustness, accuracy, societal and environmental well-being and accountability. The following processes have been embedded in product design at a selected SME based in UK and Ireland that creates ethical AI products.

Before operationalizing ethical principles, an SME or product team might be tempted to start with ethical theories and then attempt to apply them to particular AI tools or their corresponding use cases [10]. This approach seems logical since ethical theories explain the nature and justification of particular ethical values and principles. When referring to ethical theories, most ethicists (at least those in the West) will refer to consequentialism, deontology, or virtue theory.

Consequentialism states that the right thing to do is whatever brings about the best consequences overall [11]. This approach clearly aligns with the importance of risk assessments since risks are potential consequences of our decisions or actions. Deontology states that an act is right or wrong independently of its consequences. Instead, an act is right or wrong if it adheres to a set of rules or principles. Traditionally, these rules or principles are grounded in the imperative to respect a person's dignity [12]. Philosophers have drawn a straight line from deontological theories to rights-based theories [13], whereby we see the direct relevance of deontology to certain rights-based principles in the AI Act and other EU policies, such as the GDPR. Virtue ethics states that the right thing to do is that which the virtuous person does. One ought to ask oneself, what would the generous or courageous person do, and attempt to follow this course of action. Although this theory might appear less directly applicable to the development of new technologies, scholars have shown the relevance of a virtue theory to understanding the ethical issues embedded in the development and use of new and emerging technologies [14].

Despite many attempts to apply these theories to operational activities to the development of new technologies, including AI tools [15], it is not necessary to do so, and can in fact lead to confusion among product developers [16]. Building an AI tool requires thinking about consequences and rights, and yet, the corresponding ethical theories are incompatible with one another. Asking a development team to pause to engage in discussions about ethical theory (as interesting as these discussions may be) can be an unhelpful distraction. Moreover, these theories help explain why one might consider certain actions to be right and wrong, but they do not always provide clear decision-making procedures. They provide a general principle, which one ought to follow. However, applying these general principles to the nuanced processes involved in developing and using AI is not straightforward. To address the complex ethical issues concerning AI development and use, and to operationalize concrete steps that result in compliance with the AI Act, more practical processes can be adopted.

4.2. Recommendations

As stated above, the following recommendations for SMEs to operationalize the AI Act are derived in large part by processes embedded at a selected SME based in the UK and Ireland. Each recommendation reflects concrete steps that SMEs can take. Simultaneously, the recommendations are interrelated, meaning that pursuing and achieving a single recommendation may require pursuing and achieving the others. Taken together, these recommendations provide the means for compliance with the AI Act, but also reach beyond compliance to constitute a set of best practices for SMEs to adopt.

- **Accountability**: Owing to both the less burdensome regulations facing SMEs as well as their opportunities to influence and issue recommendations on matters related to the implementation of the AI Act (both of which were detailed in Sections 2 and 3), accountability is paramount for SMEs. Generally speaking, "accountability" refers to an ethical expectation that individuals or organisations take ownership of their actions or conduct, and they explain reasons for which decisions and actions were taken. When mistakes or errors are made, it also implies taking action to ensure a better outcome in the future. To achieve these ends, SMEs ought to establish mechanisms and procedures to ensure responsibility and accountability for AI systems during development and in use. Putting clear lines of accountability in place helps promote transparency and trust as both clients and the public can feel confident that AI products and services are planned and monitored, that someone is responsible for the ethical development of the tools, and that one can request an explanation for why the organization made particular development choices. In these ways, accountability helps SMEs achieve many other ethical principles mandated in the AI Act including transparency, explainability and redress.
- **Organizational support**: To help establish lines of accountability and to further institutionally embed the processes of an impact assessment into product development, SMEs ought to have organizational support for AI compliance and responsible AI from the C-suite on down. Company executives ought to embrace the steps needed to operationalize the AI Act including implementing ethics and ethics discussions into regular meetings. For example, organizing a weekly "ethics meeting" where employees can discuss ethical issues, stories taken from the news about ethical and unethical AI, or concrete issues they are confronting in the development of a new product helps place the AI Act's ethical principles at the forefront of an SME's approach to product development. Furthermore, employees ought to have a clear understanding of their responsibilities and the shared responsibilities of clients during co-design or after products have been deployed. It is integral to establish feedback mechanisms so that any employee can communicate their opinions and ask questions. Additionally, empowering employees to make decisions and be accountable to them can facilitate a sense of ownership in the technical solutions that they are building.
- **Company values/pledges**: Defining company values, and making a pledge to adhere to them, can have a real and significant effect on employee uptake of those values [17]. Adopting responsible and trustworthy AI, as it is defined by the EU AI-HLEG and embodied in the AI Act principles, as company values communicates both to clients and to employees their importance to the SME. It is a simple, but effective, way of facilitating understanding among employees of the company's objectives and motivating them to help advance them. When seeking guidance concerning how to make decisions to achieve company objectives, employees can refer to these company values and the company's pledge to adhere to them. By internalizing responsible and trustworthy AI, the corresponding values become goals rather than compliance-related hindrances to innovation.
- **Sociotech Approach**: AI devices influence and impact our lives in numerous ways. From receiving automated recommendations from movie or music streaming services and unlocking mobile phones with facial images or fingerprints, to diagnosing cancer as well as decision making and information gathering in the areas such as employment, bank loans, marking student's tests, policing and safeguarding, defense and security, AI devices are pervasive in human society and individual lives.

  These technologies are not value neutral. For example, when a company decides to design, develop and sell a product because it is more efficient, it has valued efficiency. Or if a country installs technology at its border to improve security and accuracy of identity verification, then it has valued security and accuracy. Furthermore, because, ethical values can conflict with one another, when an individual or organization chooses to pursue a particular value, it often chooses to pursue it at the cost of another value. In a streaming platform people might choose (or allow the choice of) entertainment over privacy, or in the field of border security, security over inclusivity [18]. Furthermore, AI devices do not always function individually, or in silos. They often interact with one another such as in smart home systems, vehicular communication systems and "the "internet of military things".

This interaction creates a socio-technological complexity. Technology impacts society and social relations, and society and social relations influence which technologies are developed, marketed and used [19]. This intermingling of the social and the technological requires SMEs to adopt a sociotechnical approach to understanding their AI products. This approach requires data scientists, ethics experts, philosophers, lawyers, and subject matter experts to be incorporated into product design and development. Including these individuals in product design and development facilitates a robust and efficient due diligence process.

- **Structure/Content Distinction**: Companies may not know precisely how to begin adopting and operationalizing its principles. A crucial distinction for understanding and operationalizing responsible and trustworthy AI as its described in the AI Act, is structure versus content [20]. Policies, procedures, and role-determined responsibilities that identify and mitigate AI ethical risks in the design, development, and procurement of AI constitute the structure a company has established to comply with the AI Act or otherwise promote responsible and trustworthy AI. These processes and procedures are the formal components that have been endorsed by the company's C-suite, reflect the company's values, and provide practical means for achieving AI Act compliance.
  On the other hand, the ethical risks or benefits that a company aims to address are the content of its approach. For example, prioritizing the identification and minimization of bias in an AI model reflects the content of a company's responsible AI program.
  Recognizing this distinction between structure and content is a useful method for organizations to adopt to be able to advance the operationalization of the AI Act. Establishing an explicit distinction between *how* a company approaches ethical issues and *what* those issues are, can help demystify the AI Act and provide clear steps to operationalize its principles.
- **Content workshops**: After the structure has been established, validated, and endorsed internally, the product development team—which includes ethics experts, lawyers, and subject matter experts in addition to data scientists—ought to hold regular internal workshops corresponding to the different ethical principles listed in the AI Act. For example, bias. While the overall aim of the principles in the AI Act is to protect society from the risks of AI, its mandatory requirements speak loudly to safeguarding against bias; organizations must clearly evidence the steps that they have taken to safeguard against bias and other associated risks at each stage of the AI lifecycle. Bias occurs where results or outputs from an AI system are disproportionately and unfairly in favor of or against an idea, group or person. Bias can occur at any point in the AI life cycle. To identify and mitigate bias in an AI tool, product development teams ought to map out all of the potential kinds of bias and how they can arise from the planning and assessment phase to data collection and processing, model training and validation, and deployment of the tool. By holding content workshops for each of the AI Act ethical principles—transparency, explainability, human oversight, etc.—organizations embed compliance into their structure.
- **AI ethics expertise**: The AI Act is a complex piece of legislation and understanding ethics, conflicts among values, and how to evaluate value trade-offs requires specialization in these areas. Just as the arrival of the GDPR resulted in company's hiring DPOs or outsourcing their GDPR compliance work to certified professionals, companies are advised to seek comparable expertise to ensure compliance with the AI Act. If company resources permit, they could either hire an internal AI ethics and/or compliance officer, form an AI ethics board or outsource their compliance needs to SMEs providing consultancy and client services in the AI assurance ecosystem. These experts can establish the structure, identify the content, and carry out the risk assessments needed for AI Act compliance.

4.3. Frameworks and Best Practices

In addition to the recommendations made in the preceding subsection, SMEs may require additional guidance to operationalize the AI Act's ethical principles. Importantly, these principles must be addressed throughout the AI lifecycle as risks to their fulfilment can arise repeatedly and in different ways for the duration of product development.

AI Lifecycle Phases	Examples
Assessment and planning	• AI impact assessments [9] – a diverse field of methods for identifying positive and negative impacts to safeguard benefits and avoid downsides. Can be based on data protection, fundamental rights, ethics, or other perspectives, and shall be conducted throughout the AI lifecycle. • Funding constraints – The EU's Horizon Europe innovation research funding programme makes technical robustness an evaluation criterion for AI projects.
Design	• Value-sensitive design, privacy-by-design, Ethics-by-Design – various methods for including principles into design processes (IEEE ethical-aligned design [21]) • Methods for raising awareness of ethical issues for designers, e.g., ODI data ethics canvas [22] • Participatory design approaches (e.g., NESTA Participatory AI for humanitarian innovation [23]) that bring in principles via involving impacted stakeholders
Development and procurement	• UK government, Office of AI, Guidelines for AI procurement [24]– a summary of best practices addressing specific challenges of acquiring Artificial Intelligence technologies in government. • World Economic Forum, AI Procurement in a Box [25] • Data sheets for data sets [26]– a process for documenting datasets used for ML tools intended for high-risk environments. • Security guidance (e.g., German BSI Security of AI systems [27]) • Test infrastructure (e.g., NIST Facial recognition vendor test [28])
Deployment, Monitoring and Control	• AI Audits, either bespoke Trustworthiness auditing or general audits that refer to principles. E.g., ICO AI audit [29]. • UK government Central Digital and Data Office and Centre for Data Ethics and Innovation's Algorithmic transparency reporting standard [30]. • AI Incident databases – databases [31] aiming to collect incidents of harm or near harm from AI systems to help researchers and developers avoid repeated unwanted outcomes.
Decommissioning/ Retirement	• There is limited stand-alone guidance for this phase, which is sometimes part of full-lifecycle guidance. • ICO guidance on personal data storage limitation and retention policies [32]

Table 1: Frameworks and tools to operationalize the AI Act's ethical principles

SMEs may encounter challenges implementing risk assessment practices and operationalizing ethical principles in product development. Nevertheless, these tools and frameworks can provide the means and the practical steps to achieve these goals. Given the complexity involved with understanding and applying ethical principles and risk assessment practices, the application of these frameworks ought to be combined with the ethics expertise described in the preceding subsection.

5. Conclusion

To avoid overburdening SMEs with excessive regulation which could hinder innovation, the AI Act contains explicit provisions as well as purposeful exclusions intended to foster innovation among these companies. Although these provisions and exclusions create significant innovation opportunities for SMEs, they are still required to create a compliance governance structure to comply with most of the requirements of the Act. Moreover, given the Act's comparative leniency towards SMEs, these

companies confront a heightened need for self-accountability to ensure their processes and products produce responsible and trustworthy AI. To align with the AI Act's risk-based approach, and to operationalize the Act's principles and provisions, SME's will need to conduct risk assessments at minimum, and impact assessments as best practice and to identify ethical benefits and commercial opportunities. However, putting these assessments into action, embedding their objectives into company procedures, and successfully aligning their methods with product development are significant challenges. Accordingly, this paper has described a preliminary position recommending concrete steps for SMEs to operationalize the AI Act.

6. Acknowledgements

This Word template was created by Aleksandr Ometov, TAU, Finland. The template is made available under a Creative Commons License Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

7. References

[1] O. Noyan, EU Parliament, countries want mor innovation, less burden in AI Act, 2021. URL: https://www.euractiv.com/section/digital/news/eu-parliament-countries-want-more-innovation-less-burden-in-ai-act/.

[2] Council of the European Union, Interinstitutional File: 2021/0106(COD), 2022. URL: https://data.consilium.europa.eu/doc/document/ST-15698-2022-INIT/EN/pdf.

[3] European Parliament Committee on the Internal Market and Consumer Protection and Committee on Civil Liberties, Justice and Home Affairs, COM(2021)0206 – C9 0146.2021 – 2021/0106(COD), 2023. URL: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf.

[4] SIDLEY, Proposal for EU Artificial Intelligence Act Passes Next Level – Where Do We Stand and What's Next?, 2022. URL: https://www.sidley.com/en/insights/newsupdates/2022/12/proposal-for-eu-artificial-intelligence-act-passes-next-level.

[5] R. Morrison, Government backs UK AI regulatory sandbox. URL: https://techmonitor.ai/technology/ai-and-automation/government-backs-ai-regulatory-sandbox.

[6] S. Ranchordas. "Experimental Regulations for AI: Sandboxes for Morals and More." University of Groningen Faculty of Law Research Paper No. 7. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3839744.

[7] W. Reijers, D. Wright, P. Brey, K. Weber, R. Rodrigues, D. O'Sullivan, & B. Gordijn. "Methods for Practising Ethics in Research and Innovation: A Literature Review, Critical Analysis and Recommendations." Science and Engineering Ethics, 24.5 (2018): 1437-1481.

[8] R. Koivisto, D. Douglas. "Principles and Approaches in Ethics Assessment." Satori Project, 2015. URL: https://satoriproject.eu/media/1.h-Ethics-and-Risk1.pdf.

[9] B.C. Stahl, J. Antoniou, N. Bhalla, et al. "A systematic review of artificial intelligence impact assessments." Artif Intell Rev (2023). URL: https://doi.org/10.1007/s10462-023-10420-8.

[10] Z. Goldberg. How to Conduct an Ethics Assessment of AI in Policing, Proceedings, 14th ACM Web Science Conference June 2022, pp. 466-470. Doi: https://doi.org/10.1145/3501247.3539509.

[11] J.S. Mill, Utilitarianism. Parker, son, and Bourn, London, 1863.

[12] I. Kant. Grounding for the Metaphysics of Morals. Hackett, Cambridge, MA, 1993.

[13] L. Alexander, M. Moore, Deontological Ethics, in E.N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, Winter 2021 ed., 2021, URL: https://plato.stanford.edu/archives/win2021/entries/ethics-deontological/.

[14] S. Vallor. Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting. Oxford University Press, Oxford, 2016.

[15] J. Donia, J.A. Shaw. "Ethics and Values in Design: A Structured Review and Theoretical Critique." Sci Eng Ethics 27.57 (2021). https://doi.org/10.1007/s11948-021-00329-2.

[16] R. Blackman. Ethical Machines: Your Concise Guide to Totally Unbiased, Transparent, and Respectful AI. Harvard Business review Press, Cambridge, MA, 2022.
[17] B. Gleeson. Why Core Values Matter (And How To Get Your Team Excited About Them). Forbes. March 30, 2021. URL: https://www.forbes.com/sites/brentgleeson/2021/03/30/why-core-values-matter-and-how-to-get-your-team-excited-about-them/.
[18] Z. Goldberg. What Moral Philosophy Brings to a Technological Society. TedX. Video, 2021. URL: https://www.youtube.com/watch?v=oX-IoIB9QCg.
[19] G. Baxter, I. Sommerville. "Socio-technicalsystems: From design methods to systems engineering." Interacting with computers, 23.1, 2011: 4-17.
[20] R. Blackman. Ethical Machines: Your Concise Guide to Totally Unbiased, Transparent, and Respectful AI. Harvard Business review Press, Cambridge, MA, 2022: 16.
[21] IEEE. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous Intelligent Systems. URL: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf.
[22] Open Data Institute. Data Ethics Canvas. URL: https://www.theodi.org/article/the-data-ethics-canvas-2021/.
[23] NESTA. "Participatory AI for Human Intervention: A Briefing Paper", 2021. URL: https://media.nesta.org.uk/documents/Nesta_Participatory_AI_for_humanitarian_innovation_Final.pdf.
[24] Office for AI. Guidelines for AI Procurement. URL: https://www.gov.uk/government/publications/guidelines-for-ai-procurement.
[25] World Economic Forum. AI Procurement in a Box. URL: https://www.weforum.org/reports/ai-procurement-in-a-box/.
[26] T. Gebru, et al. Data Sheets for Datasets, 2021. URL: https://dl.acm.org/doi/10.1145/3458723.
[27] German BSI. Security of AI Systems. URL: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/KI/Security-of-AI-systems_fundamentals.pdf?__blob=publicationFile&v=4.
[28] NIST. Facial Recognition Vendor Test. URL: https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt.
[29] ICO. A Guide to ICO AI Audits. URL: https://ico.org.uk/media/for-organisations/documents/4022651/a-guide-to-ai-audits.pdf.
[30] CDDO/CDEI. Algorithmic Transparency Recording Standard. URL: https://ico.org.uk/media/for-organisations/documents/4022651/a-guide-to-ai-audits.pdf.
[31] AI Incident Database. URL: https://partnershiponai.org/workstream/ai-incidents-database/.
[32] ICO. Guidance on Data Storage Limitation. URL: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/the-principles/storage-limitation/.