

**ETHICS**

# On the ROI of AI Ethics and Governance Investments: From Loss Aversion to Value Generation

by Heather Domin, Francesca Rossi, Brian Goehring, Marianna Ganapini, Nicholas Berente, and Marialena Bevilacqua

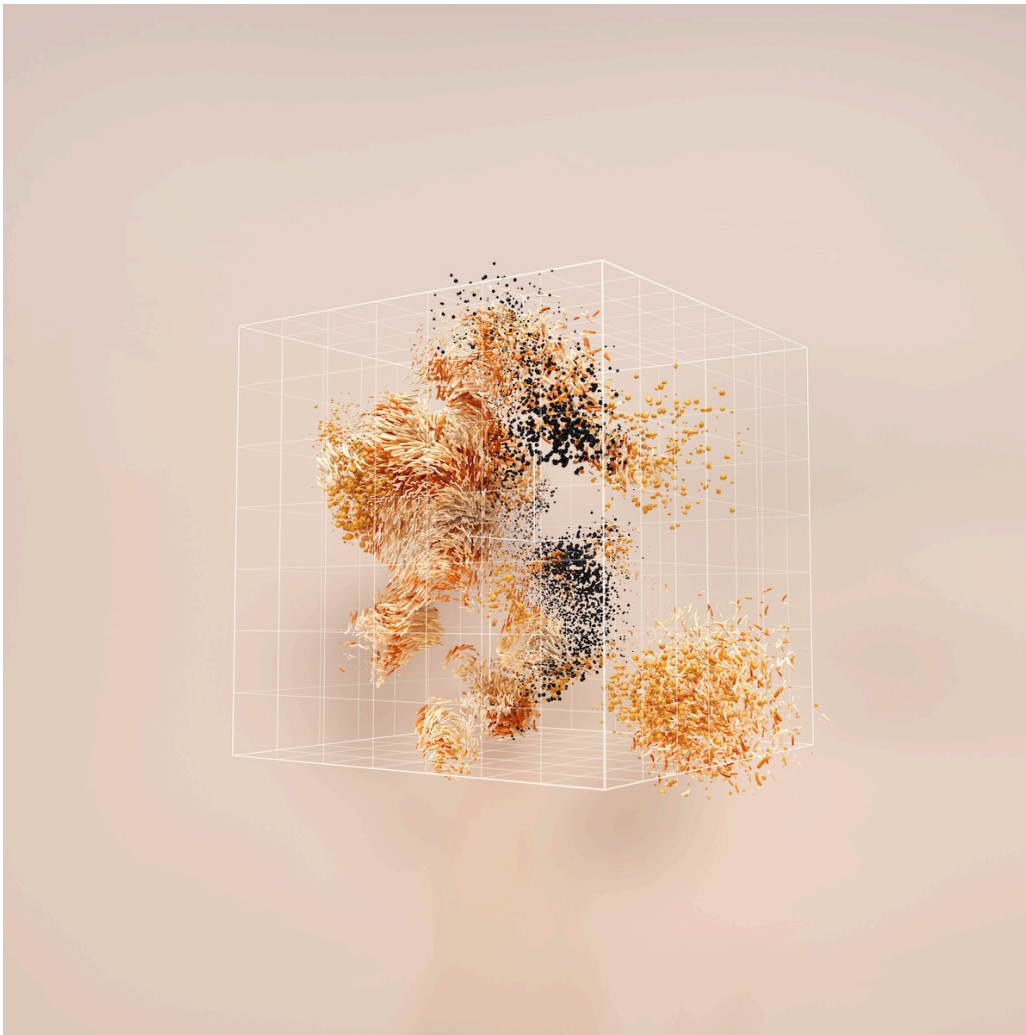


Image Credit | Google DeepMind

*Tackling the ROI question around AI ethics and governance to proactively generate value.*

AI presents an opportunity for growth and productivity improvements for organizations around the world, and at this point, many recognize its tremendous potential. Because of this, AI investment is poised to approach \$200 billion globally by 2025.<sup>1</sup> Generative AI is expected to be a major part of that growth, adding as much as \$4.4 trillion annually to the global economy.<sup>2</sup>

#### RELATED CMR ARTICLES

**“Beyond the “Win-Win”: Creating Shared Value Requires Ethical Frameworks”** by Gastón de los Reyes, Jr., Markus Scholz, and N. Craig Smith

**“Data-Driven, Data-Informed, Data-Augmented: How Ubisoft’s Ghost Recon Wildlands Live Unit Uses Data for Continuous Product Innovation”** by Karl Werder, Stefan Seidel, Jan Recker, Nicholas Berente, John Gibbs, Nouredine Abboud, and Yossef Benzeghadi

However, the ethical issues that arise with the AI technologies in terms of privacy, bias, explainability, misinformation, and their impact on societal structures such as the job market, education, and government, present legitimate concerns. Whether they build or use AI, companies increasingly realize that they cannot simply develop and use this technology without careful attention to how to identify the relevant ethics issues and build governance mechanisms to ensure that they are mitigated.

Often, civil society organizations are skeptical as to whether organizations are doing enough in AI ethics and governance to ensure the ethical development and use of AI. Indeed, studies show that many organizations are investing rather extensively in AI ethics and governance, but many argue that these investments are insufficient.<sup>3, 4, 5</sup> But how should managers determine, justify, and evaluate investments in AI ethics and governance?

Through our research with organizations across a variety of fields, we found that many organizations take a “loss aversion” approach to justify investments in AI ethics and governance. This stands to reason, because loss aversion is a powerful decision criterion for all sorts of organizational decisions.<sup>6, 7</sup> This approach often suits environments where organizations are focused on the near-term tactical decisions, and greater certainty is required before investing. However, ethically related investments that emphasize only loss aversion may be myopic and stand in the way of potential value generation opportunities, which are often longer-term and therefore more uncertain in their nature.<sup>8</sup> Therefore, it is important to understand how organizations can generate value from AI ethics and governance investments, as well as how they can reduce risks and direct costs associated with loss aversion approaches.

A holistic approach to the return on investment in AI ethics and governance, which addresses both direct returns, often motivated by loss aversion, and also the value associated with indirect returns, often requiring a value generation approach is needed.<sup>9</sup> The loss averse approach to AI investments involves avoiding costs such as those associated with regulatory compliance, as well as keeping customers or other stakeholders. The value generation approach involves strengthening intangibles and capabilities for potential competitive differentiation and longer-horizon innovation.

We assume that organizations are determining how important AI is to their business strategy – from its use in operations to products and services to business models.<sup>5</sup> In some cases they use a tactical approach, and in others a more strategic approach to AI. In this paper, we provide concrete examples of both loss averse and value generating justifications for AI ethics and governance investments. Together, these views can help determine the degree to which AI ethics and governance are a strategic imperative for an organization. We then provide recommendations on what companies should do to get the best (not just the most) out of their AI ethics and governance investments, both for them and for society at large.

## **Loss Aversion Justifications for AI Ethics and Governance Investments: *Reactive but More Certain***

There is a growing business case for implementing AI with the appropriate level of governance to ensure ethical development, deployment, and use. For example, companies that focus on implementing generative AI with the appropriate guardrails may be 27% more likely to achieve higher revenue performance than their peers that do not.<sup>10</sup> Implementing AI without governance for certain systems, particularly those that are higher risk, could quite literally block the sale of a product from an entire market. For example, when the EU AI Act goes into effect, high-risk AI systems that do not conform to the requirements of the act cannot be placed on the market in the EU.<sup>11</sup> Even without laws and clear legal standards, ethical concerns can drive significant costs to an organization if they result in lawsuits and legal challenges, or damage to the organization's reputation.

To understand the various dimensions of these kind of justifications, we have interviewed leaders across industries as part of an ongoing collaborative project supported by the Notre Dame – IBM Tech Ethics Lab. Our initial research shows that justifications for AI ethics and governance investments are frequently made when the organization acknowledges the importance of AI to their strategy. Furthermore, decision makers within these organizations have reasoned that if AI is important to the organization, then they must also invest in proper AI ethics and governance efforts that help to maximize the overall ROI in AI investments.

However, we observe that organizations often justify AI ethics and governance investments with a focus on loss aversion that generates what can be called “reactive” justifications, since they react to a risk posed by some regulation, or some trend in the market, or other events. Reactive justifications emphasize, for

example, regulatory compliance (if a company is not compliant, it may face a fine), client trust (if a company is not trusted, its clients could buy from some other vendors), or brand reputation protection (if not doing enough around AI societal impact, a company may be portrayed negatively on the media). All these considerations are around avoiding costs or reducing loss, rather than proactively making investments that go beyond required actions.

More precisely, regulatory compliance justifications assume that AI ethics and governance investments are a necessary cost to lower the risk of noncompliance fines. This includes a need to respond to regulations specifically regarding AI, and also other relevant regulations such as those focusing on data and privacy. It can also involve preparation for regulatory compliance during periods of anticipated regulatory activity – such as when AI-based breakthroughs dominate the public discourse and draw attention from regulators.

Examples of regulatory compliance justifications:

- Avoid a regulatory fine.
- Avoid legal costs.
- Implement required technical compliance mechanism.
- Enable business for required compliance.

In addition to regulatory compliance, organizations often use trust justifications relating to clients or partners, which are also reactive in nature. Trust justifications for investments in AI ethics are typically underpinned by concern that not taking the necessary action could result in a loss of customers or other another important relationship with a stakeholder. These justifications can also arise in response to a specific client need that is tied to the adoption of AI for a specific use case or technology like generative AI, or in response to widespread concern from stakeholders about a product that threatens an organization's business model and attracts the attention of senior leaders. Organizations can therefore often justify AI ethics investments in response to a trust analysis concerning AI.

Examples of trust justifications:

- Respond to stakeholder concerns.
- Avoid threat to business model.
- Meet specific client request or need.
- Avoid risk to brand reputation.
- Maintain trust of clients and stakeholders.

# Value Generation Justifications for AI Ethics and Governance Investments: *Proactive but Less Certain*

Focusing on compliance and trust can be a result of viewing investments in AI ethics and governance narrowly, in terms of cost avoidance only. However, the benefits of investments in this area are much broader. Several organizations recognize this and put forward “proactive” justifications for such investments. Of course, it is often difficult to take on a proactive approach upfront, so many organizations start by justifying AI ethics and governance investments using reactive approaches at first, and then move to a more proactive one as they mature in their AI and AI ethics and governance journey. In fact, as the AI maturity level increases, organizations may be willing to consider actions with a longer-term impact and may feel more confident in making AI ethics and governance an integral part of their operations.

Indeed, proactive justification usually involve long-term planning to scale the use of AI, an organizational desire to help ensure its trustworthiness over time, and a willingness to enhance the organization’s reputation also to gain competitive advantages. Proactive approaches are focused on value generation since they often involve building unique organizational capabilities. These capabilities can involve learning deeply about AI and their applications which can translate, in the shorter term to improving employee efficiency or productivity, responsiveness to customers, and client satisfaction, when combined with responsible leadership in implementing AI technologies that are supported by integrated guardrails. In the longer term, these capabilities can enable the organization to anticipate, sense, and adroitly respond to changes in technologies, business conditions, and regulatory changes.

Further, proactive approaches tend to emphasize ways in which investments drive important intangibles that are reflected in a differentiated reputation and robust corporate culture. Mature firms link these intangible outcomes directly to investments in AI ethics and governance, using criteria such as recognitions of industry leadership, support of Environmental, Social and Governance (ESG) efforts, as well as the long-term ability to manage risks. Organizations may also identify enhanced revenue opportunities that could be supported by training and developing AI ethics and governance experts. Furthermore, and most significant, many organizations proactively justify AI ethics and governance investments due to a sincere commitment to values-based leadership and the recognition of an inherent social contract with all societal stakeholders, as well as a desire to protect vulnerable individuals and communities, and to have a positive impact on the world. In this sense, significant investments in AI ethics and governance, supported by a proactive and value generation approach, align with organizations that are authentically committed to their espoused values and they believe that AI is a very powerful and impactful technology, and that investing in AI ethics and governance is the right thing to do.

Examples of value generation justifications:

- Enable long-term plan to scale AI responsibly.

- Build organizational capabilities.
- Improve employee efficiency or productivity.
- Align with values to advance as an industry leader.
- Seize a market opportunity.
- Protect vulnerable individuals and communities.
- Increase customer satisfaction.
- Demonstrate trustworthiness and maturity.
- Support Environmental, Social and Governance (ESG) efforts.
- Increase ability to manage risk over long-term.
- Innovate and obtain a competitive advantage.

## Evaluating the ROI of AI Ethics and Governance Investments: *What and Why*

Our research has a number of implications for how organizations can justify and then evaluate AI ethics and governance investments. The holistic ROI framework we developed addresses both direct returns on investments, and also the value associated with indirect returns to these investments.<sup>9</sup> Direct returns on AI ethics and governance involve reducing costs such as those associated with regulatory compliance, as well as increasing revenues when strong AI ethics and governance practices enable organizations to gain or keep customers or other stakeholders. Indirect returns involve strengthening intangibles and capabilities. Intangibles refer to important yet difficult to quantify elements such as an organization's brand and culture that support positive returns. Capabilities refer to the organizational knowledge and technical infrastructure that, once in place for governance, offers options for being leveraged and repurposed to provide value throughout an organization.

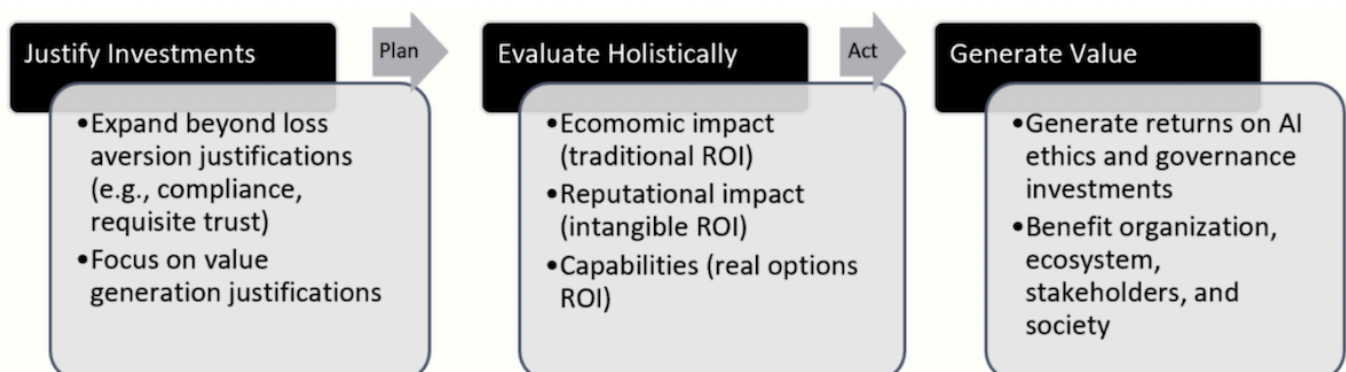
AI ethics and governance investments can include an AI Ethics Board, an Ethics by Design framework, an Integrated Governance Program, and training programs covering AI ethics and governance.<sup>12, 13, 14</sup> For an investment in an AI Ethics Board, committee, or team that functions like IBM's AI Ethics Board, an organization might determine that these efforts help prevent regulatory fines (traditional ROI), increase client trust and partner endorsements resulting in direct leads (intangible ROI), and help enable the development of management system tooling that improves capabilities for more automated documentation and data management (real option ROI). Other common investments include design and review activities during AI development such as the red-teaming of AI models and notifications regarding the use or potential risks associated with AI.<sup>4</sup>

According to research conducted by the IBM Institute for Business Value, the number of organizations that embrace both AI and AI ethics is expected to grow, and a majority of companies expect both of these things to be important strategic enablers.<sup>15</sup> Seventy-five percent of executives surveyed viewed ethics as an important source of creating competitive differentiation.<sup>15</sup> A study from the Economist Intelligence

Unit had similar findings, including that that responsible AI practices can build competitive advantage through factors such as product quality, ability to attract and retain talent, and expanding revenue streams.<sup>16</sup> These studies highlight the importance embracing proactive approaches to AI ethics and governance, particularly in an environment where AI is of increasingly strategic importance to many organizations. As more AI regulations come into force around the globe, we can expect that a basic level of regulatory compliance and loss aversion justifications will become common. Focusing on value generation can provide a competitive advantage in an environment where regulatory compliance is the norm.

## Moving from Loss Aversion to Value Generation: *How*

Organizations that have identified AI as strategically important can move towards proactive approaches to AI ethics and governance by focusing on value generation justifications and measuring ROI holistically. This can be done through a process of identifying the relevant value generation justifications as the organization plans and then evaluates its potential investments, and then acting upon the results of the holistic evaluation to generate the value to the organization. To start, it is important for an organization to identify the specific value generation justifications for AI ethics and governance that may apply to the AI use case (e.g., ability to responsibly improve the quality of responses to customer questions, increased employee productivity and job satisfaction). Leaders within organizations can benefit from thinking through the anticipated stakeholder impacts of the AI use case and identifying potential metrics or indicators of direct economic returns (e.g., value of expanded customer base, reduced employee retention costs), intangible reputational returns (e.g., earned media value of customer reviews), and capabilities and knowledge returns from real options (e.g., improved customer response quality that leads to more first contact resolutions). Rooting an AI implementation strategy in value generation justifications and identifying the potential returns holistically can help organizations move from approaches focused on loss aversion to value generation. Furthermore, it can help maximize the potential returns on their investments in AI ethics and governance while simultaneously benefitting their stakeholders, the ecosystem in which they operate, and society as a whole.





# Conclusion

Artificial intelligence (AI) has caught the public imagination for decades in movies and popular culture. Still, it is only since the release of ChatGPT in November 2022 that we hear phrases like “AI is here” more frequently. The term AI has been used for decades, but we are only just now gaining widespread agreement on how to define it. For example, there is a recently updated OECD definition of AI that provides a basis for the definition within the European Artificial Intelligence Act (EU AI Act) and is ultimately expected to be adopted by many other countries and organizations.<sup>17</sup> In January 2024, the theme “Artificial Intelligence as a Driving Force for the Economy and Society” was one of four key themes at the World Economic Forum Annual Meeting, where world leaders also discussed how to ensure responsible use and deployment of this technology.<sup>18</sup> Within the last few years, we have observed a shift from projected growth in AI and principles to rapid adoption and concrete actions.

As the rapid adoption of AI and generative AI applications like ChatGPT capture the attention and imagination of leaders around the world, many are asking whether AI and its governance are a strategic imperative for their organization. The answer can vary based on factors like operating model, industry, use case, and organizational readiness. However, many companies are determining that AI and its governance are a strategic imperative and have actively taken steps to embrace it. In fact, it is not just companies, but also governments and other organizations that are making it a strategic imperative. Governments and other organizations around the world recognize that the use of AI presents a critical opportunity for economic growth and that the risk of not using AI can result in a widening gap in economic advancement between areas of the world that leverage AI and those that use it less, while at the same time acknowledging it has other risks associated with adoption.<sup>19</sup> These challenges are particularly acute in areas in the Global South (Okolo, C., 2023).<sup>20</sup> In fact, organizations (and countries) that view AI strategically and AI ethics and governance proactively may view their investments as not just an organizational imperative but also a societal imperative.

# References

1. Goldman Sachs. (2023). AI investment forecast to approach \$200 billion globally by 2025. <https://www.goldmansachs.com/intelligence/pages/ai-investment-forecast-to-approach-200-billion-globally-by-2025.html>
2. McKinsey & Company. (2023). The economic potential of generative AI: The next productivity frontier. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier?gclid=Cj0KCQjwl8anBhCFARIsAKbbpyTeLnz0c6i4X2UTnmWdO1KGQnE1mUR8ErJSrM0eMnWDxgfaZukl>



3. Kiron, D. & Mills, S. (2023). Is your organization investing enough in responsible AI? 'Probably not,' says our data. <https://sloanreview.mit.edu/article/is-your-organization-investing-enough-in-responsible-ai-probably-not-says-our-data/>
4. Dotan, R., Rosenthal, G., Buckley, T., Scarpino, J., Patterson, L., Bristow, T. (2023). Evaluating AI governance: Insights from public disclosures. <https://www.ravitdotan.com/evaluating-ai-governance>
5. IBM Institute for Business Value. (2023). CEO decision-making in the age of AI. <https://www.ibm.com/thought-leadership/institute-business-value/c-suite-study/ceo>
6. Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1), 193-206.
7. Greve, H. R., Rudi, N., & Walvekar, A. (2021). Rational Fouls? Loss aversion on organizational and individual goals influence decision quality. *Organization Studies*, 42(7), 1031-1051.
8. Camilli, R., Mechelli, A., Stefanoni, A., & Rossi, F. (2023). Addressing Managerial Loss Aversion for the Corporate Value Creation Process: A Critical Analysis of the Literature and Preliminary Approaches. *Administrative Sciences*, 14(1), 5.
9. Bevilacqua, M., Berente, N., Domin, H., Goehring, B., Rossi, F. (2023). The return on investment in AI ethics: A holistic framework. <https://arxiv.org/abs/2309.13057>
10. IBM Institute for Business Value. (2023). The CEO's guide to generative AI: customer and employee experience. <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ceo-generative-ai/employee-customer-experience>
11. European Parliament. (2023). AI Act: A step closer to the first rules on Artificial Intelligence: News: European parliament. <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>
12. IBM. (n.d.). Building trust in AI. <https://www.ibm.com/case-studies/ibm-cpo-pims>
13. IBM. (2023). How IBM's Chief Privacy Office is building on the company's trustworthy AI framework and scaling automation to address AI regulatory requirements. <https://www.ibm.com/downloads/cas/7PWAWRQN>
14. IBM. (2023). IBM commits to train 2 million in artificial intelligence in three years, with a focus on underrepresented communities. <https://newsroom.ibm.com/2023-09-18-IBM-Commits-to-Train-2-Million-in-Artificial-Intelligence-in-Three-Years,-with-a-Focus-on-Underrepresented-Communities>

15. IBM Institute for Business Value. (2022). AI ethics in action. <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ai-ethics-in-action>
16. Economist Intelligence Unit. (2020). Staying ahead of the curve – The business case for responsible AI. <https://www.eiu.com/n/staying-ahead-of-the-curve-the-business-case-for-responsible-ai/>
17. Bertuzzi, L. (2023). EU lawmakers set to settle on OECD definition for Artificial Intelligence. <https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/>
18. World Economic Forum. (2024). World Economic Forum annual meeting. <https://www.weforum.org/events/world-economic-forum-annual-meeting-2024/>
19. World Economic Forum. (2024). Global risks report 2024. <https://www.weforum.org/publications/global-risks-report-2024/digest/>
20. Okolo, C. (2023). AI in the Global South: Opportunities and challenges towards more inclusive governance. <https://www.brookings.edu/articles/ai-in-the-global-south-opportunities-and-challenges-towards-more-inclusive-governance/>

## Acknowledgements

The authors wish to acknowledge support from the Notre Dame – IBM Technology Ethics Lab and to thank Nicole McAlee and Jungmin Lee for their research program assistance.

---



Heather Domin [Follow](#)

Heather Domin is Global Leader, Responsible AI Initiatives at IBM and Associate Director, Notre Dame - IBM Tech Ethics Lab. She has been instrumental in developing and executing AI ethics and governance practices at IBM and within the industry and has served as a World Economic Forum Fellow.



Francesca Rossi [Follow](#)

Francesca Rossi is the IBM AI Ethics Global Leader. She is based at the T.J. Watson IBM Research Lab, New York, USA, where she leads AI research projects and she co-chairs the IBM AI Ethics board. She is the current president of AAAI, the world-wide association of AI researchers.



Brian Goehring [Follow](#)

Brian Goehring, global AI research leader for IBM's Institute for Business Value, brings over 25 years' experience in strategy consulting to developing data-driven thought leadership with academics, clients, and other experts. He received an A.B. in Philosophy from Princeton University with certificates in Cognitive Studies and German.



Marianna Ganapini [Follow](#)

Marianna B. Ganapini is an associate professor of philosophy at Union College. She has numerous prestigious publications and is currently a Visiting Fellow at the Notre Dame-IBM Tech Ethics Lab. She has collaborated with IBM research on the project "Thinking Fast and Slow in AI". She is co-founder of LogicaNow.



Nicholas Berente [Follow](#)

Nicholas Berente is Professor of IT, Analytics, and Operations and Senior Associate Dean of Academic Programs at the University of Notre Dame's Mendoza College of Business. He is Senior Editor at MIS Quarterly and at Information and Organization.



Marialena Bevilacqua [Follow](#)

Marialena Bevilacqua is a third-year PhD student in Analytics at the University of Notre Dame's Mendoza College of Business. Her research focuses on the intersection of AI technologies, ethics, and moral decisions, in both domestic and international domains.