frontiers | Frontiers in Computer Science

Check for updates

Challenges and best practices in corporate AI governance: Lessons from the biopharmaceutical industry

Jakob Mökander[1,2]*†, Margi Sheth[3]*†, Mimmi Gersbro-Sundler[3], Peder Blomgren[3] and Luciano Floridi[1,4]

[1]Oxford Internet Institute, University of Oxford, Oxford, United Kingdom, [2]Center for Information Technology Policy, Princeton University, Princeton, NJ, United States, [3]R&D Data Office, Data Sciences and AI, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom, [4]Department of Legal Studies, University of Bologna, Bologna, Italy

While the use of artificial intelligence (AI) systems promises to bring significant economic and social benefits, it is also coupled with ethical, legal, and technical challenges. Business leaders thus face the question of how to best reap the benefits of automation whilst managing the associated risks. As a first step, many companies have committed themselves to various sets of ethics principles aimed at guiding the design and use of AI systems. So far so good. But how can well-intentioned ethical principles be translated into effective practice? And what challenges await companies that attempt to operationalize AI governance? In this article, we address these questions by drawing on our first-hand experience of shaping and driving the roll-out of AI governance within AstraZeneca, a biopharmaceutical company. The examples we discuss highlight challenges that any organization attempting to operationalize AI governance will have to face. These include questions concerning how to define the material scope of AI governance, how to harmonize standards across decentralized organizations, and how to measure the impact of specific AI governance initiatives. By showcasing how AstraZeneca managed these operational questions, we hope to provide project managers, CIOs, AI practitioners, and data privacy officers responsible for designing and implementing AI governance frameworks within other organizations with generalizable best practices. In essence, companies seeking to operationalize AI governance are encouraged to build on existing policies and governance structures, use pragmatic and action-oriented terminology, focus on risk management in development and procurement, and empower employees through continuous education and change management.

Introduction—The need for corporate AI governance

There are two main reasons why artificial intelligence (AI) might be giving business leaders around the world sleepless nights. The first is sheer excitement at the possibilities of the technology.[1] The second is abject fear about the implications of getting it wrong.[2] The potential of autonomous and self-learning technologies to revolutionize industries is only just starting to be realized, with far-reaching implications for human development, economic prosperity, and the financial prospects of individual firms (Floridi et al., 2018). Unsurprisingly, companies are scrambling to implement AI-powered solutions, conscious no doubt that their competitors will be doing the same. But in the rush to join the revolution, there is great danger of organizations making missteps that could lead them into legal and ethical minefields, where they might not only suffer reputational damage but also cause real-world harm (Silverman, 2021).

Consider healthcare as an example of both the benefits and the risks associated with AI. Across the industry, AI systems are already saving lives and increasing life quality by aiding medical diagnostics, driving service improvements through better forecasting, and enabling more effective drug discovery processes (Schneider, 2019; Topol, 2019). However, along with excitement and opportunity, the use of AI systems in healthcare is coupled with serious ethical challenges. Such systems may leave users vulnerable to discrimination and privacy violations (Laurie et al., 2015). They can also erode human self-determination and enable wrongdoing (Tsamados et al., 2021).

Recently, public outcry against specific AI use cases (Holweg et al., 2022) and proposals for "hard" legislation in both the EU[3] and the U.S.[4] have stressed the urgency of addressing these challenges. In that context, defining and communicating ethical principles is a crucial first step toward implementing effective AI governance. It is also a step that many organizations in both the private and the public sector have already taken.[5] However, organizations still lack effective ways of translating those abstract principles into concrete actions that will enable them to achieve the benefits of AI in ways that are ethical, legal, and safe (Morley et al., 2020; Gianni et al., 2022).

As a team consisting of both industry practitioners[6] and academic researchers,[7] we set out to fill that gap. Over 12 months, we helped guide and coordinate the roll-out of AI governance within AstraZeneca, a multinational biopharmaceutical company. Throughout the process, we documented the challenges the organization faced and the best practices that were developed to manage the different tensions that arose. Our findings—which are summarized in this article—include insights that are broadly applicable across industries, offering answers to some of the questions that may be troubling the sleepless manager: What challenges await when implementing AI governance? How can well-intentioned ethical principles be translated into effective practice? And what initial investments are required to reap long-term benefits?

Before proceeding, is should be noted that the pharmaceutical industry—from which our case study is drawn—is well-positioned to pioneer the operationalization of AI governance for three reasons. First, the pharmaceutical industry has a long history of dealing with sensitive data. As a result, governance structures already exist to identify and mitigate technology-related risks. Second, pharmaceutical companies have always operated in an environment governed by laws as well as trust. Indeed, many approaches to AI ethics are based on the classical principles of bioethics (Blasimme and Vayena, 2021). Finally, developing new drugs is not only a data-driven but also a resource-intense endeavour. Globally, over \$200 bn is spent on pharmaceutical R&D each year.[8] Hence, there are strong incentives to use AI systems in ways that avoid regulatory red tape.

Case study—AstraZeneca's AI governance journey

As an R&D-driven organization, AstraZeneca's core business is to use science and innovation to improve health outcomes

---

1 This bullish mood was reflected in a much-cited report by PwC (2017), which suggested that organizations will be 40% more efficient by 2035 thanks to their adoption of AI systems. While these numbers are debatable, the use of AI systems can lower costs, increase consistency, and enable novel solutions to complex problems Taddeo and Floridi (2018).

2 A survey by McKinsey & Company (2021) found that business leaders worry that their companies lack the capacity to address the full range of risks posed by AI. The same survey also found that executives are particularly concerned about risks relating to cybersecurity, data privacy, and regulatory compliance.

3 See e.g. The Artificial Intelligence Act (European Commission, 2021).

4 See e.g. The Algorithmic Accountability Act of 2022 (Office of U.S. Senator Ron Wyden., 2022).

5 Institutions like the European Commission (AI HLEG, 2019), the (OECD, 2019) and the (IEEE, 2019) have all published principles with the aim of guiding the design and use AI systems. In parallel, many companies have chosen to develop and publish their own sets of AI ethics principles.

6 As members of AstraZeneca's R&D Data Office, MS, MSG, and PB have been instrumental in shaping, coordinating, and driving the operationalization of AI governance within the organization.

7 LF is a professor of philosophy and ethics of information and a former member of the European Commission's High-Level Expert Group on AI. JM is a PhD candidate at the Oxford Internet Institute. Over a period of two years, from Q3 2020 to Q3 2022, LF and JM observed and documented AstraZeneca's internal activities related to AI governance.

8 www.iqvia.com/insights/the-iqvia-institute/reports/global-trends-in-r-and-d-2022

through more effective treatment and the prevention of complex diseases. In doing so, the company uses AI systems in many different ways. These include biological insight knowledge graphs to improve drug discovery processes, machine-learning powered image recognition software for faster and more accurate medical analysis, and natural language processing models to prioritize adverse event reports (Crowe, 2020; Lea et al., 2021).

All these use cases change how the company collects, analyzes, and utilizes data. And as technologies and ways of working evolve, so must organizational governance. In November 2020, AstraZeneca's board moved toward addressing that need by publishing a set of Principles for Ethical Data and AI.[9] Those principles stipulate that the use of AI systems should be private and secure, explainable and transparent, fair, accountable, human-centric, and socially beneficial.

AstraZeneca's ethics principles aim to help employees and partners navigate the risks associated with AI systems. Yet principles alone cannot ensure that such systems are designed and used ethically (Mittelstadt, 2019), and their implementation is never straightforward (Ryan et al., 2021). Moreover, like many other multinational corporations, AstraZeneca is a decentralized organization. Different business areas were thus allowed to develop their own AI governance structures to reflect variations in objectives, digital maturity, and ways of working.

To support and unify local activities, the company launched four enterprise-wide initiatives:

- The creation of overarching *compliance* and *guidance documents*. The aim thereby was to break down each high-level principle into more tangible and actionable formulations.[10]
- The development of a *Responsible AI playbook*. The purpose of the playbook was to provide detailed, end-to-end guidance on developing, testing, and deploying AI systems within AstraZeneca.[11]
- The establishment of an internal *Responsible AI Consultancy Service,* and an *AI resolution Board*. These new organizational functions were established to (i) facilitate the sharing of best practices, (ii) educate staff, and (iii) monitor the governance of AI projects.

- The commissioning of an *AI audit* conducted in collaboration with an independent party. By subjecting itself to external review, AstraZeneca got valuable feedback on how to improve its existing and emerging AI governance structures.[12]

The above-listed initiatives may appear straightforward. But they only emerged out of extended—and sometimes *ad hoc*—internal processes that came up against both conceptual difficulties and organizational tensions. In the remainder of this article, we highlight the challenges AstraZeneca faced in its efforts to operationalize AI governance and discuss lessons learned, i.e., how these challenges can be managed in pragmatic and constructive ways. The aim thereby is to distill generalizable best practices for how to implement any set of AI ethics principles in practice.

Implementation challenges—What to be prepared for?

Organizations seeking to operationalize AI governance face both conceptual and practical difficulties. Our research and first-hand experience suggest that there are four main challenges:

Balancing interests

The first challenge concerns the tension between risk management and innovation. The use of AI systems in the pharmaceutical industry gives a stark illustration of that issue. Obviously, the industry must put patients' safety first. Often, that means using available technologies to develop new drugs or to diagnose and intervene early in the course of a disease. To ensure that such drugs are safe, AstraZeneca trains AI systems to detect treatment response patterns (Nadler et al., 2021). But red tape related to AI governance could restrict the development of such potentially lifesaving procedures. In such circumstances, how does a company "err on the safe side?" There is no simple, one-size-fits-all answer. Instead, organizations seeking to operationalize AI governance should prioritize defining and controlling the risk appetite of different projects.

Defining "AI"

Second, every policy needs to define its scope, but how can you do this when there is no universally accepted definition of AI?[13] Determining the scope of AI governance is especially difficult because the technology is always embedded in larger

---

sociotechnical systems, in which processes driven by humans and by machines overlap (Lauer, 2020). That is why establishing the scope of AI governance is a balancing act. Make the scope overinclusive, and you create unnecessary administrative burdens. Make it underinclusive, and risks will go under the radar (Danks and London, 2017). In AstraZeneca's case, countless meetings were spent on discussing how to best strike that balance. They key to move beyond such discussion is to realize that there is a three-way trade-off between how precisely you define the scope of your AI governance, how easy it is to apply it, and how generalizable it is.

Harmonizing standards

Third, the same requirements must apply to all AI systems used by an organization. If not, corporate AI governance may simply persuade managers to outsource unethical or risky projects (Floridi, 2019). But the drive to impose uniform requirements creates new tensions. Large organizations often comprise distinct business areas that operate independently. The cycle of designing and training AI systems often involves multiple organizations. For example, AstraZeneca collaborates with BenevolentAI, a British start-up, to identify treatments against chronic kidney disease by using the former's rich datasets to build biological insight knowledge graphs.[14] This is not an exception but the rule: AI systems result from supply chains spanning multiple actors and geographic regions (Crawford, 2021). Harmonizing standards means treating all AI systems equally, regardless of whether they have been developed in-house or procured from third parties.

Measuring results

Fourth, ethics is hard to quantify, and it is not clear how organizations seeking to operationalize AI governance can measure and demonstrate their success. One option is assessing how AI systems operate in terms of fairness, transparency, and accountability. However, it is hard to find ways to quantify and measure these in practice (Kleinberg, 2018). And, as Goodhart's Law reminds us, when a metric becomes a target, it ceases to be a good measure (Strathern, 1997). Alternatively, organizations could focus on designing process-based KPIs to capture the mechanisms in place to mitigate technology-related risks. Yet such checklists tend to reduce AI governance to a box-ticking exercise. Perhaps the solution here comes down to a question of mindset: In an AI governance context, the main purpose of KPIs

is not to assess whether a specific system is "ethical" but rather to spark debates about ethics that inform design choices.

Discussion—Best practices and lessons learned

While the challenges discussed above are real and important, they are not insurmountable. Our experiences from coordinating and observing AstraZeneca's efforts to operationalize AI governance can be condensed into four transferable best practices:

Build on existing policies and governance structures

AI governance is most likely to be effective when integrated into existing governance structures and business processes (Hodges, 2015). Policies that duplicate existing structures may be perceived as unnecessary by the people expected to implement them. Rather than adding steps to their software development processes, organizations can simply update them so that solution requirements align with the objectives of AI governance. This makes is easier for employees to understand what is expected of them and how any new measures related to AI governance impacts their daily tasks. For example, AstraZeneca's software developers and clinical experts found that trying to implement the company's AI ethics principles pushed them to think about their projects in new ways that can help organizations improve their processes and workflows. In our experience, AI governance is most effectively operationalized when such advantages are clearly communicated.

Use pragmatic and action-oriented terminology

It is less important to define what AI is in abstract terms and more important to establish processes for identifying those systems that require additional layers of governance.[15] Rather than struggling to pin down a precise definition of AI, AstraZeneca created a guidance document that describes the characteristics of the systems to which their ethics principles apply. A list of examples does not constitute a definition of AI, but it can nevertheless help employees determine whether a specific use case is in scope. Also, following the

---

13 See Wang (2019) for an excellent overview of different, partly conflicting, definitions of AI.

14 www.benevolent.com/news/astrazeneca-starts-artificial-intelligence-collaboration-to-accelerate-drug-discovery

---

15 Several recent publications have proposed pragmatic ways of classifying AI systems for the purpose of implementing corporate AI governance. See e.g. Aiken (2021), Mökander et al. (2022) and OECD (2022).

European Commission (2021), AstraZeneca adopted a risk-based approach, classifying systems as either low-, medium- or high-risk with proportionate governance requirements attached to each level. Using a familiar concept like "risk assessment" helps organizations integrate AI governance into their existing quality management processes. A risk-based approach is also future-proof because it avoids the trap of committing to a definition that could become obsolete as the technology rapidly develops.

Focus on risk management in development and procurement

Distinguishing between compliance assurance and risk assurance helps to harmonize standards across organizations. While compliance assurance compares organizational procedures to existing laws and regulations, risk assurance asks open-ended questions about how different business areas work to identify and manage risk. Because regulations vary across jurisdictions and sectors (Viljanen and Parviainen, 2022), it is not always practically feasible to audit all parts of a large, multinational organization for compliance against the same substantive standards. In contrast, risk assurance can be adapted locally to reflect how different business areas understand risk. Because they leave space for managers in different regions and business areas to justify their governance design choices, it is both possible and desirable to subject all parts of an organization to harmonized AI risk audits. Again, this does not necessarily require the creation of additional layers of governance. Organizations should simply focus on finding any gaps in their existing development and procurement processes and filling them by adding ethics-based criteria for evaluating AI systems.

Empower employees through continuous education and change management

Because corporate AI governance is about change management, internal communication and training efforts are key. In AstraZeneca's case, these efforts were continuous and happened simultaneously on several different levels. For example, the ethics principles were agreed upon through a bottom-up process that included extensive consultations with employees. An important aspect of this process was anchoring the ethics principles with internal stakeholders. After all, ensuring that AI systems are designed and used legally, ethically, and safely requires organizations not only to have the right tools in place but also to make their employees aware of them and willing to use them.

Admittedly, change management is no easy task and the implementation of AI governance is no exemption: Humans have limited attention spans, and employees are frequently bombarded with information about different governance initiatives (Baldwin and Cave, 1999). That said, our first-hand experiences suggest that much can be done to facilitate a successful implementation of AI governance. To start with, communication concerning AI governance is most effective when supported by senior executives. AI governance is also more likely to be implemented when aligned with incentives for individuals and business areas. Put differently, employees must be enabled and supported to do the right thing. That includes training and education as well as channels through which employees can seek escalate issues without fear of being blamed. Finally, tools such as impact assessments and model testing protocols may be developed by individual teams but should be shared widely to encourage the harmonization of practices and prevent the duplication of efforts.

Concluding remarks—Upfront investments vs. long-term benefits

Our case study of AstraZeneca shows that the most important step toward good corporate AI governance is to ensure procedural regularity and transparency. To do so, organizations do not need to invent or impose new corporate governance structures. For example, while many useful tools such as model cards (Mitchell et al., 2019) and datasheets (Gebru et al., 2018) and methods, like conformity assessments (Floridi et al., 2022), have already been developed, their use is typically neither coordinated nor enforced. That is why the immediate goal of corporate AI governance should be to interlink existing structures, tools, and methods as well as to encourage and inform ethical deliberation through all stages of the AI lifecycle (Mökander and Axente, 2021).

Efforts to operationalize AI governance incur costs, both financial and administrative. To start with, formulating organizational values bottom-up is a time-consuming activity. In AstraZeneca's case, the process of drafting the ethics principles also included multiple consultations with internal executive leaders on strategy and with academic researchers offering external feedback. Since the publication of its ethics principles in 2020, approximately four full-time staff have been working on implementing AI governance across AstraZeneca. Also, in Q4 2021, the company conducted an "AI audit" in collaboration with an independent third party. Added to the costs of procuring that service, AstraZeneca employees invested around 2,000 person-hours in the audit.

To put those numbers into perspective, consider the costs associated with certification and compliance with hard legislation. According to the European Commission, obtaining certification for an AI system in line with the proposed EU

legislation on AI will cost on average EUR 20.000, corresponding to approximately 12% of the development cost (Renda et al., 2021). At the same time, one of the main reasons why technology providers engage with auditors is that it is often cheaper and easier to address system vulnerabilities early in the development process. In addition, good AI governance can help organizations improve several business metrics, including data security, brand management, and talent acquisition (EIU, 2020). This shows that—despite the associated costs—businesses have clear incentives to implement effective corporate AI governance structures.

Our discussion in this article has centered on lessons from the pharmaceutical industry. However, AstraZeneca's situation seems highly representative of the many firms that have recently adopted ethics principles for the design and use of AI systems. That is why the challenges and best practices discussed above should be relevant to any organization seeking to operationalize AI governance. It is vital to remember that such governance will not, and should not, replace the need for the designers, operators, and users of AI systems to continuously reflect on the ethics of their actions. Nevertheless, governance that follows the best practices outlined in this article can help organizations manage the ethical risks posed by AI systems while reaping the economic and social benefits of automation.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

All listed authors have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Conflict of interest

Authors MS, MG-S, and PB are employees of AstraZeneca plc. Author JM is a Ph.D. student at the Oxford Internet Institute. His research was supported by AstraZeneca plc.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

AI HLEG (2019). *European Commission's Ethics Guidelines for Trustworthy Artificial Intelligence* (Issue May). Available online at: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1 (accessed September 5, 2022).

Aiken, C. (2021). *Classifying AI Systems CSET Data Brief.* Available online at: https://cset.georgetown.edu/publication/classifying-ai-systems/ (accessed September 5, 2022).

Baldwin, R., and Cave, M. (1999). *Understanding Regulation : Theory, Strategy, and Practice.* Oxford, UK: Oxford University Press.

Blasimme, A., and Vayena, E. (2021). "The ethics of AI in biomedical research, patient care, and public health [Bookitem]," *The Oxford Handbook of Ethics of AI*, eds in M. Dubber, F. Pasquale, and S. Das. Oxford, UK: Oxford University Press.

Crawford, K. (2021). *The Atlas of AI [Book].* New Haven, CT: Yale University Press.

Crowe, D. (2020). *Modelling Biomedical Data for a Drug Discovery Knowledge Graph. Towards Data Science.* Available online at: https://towardsdatascience.com/modelling-biomedical-data-for-a-drug-discovery-knowledge-graph-a709be653168

Danks, D., and London, A. J. (2017). Algorithmic bias in autonomous systems. *IJCAI Int. Joint Conf. Artif. Intell.* 0, 4691–4697. doi: 10.24963/ijcai.2017/654

EIU (2020). *Staying Ahead of the Curve–The Business Case for Responsible AI.* Available online at: https://www.eiu.com/n/staying-ahead-of-the-curve-the-business-case-for-responsible-ai/ (accessed September 5, 2022).

European Commission (2021). *Proposal for Regulation of the European Parliament and of the Council—Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.*

Floridi, L. (2019). Translating principles into practices of digital ethics: five risks of being unethical. *Philos. Technol.* 32, 185–193. doi: 10.1007/s13347-019-00354-x

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. (2018). AI4People—An ethical framework for a good ai society: opportunities, risks, principles, and recommendations. *Minds Mach.* 28, 689–707. doi: 10.1007/s11023-018-9482-5

Floridi, L., Holweg, M., Taddeo, M., Amaya Silva, J., Mökander, J., Wen, Y., et al. (2022). *capAI — A Procedure for Conducting Conformity Assessment of AI Systems in Line With the EU Artificial Intelligence Act.*

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., et al. (2018). *Datasheets for Datasets.* Available online at: http://arxiv.org/abs/1803,09010 (accessed September 5, 2022).

Gianni, R., Lehtinen, S., and Nieminen, M. (2022). Governance of responsible ai: from ethical guidelines to cooperative policies. *Front. Comput. Sci.* 4, 873437. doi: 10.3389/fcomp.2022.873437

Hodges, C. (2015). Ethics in business practice and regulation. *Law and Corporate Behaviour : Integrating Theories of Regulation, Enforcement, Compliance and Ethics* (London, UK: Hart Publishing), 1–21.

Holweg, M., Younger, R., and Wen, Y. (2022). The reputational risks of AI. *Calif. Manage. Rev.* 64, 1–12. Available online at: www.cmr.berkeley.edu/2022/01/the-reputational-risks-of-ai/

IEEE (2019). "The IEEE global initiative on ethics of autonomous and intelligent systems," in *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition.* Available online at: https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html

Kleinberg, J. (2018). "Inherent trade-offs in algorithmic fairness," in *Abstracts of the 2018 ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '18)* (New York, NY: Association for Computing Machinery). doi: 10.1145/3219617.3219634

Lauer, D. (2020). You cannot have AI ethics without ethics. *AI Ethics*, 0123456789, 1–5. doi: 10.1007/s43681-020-00013-4

Laurie, G., Stevens, L., Jones, K. H., and Dobbs, C. (2015). *A Review of Evidence Relating to Harm Resulting from Uses of Health and Biomedical Data.* Nuffield Council on Bioethics. Available online at: http://nuffieldbioethics.org/wp-content/uploads/A-Review-of-Evidence-Relating-to-Harms-Resulting-from-Uses-of-Health-and-Biomedical-Data-FINAL.pdf

Lea, H., Hutchinson, E., Meeson, A., Nampally, S., Dennis, G., Wallander, M., et al. (2021). Can machine learning augment clinician adjudication of events in cardiovascular trials? A case study of major adverse cardiovascular events (MACE) across CVRM trials. *Eur. Heart J.* 42. doi: 10.1093/eurheartj/ehab724.3061

McKinsey & Company (2021). *Global Survey: The State of AI in 2021.* Available online at: https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2021 (accessed September 5, 2022).

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., et al. (2019). "Model cards for model reporting," in *FAT* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency, Figure, vol. 2,* pp. 220–229.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nat. Mach. Intell.* 1, 501–507. doi: 10.1038/s42256-019-0114-4

Mökander, J., and Axente, M. (2021). Ethics-based auditing of automated decision-making systems : intervention points and policy implications. *AI Soc.* 0123456789. doi: 10.1007/s00146-021-01286-x

Mökander, J., and Floridi, L. (2022). Operationalising AI Governance through Ethics-based Auditing: An Industry Case Study. *AI Ethics.* doi: 10.1007/s43681-022-00171-7

Mökander, J., Sheth, M., Watson, D., and Floridi, L. (2022). *Models for Classifying AI Systems: The Switch, the Ladder, and the Matrix.* 1–31. Available at SSRN: https://ssrn.com/abstract=4141677

Morley, J., Floridi, L., Kinsey, L., and Elhalal, A. (2020). From what to how: an initial review of publicly available ai ethics tools, methods and research to translate principles into practices. *Sci. Eng. Ethics* 26, 2141–2168. doi: 10.1007/s11948-019-00165-5

Nadler, E., Arondekar, B., Aguilar, K. M., Zhou, J., Chang, J., Zhang, X., et al. (2021). Treatment patterns and clinical outcomes in patients with advanced non-small cell lung cancer initiating first-line treatment in the US community oncology setting: a real-world retrospective observational study. *J. Cancer Res. Clin. Oncol.* 147, 671–690. doi: 10.1007/s00432-020-03414-4

OECD (2019). *Recommendation of the Council on Artificial Intelligence.* OECD/LEGAL/0449.

OECD (2022). "OECD framework for the classification of AI systems", in *OECD Digital Economy Papers, No. 323* (Paris: OECD Publishing). doi: 10.1787/cb6d9eca-en

Office of U.S. Senator Ron Wyden (2022). *Text – H.R.6580 – 117th Congress (2021–2022): Algorithmic Accountability Act of 2022.* Available online at: http://www.congress.gov/

PwC (2017). *Sizing the Price: What's the Real Value of AI for your Business and how can you Capitalise?* Available online at: www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf (accessed September 5, 2022).

Renda, A., Arroyo, J., Fanni, R., Laurer, M., Maridis, G., Devenyi, V., et al. (2021). *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe.* Available online at: https://digital-strategy.ec.europa.eu/en/library/impact-assessment-regulation-artificial-intelligence (accessed September 5, 2022).

Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., Stahl, B., et al. (2021). Research and Practice of AI Ethics: A Case Study Approach Juxtaposing Academic Discourse with Organisational Reality. *Sci. Eng. Ethics* 27, 1–29. doi: 10.1007/s11948-021-00293-x

Schneider, G. (2019). Mind and machine in drug design. *Nat. Mach. Intell.* 1, 128–130. doi: 10.1038/s42256-019-0030-7

Silverman, K. (2021). Why your board needs a plan for ai oversight. *MIT Sloan Manage. Rev.* 62, 14–17. Available online at: www.sloanreview.mit.edu/article/why-your-board-needs-a-plan-for-ai-oversight/

Strathern, M. (1997). "Improving ratings": audit in the British university system [Article]. *Eur. Rev. (Chichester, England),* 5, 305–321. doi: 10.1002/(SICI)1234-981X(199707)5:3<305::AID-EURO184>3.0.CO;2-4

Taddeo, M., and Floridi, L. (2018). How AI can be a force for good. *Science* 361, 751–752. doi: 10.1126/science.aat5991

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* 25, 44–56. doi: 10.1038/s41591-018-0300-7

Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M., et al. (2021). The ethics of algorithms: key problems and solutions. *AI Soc.* 37, 215–230. doi: 10.1007/s00146-021-01154-8

Viljanen, M., and Parviainen, H. (2022). AI applications and regulation: Mapping the regulatory strata. *Front. Comput. Sci.* 3. doi: 10.3389/fcomp.2021.779957

Wang, P. (2019). On defining artificial intelligence. *J. Artif. General Intell.* 10, 1–37. doi: 10.2478/jagi-2019-0002