

Research Brief – October 2020



AI Ethics: Why does it matter?

As the IEAI celebrates its one-year anniversary, we have collected some reflections on the most relevant aspects of AI ethics research from our project leaders, partners of the Global AI Ethics Consortium and RAIN-Africa, as well as guest speakers from our Speakers Series and other workshops. In this Research Brief, 19 experts from across diverse fields, sectors and locations shed a light on the complexity and current relevance of this issue.

AI is a powerful force transforming lives, interactions, environments and societies. We use AI every day, mostly without noticing. Self-driving vehicles, digital assistants, chatbots, face recognition and personalized recommendations are just a few examples of AI applications. AI is here to stay and will continue to transform society substantially. The most pressing questions are how this transformation will take place and what the repercussions will be.

The introduction of increasingly sophisticated AI-enabled systems in many areas of everyday life comes with promises of improvements in social, physical and environmental well-being. Nevertheless, growing concerns about ethical challenges accompany the development of these cutting-edge tools, such as bias, inequity and loss of autonomy. Consideration of these challenges will help to ensure the development of advanced and “trustworthy” AI systems. Achieving the dual advantage of “ethical AI”, where technological innovations capitalize upon opportunities to improve well-being, sustainability and justice and, at the same time, foresee, minimize and mitigate associated risks¹, requires the consideration of several aspects of the development of these systems: from the design, to the context of deployment, to the users and needed data.²

The Institute for Ethics in Artificial Intelligence (IEAI) seeks to be a leader in this process by exploring the ethical issues related to the development, use and impact of AI-enabled technologies. Founded in October 2019, the IEAI celebrates its first anniversary this autumn with a special Research Brief issue examining “AI Ethics: Why does it Matter?”

6 questions and 19 experts from around the globe

As the IEAI’s one-year anniversary approached, we decided to collect some reflections on the most relevant aspects of the AI ethics research through a questionnaire put out to our project leaders,

¹ Floridi, L., Cowsls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Schafer, B. (2018). AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707.

partners of the [Global AI Ethics Consortium](#) and [RAIN-Africa](#), as well as guest speakers from our [Speakers Series](#) and other workshops. These questions were purposefully broad and open-ended in order for each respondent to answer based on their interests and expertise.

**“AI is a global challenge.
So is AI ethics.”**

Prof. Christoph Lütge
Director, TUM IEAI

The various uses of AI, as well as the ethical challenges related to them, require multidisciplinary and multi-stakeholder engagement, and international collaboration on developing AI governance. As TUM’s strategic partnership initiative articulates: “No single university or country can master today’s scientific challenges on its own.”

The IEAI holds public events and supports and participates in several networks related to AI ethics. By joining forces with researchers and practitioners worldwide, the IEAI aims to address real-world challenges and contribute to the broader conversations and concerns surrounding ethics and AI on an international level.

An amazing 19 experts from across diverse fields, sectors and locations responded. They include members of academia, directors of research centers, entrepreneurs and partners of our Global AI Ethics Consortium and RAIN-Africa ([See Appendix](#)). Their diverse backgrounds, interests and expertise shed a light on the complexity and current relevance of this issue.

² Cath, C., Zimmer, M., Lomborg, S., & Zevenbergen, B. (2018). Association of Internet Researchers (AoIR) roundtable summary: artificial intelligence and the good society workshop proceedings. *Philosophy & Technology*, 31(1), 155-162.

This Brief is an attempt to further understand the importance of AI ethics, the dilemmas and misconceptions associated with AI and the role of the various stakeholders when it comes to the ethics and governance of AI as viewed by the experts in the field themselves. In the following sections, we outline the commonalities, varying viewpoints and distinctions of the experts' responses. The final section reflects upon and summarizes their observations.

Why and where AI ethics matters today

Respondents were asked: Why has AI ethics become such an important topic today? Why do we need AI ethics? And what are the most important ethical dilemmas associated with AI right now?

There was a general agreement that the exponential growth of the application and impact of AI in many areas of everyday life has determined the increasing relevance of the field of AI ethics. "The question (also) arises whether and where exactly AI creates genuinely new problems or whether problems that are already qualitatively known become more urgent" (Uhl). Hence, exploring the consequences of the deployment of AI systems in society is fundamental to ensuring such systems develop in a way that supports the well-being of humans, which should be the ultimate goal. "While AI-based decision-making opens up possibilities that were previously not thought of, it does not operate outside of the domain of ethical responsibilities" (Baumert). "Thus, it is extremely important that such a powerful 'tool' is accepted and understood" (Fottner).

"The questions we ask should no longer be along the lines 'can AI do this?', but rather, 'should AI do this, and how?'"

Christian Adumatta Gyampomah
Kwame Nkrumah University of Science
and Technology

Moreover, the participants believed that the role of the international community is of great importance in setting out universal principles that can help and safeguard humanity as a whole in order to respond to the increasing number of dilemmas and

controversies that inevitably emerge with technological advancement.

The most important ethical dilemmas related to the development of AI reported by respondents included accountability, autonomy, privacy, fairness, (public) trust, explainability and interpretability. However, respondents acknowledge the challenge of trade-offs that are likely to emerge and complicate the relationship between desired solutions and ones that can be applied in reality. "It becomes a challenging task to find the right balance" (Roosen). Hence, there is an urgent need to discuss real-world scenarios and "consider ethics at every stage of AI design, development and deployment" (Weller).

The future of AI ethics as a field

Respondents were next asked: Where do you see the field of AI, or AI ethics, going in the coming years?

"If AI is to genuinely contribute to social good, then there needs to be in place some actionable norms and values which endorse the social dimensions of new technology."

Mark Findlay
Singapore Management University

Our experts heavily agreed on the importance of interdisciplinary research in order to address the issues that emerge with the development of AI systems. Engineers and computer scientists will need to work together with experts from other disciplines to understand the legal, psychological and societal implications of the use of AI-enabled technologies. Moreover, we need to "place humans in the center of the picture in order to [...] have it be a real help for us" (Jacquet).

In the future, research on human interaction with AI will be an interesting area of exploration, assessing how these systems can positively impact human lives, while mitigating potential harms. Hence, AI ethics play a fundamental role in defining the necessary steps to ensure the "trustworthy" development of AI tools. Among

others, priorities include protecting individual interest and privacy, ensuring social justice and digital equity, reducing disparities, protecting the environment as well as defining accountability, responsibility and transparency.

To successfully address these issues in the future, cooperation between all the actors involved in the development of AI systems is fundamental - a point elaborated on in the next section.

What are we missing?

We next asked: Is there anything you think policymakers, developers or the AI ethics community are overlooking at the moment?

Moving from theory to practice: Integrating ethics into AI-enabled tools requires collaboration and mutual understanding between the involved stakeholders in every part of the process of developing an AI system. Nevertheless, there is still a gap between general AI principles and practical applications. The responsibility for ensuring that this gap narrows falls on everyone's shoulders.

“We have to deal with concrete ethical problems and then find solutions that are ethically appropriate but can be implemented by software developers at the same time.”

Markus Lienkamp
Technical University of Munich

Responses highlighted differing focuses or approaches to dealing with ethical issues between disciplines as one significant gap. For example, those from technical disciplines look for ways in which ethical issues can be coded into the AI enabled systems, often in ways that are detached from those who are focused on more abstract ethical issues and dilemmas that have difficult to implement solutions. Therefore, the challenge continues in terms of bridging the identification of ethical issues or the formulation of guidelines, and the tangible implementation of these findings into

AI systems themselves. “Accordingly, the ethics of AI clearly needs a stronger empirical focus” (Uhl).

“Policymakers and developers aren’t thinking enough about how AI can become part of a viable and sustainable social project that enables resilience, human solidarity and shared flourishing in the face of these changes.”

Shannon Vallor
University of Edinburgh

Research based on a single discipline cannot meet the growing need to determine potential and implementable solutions to real-world problems involving the development of AI systems.

Moving from policy to practice: Many of the respondents underlined policymakers' inadequate knowledge of the AI-related issues and challenges. Policy guidelines are often abstract. As a result, regulations and standards lack the empirical evaluations and technical applicability necessary for the developers of AI-enabled systems to integrate principles into applicable solutions. The disconnect between policy and practice, or technical feasibility, is probably a consequence of the absence of a proper understanding of how these systems operate and the ethical implications that emerge from their use. As Darlington Akogo put it, “It’s important that policies, regulations and standards around AI focus on the technical feasibility as well.” “This requires bringing together both technical and legal expertise” (Nemitz). Policymakers and developers need to work together to find effective ways to achieve sound governance that envisions future scenarios. This will also involve not shying away from problems that have to do with conflicting ethical principles.

Moreover, as implementation is discussed, it becomes clear that context matters and much of what is going on in practice is happening below the regional or national level. Local solutions for the deployment of AI systems may offer a more effective process, by achieving a higher coordination between the different actors involved

in the design and deployment of AI tools. The involvement of citizens allows for the development of AI systems that align with the values of actual users and those affected, achieving a higher trust in the deployment of such systems. This viewpoint also has farther-reaching effects, as GAIEC partner Stefaan Verhulst points out, “Regulating AI companies locally has potential knock-on effects on how technology is designed and deployed globally.”

The role of academia

Finally, as an academic institution, we asked: What role can academia, research institutions and other centers play when it comes to the ethics and governance of AI?

“We cannot rely on existing laws and juridical principles that have been written in previous epochs, where the world and the problems were very far away from our current digitized world.”

Jean-Gabriel Ganascia
Sorbonne University

Academia and research centers are crucial stakeholders when it comes to the ethics and governance of AI: from questioning the status quo, seeking answers beyond what is known and thinking ahead to solve issues in a constructive

“Academics can help identify the areas where AI can help social and individual wellbeing, and bring the focus back to that.”

Rafael A. Calvo
Imperial College London

manner. As Markus Lienkamp states, “In academia, two important things come together: independence of interests with the competence in concrete implementation.” Academic research has the ability to investigate over longer time spans and conduct more in-depth examinations, providing independent, unbiased and nuanced in-

sights into ethical challenges related to AI-enabled technologies. “In this way, researchers sometimes come up with new and creative solutions, [...] grasping problems in a more profound way” (Djeffal). It also has the potential to improve the “relation(ship) between democracy and technology” (Nemitz).

There was common ground among the respondents that an interdisciplinary approach is fundamental in order to analyze and tackle the challenges raised by AI holistically, from a technological, sociological, managerial, ethical and legal point of view. Moreover, academia, with its broad access across different disciplines and cultures, can connect and build international research communities, collaborative networks and global consortiums working together on AI ethics-related issues.

“There is a growing and generally highly collaborative network of researchers working in AI ethics. The different research agendas and projects across these organizations can provide a spotlight on and conscience of otherwise unchecked technological expansion.”

Jeannie Paterson
University of Melbourne

Another important point focused on the role of universities to “educate our students with an interdisciplinary approach” (Pouillet) to AI development and to promote courses, workshops and conferences to raise awareness regarding the ethical use of AI.

The thoughts of our respondents lead to some crucial questions that academia needs to answer in order to fulfill its crucial role: What kind of academic community needs to be built in order to meet the ethical challenges related to AI? And how can the members of academia, research institutions and other centers make sure that their recommendations on AI ethics do not only circulate amongst researchers, but actually reach the public?

Conclusion: **Reflecting on Why AI Ethics Matters**

The ideas and reflections of our expert respondents shed a clear light on the importance of AI ethics, along with the biggest challenges and opportunities that lay ahead. Some clear highlights emerged from their responses:

- The importance of AI ethics is only growing, and while there are a variety of questions to be answered, acknowledging trade-offs and ethical challenges early on in the development process is of the utmost importance.
- AI ethics as a field has a fundamental role to play in developing “trustworthy” AI tools. Priorities in the field include protecting individual interest and privacy, ensuring social justice and environmental protection, as well as defining accountability, responsibility and transparency.
- There is a strong need for an interdisciplinary approach to research that moves out of the general/abstract and provides instruction for applying AI ethics in a way that is technically feasible.
- Moreover, there need to be clearer information channels and knowledge sharing from *research to policy to practice* in order to maximize the impact of our work.
- Academic institutions and independent research centers have unique characteristics that will allow them to carry out the needed studies and provide the needed expertise to move the field AI ethics forward, but an interdisciplinary approach is key.

- Finally, universities have an opportunity to ingrain the importance of ethics in AI into the mindsets of our future developers and policymakers early on through integrating AI ethics and the importance of interdisciplinary, multi-stakeholder approaches into core curricula for engineers and social scientists alike.

AI is a powerful tool that has the ability to amplify current positive efforts, but also entrench and increase harms and discrimination if not used responsibly. During 2020, our reliance on AI has skyrocketed. So have the discussions around AI ethics.

The COVID-19 pandemic has made clear the need for a coordinated, dedicated data and technology infrastructure and an ecosystem for tackling the societal threats. This crisis has also underlined the urgent need to discuss the ethical considerations in the use of AI and develop operational ethical frameworks in the field of AI now, before AI-enabled tools need to be rapidly employed.

AI can neither solve humanity's old problems³, nor humanity's all problems.⁴ The big questions about AI ethics are the big questions of society.⁵ Thus, as a society, we need to decide what principles are important. After all, “AI is what we make of it” (Verhulst). The effort to make sure that technologies are beneficial to society needs to be, among other things, a globally collaborative one.⁶ In this effort, academic research institutions, collaborating across cultures and disciplines have the opportunity to make a valuable and unique impact.

³ Neppel, C. (June, 2020). Reflections on AI. [Q&A with Dr. Clara Neppel]. TUM IEAI. Retrieved online from: https://ieai.mcts.tum.de/wp-content/uploads/2020/06/Reflections-on-AI-Ethics_ClaraNeppel_Final.pdf

⁴ Lütge, C. (September, 2020). Reflections on AI. [Q&A with Prof. Dr. Christoph Lütge]. TUM IEAI Retrieved online from: https://ieai.mcts.tum.de/wp-content/uploads/2020/10/Reflections-on-AI-Ethics_ChristophL%C3%B4tgeFinal.pdf

⁵ Bryson, J. (May, 2020). Reflections on AI. [Q&A with Professor Johanna Bryson]. TUM IEAI. Retrieved online from: <https://ieai.mcts.tum.de/wp-content/uploads/2020/05/Reflections-on-AI-Ethics-FINAL.pdf>

⁶ Price, H. (July, 2020). Reflections on AI. [Q&A with Prof. Dr. Huw Price]. TUM IEAI. Retrieved online from: https://ieai.mcts.tum.de/wp-content/uploads/2020/07/Reflections-on-AI-Ethics_HuwPrice_Final-1.pdf

Appendix – List of Respondents

[Darlington Akogo](#)

GUDRA AI, karaAgro AI, minoHealth AI Labs, Runmila AI Institute (Ghana)

[Prof. Dr. Anna Baumert](#)

Max Planck Institute for Research on Collective Goods and TUM School of Education, Technical University of Munich (Germany)

[Prof. Rafael A. Calvo, Ph.D.](#)

Dyson School of Design Engineering, Imperial College London (United Kingdom)

[Prof. Dr. Christian Djeffal](#)

TUM School of Governance and Munich Center for Technology in Society, Technical University of Munich (Germany)

[Prof. Mark Findlay, LL.D.](#)

Centre for AI and Data Governance, Singapore Management University (Singapore)

[Prof. Dr.-Ing. Johannes Fottner](#)

TUM Department of Mechanical Engineering and TUM School of Management, Technical University of Munich (Germany)

[Prof. Dr. Jean-Gabriel Ganascia](#)

LIP6-CNRS, Sorbonne University (France)

[Christian Adumatta Gyampomah](#)

Emerging Networks and Technologies Lab, Kwame Nkrumah University of Science and Technology – KNUST (Ghana)

[Prof. Dr. Ken Ito](#)

The University of Tokyo (Japan)

[Prof. Dr. Jean-Marie Jacquet](#)

Nadi Research Institute, University of Namur (Belgium)

[Prof. Dr. Markus Lienkamp](#)

TUM Institute of Automotive Technology, Technical University of Munich (Germany)

[Paul Nemitz](#)

European Commission (Belgium)

[Prof. Jeannie Marie Paterson, PhD](#)

Centre for AI and Digital Ethics, University of Melbourne (Australia)

[Prof. Dr. Yves Poulet](#)

Nadi Research Institute, University of Namur (Belgium)

[Prof. Dr. Jutta Roosen](#)

TUM School of Management, Technical University of Munich (Germany)

[PD Dr. Matthias Uhl](#)

TUM School of Governance, Technical University of Munich (Germany)

[Prof. Shannon Vallor, Ph.D.](#)

Edinburgh Futures Institute, University of Edinburgh (United Kingdom)

[Prof. Stefaan Verhulst, Ph.D.](#)

The GovLab, New York University (USA)

[Adrian Weller, Ph.D.](#)

University of Cambridge and The Alan Turing Institute (United Kingdom)