Critical Conversations about AI: A Glossary for Big Questions and Bold Ideas.

Contributors

Lyse Langlois, Laval University
Jérôme Duberry, Graduate Institute Geneva
Martin Wählisch, University of Birmingham
Mia Perry, Glasgow University
Sophie Toupin, Laval University
Ève Gaumond, University of Montreal
Marcela Ramos, Glasgow University

Partnership

Obvia, Laval University
Tech Hub, Geneva Graduate Institute
Center for AI in Government, Birmingham University

Produced with the financial support of the Fonds de recherche du Québec

Fourteen terms to discover

Purpose

This glossary offers a critical exploration of rarely discussed terms within the AI ecosystem, presented in an engaging and forward-thinking manner. Through its definitions, examples, and open questions, it aims to inspire reflection and dialogue in an educational setting.

Co-authored by multiple contributors, this glossary was initiated by an international group of colleagues following discussions in Geneva on science diplomacy. It will continue to evolve with input from colleagues and doctoral students, providing social and philosophical perspectives on technological developments.

If you are interested in contributing to the next edition, please contact us at **connaissances@obvia.ca.**

AI Augmentation

Although it remains a challenge to understand what the professional skills of tomorrow will be, computational thinking and the ability to interact with an AI are among them. This is referred to as AI augmentation, where AI technologies augment and enhance human efforts. In contrast, AI automation focuses on replacing human tasks entirely, often prioritizing speed, cost savings, and scalability.

EXAMPLE

A pertinent example is the widespread deployment of AI-driven automation in the workplace. While some businesses may see cost savings and productivity gains, many workers face job displacement, economic insecurity, and the need for reskilling. It is crucial to develop AI applications that augment human efforts, promoting equity and well-being instead of disruption and inequality.

QUESTIONS

- How can AI developments empower human beings?
- What are the professional skills of tomorrow?

Source:
Nguyen Thi Ngoc, T. & Elbanna, A. (2023). Understanding human-ai augmentation in business and management context: a literature review. *ECIS 2023 Research Papers.* 406. https://aisel.aisnet.org/ecis2023_rp/406

AI Narratives

Several narratives underlie many discussions of AI development. First, *the development of AI in itself is progress*, regardless of its impact or potential uses. But for whom? And how to assess its short-, medium-, and long-term impacts? Second, *AI development is inevitable*, as if technology had its own agency and could develop beyond human control. Third, *society and individuals must adapt to AI development*. But shouldn't it be the other way around?

EXAMPLE

Despite repeated calls for a more precautionary approach or a slowdown in the development of AI, given its current and potential future impact on society, AI development has continued unabated. AI is also a geopolitical power factor, given its military applications and the benefits it can bring to the economy. Under the motto "winner takes all," there is a race to be the first to introduce the most powerful AI.

QUESTIONS

- Who is leading AI development?
- What narratives underlie AI discussions and policy decisions?
- Should society and individuals adapt to AI or vice-versa?

Source:
Duberry, J. (2022). Artificial intelligence and democracy: risks and promises of AI-mediated citizen–government relations. In *Artificial Intelligence and Democracy*. Edward Elgar Publishing.

AI Subjectivity

AI, as any technology, is the result of choices and decisions made by a group of individuals and reflects values and interests that are localized in time and space, making it inherently subjective. Those with access to AI education, data, and/or computing power are in a position to choose the pace, direction, characteristics, and priorities of AI development. To date, AI designers and developers are not representative of the diversity of society, which is often cited as a major cause of AI bias and discrimination.

EXAMPLE

An example of subjectivity in AI design within women's health is the development of medical algorithms for diagnosing conditions like heart disease. Historically, many datasets used to train these algorithms were predominantly based on studies involving male patients. As a result, symptoms of heart attacks in women—such as nausea, fatigue, or back pain—were often overlooked because they differ from the "classic" male-centered symptoms like chest pain.

QUESTIONS

- What is involved in the design of AI?
- Who is harmed by AI?
- Who benefits from AI?
- How to include a broader diversity of stakeholders and populations in the design of AI to ensure it addresses the needs of all?

Sources:
Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need.* The MIT Press.
Cave, S. (2020, February). The problem with intelligence: its value-laden history and the future of AI. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 29-35).

Anthropomorphism

Anthropomorphism involves attributing human characteristics to nonhuman objects. The anthropomorphism of AI systems is a common practice. Several systems have human names, are trained to communicate in a familiar and friendly language, and some are integrated within ultra-realistic avatars. These practices are not without risk. They can contribute to unrealistic expectations about the systems and cause confusion regarding the reliability of results.

EXAMPLE

Some companies design AI systems capable of establishing more or less intimate relationships with users. Avatars can be created and customized to encourage social and intimate interactions. The research community strives to understand the impacts of these technologies on human relationships, but the evolution is so rapid that it is difficult to obtain a clear picture of its effects.

QUESTIONS

- What do we aim to achieve when systems are integrated into a human form?
- What is the intended purpose?
- Do technological models favor a particular gender or a specific voice?
- Does the system convey a stereotype?

Source:
Dubé, S., Santaguida, M., Anctil, D., Leshner, C., Vachon, D. et Johnson, A. (2021) Technology-based sexualities. *Encyclopedia of sexuality and gender.*

Artificial Knowing

From a feminist perspective

The term Artificial Knowing is used by Alison Adam to highlight the gendered assumption of “men as the norm” within symbolic AI systems. Already in the 1990s, Adam had identified that AI systems were grounded in a perspective that prioritized a detached, “view from nowhere”, ultimately centering men’s voices, thoughts, and experiences. In the absence of women’s knowledge within these systems, Adam argued that feminists needed to actively engage with AI technologies. The proposed way forward was the development of feminist AI.

EXAMPLE

Artificial feminist knowing projects have since then flourished. Catherine D’Ignazio has developed a machine learning tool to detect mentions of femicide in Spanish and English news media to support feminist activists in their efforts to collect feminicide data. The NGO Data Género in Argentina has developed AymurAI to support criminal courts in Latin America by gathering and providing data on gender-based violence. The A+ Alliance has created a feminist AI research network (F<A+1>).

QUESTIONS

- To which extent feminist imaginaries encourage us to rethink and re-envision artificial knowing in ways that are more inclusive and equitable?
- How might feminist creativity, powered by artificial knowing, be mobilized to counter technology-enhanced violence?

Sources:
Adam, Alison. (1998). *Artificial Knowing: Gender and the Thinking Machine.* London: Routledge.
D’Ignazio, Catherine. and Lauren F. Klein. (2020). *Data Feminism.* Cambridge, MA: MIT Press.

Digital Sovereignty

Digital sovereignty refers to the authority of nations, communities, and groups to govern and control their digital infrastructure, data, and online activities, ensuring alignment with their laws, policies, cultural values, and interests. It encompasses the ability to manage digital systems, technologies, and platforms within their jurisdiction and to safeguard against external dependencies or threats that could undermine their autonomy.

EXAMPLE

The European Union's General Data Protection Regulation exemplifies data sovereignty by enforcing strict rules on protecting EU citizens' data, regardless of where it is processed, ensuring control over personal data and its cross-border flow. Similarly, Indigenous data sovereignty is exemplified by the Māori Data Sovereignty principles in New Zealand, which advocate for Māori control over data related to their communities, ensuring it aligns with their cultural values, governance structures, and self-determination.

QUESTIONS

- Who should control the development and use of AI:
  - Individual nations asserting digital sovereignty to protect their interests and values?
  - Local communities advocating for data sovereignty to uphold cultural integrity?
  - Or global partnerships aiming for shared progress and universal ethical standards?
- How can nations balance the principles of digital sovereignty and data sovereignty with the need for global collaboration and the free flow of data in a highly interconnected digital world?
  - What are the challenges and opportunities for nations in asserting digital sovereignty and data sovereignty?

Sources:

Fratini, S., Hine, E., Novelli, C. & al. (2024). Digital Sovereignty: A Descriptive Analysis and a Critical Evaluation of Existing Models. DISO 3, 59 https://doi.org/10.1007/s44206-024-00146-7

Walter, M., Lovett, R., Maher, B., Williamson, b., Prehn, J., Bodkin-Andrews, G. & Lee, V. (2021). Indigenous data sovereignty in the era of big data and open data. *Australian Journal of Social Issues* 56.2 : 143-156. https://doi.org/10.1002/ajs4.141

Rainie, S., Kukutai, T., Walter, M., Figueroa-Rodriguez, O., Walker, J., & Axelsson, P. (2019). *Issues in Open Data: Indigenous data sovereignty*. Zenodo. https://doi.org/10.5281/zenodo.2677801

Digital Inclusion

Digital inclusion in the age of AI means ensuring that all individuals and communities, regardless of their socioeconomic status, location, or background, have access to AI-driven technologies and can use them effectively. This includes providing affordable internet access, AI-enabled devices, training in AI literacy, and quality support systems to help people understand and benefit from AI tools in their daily lives.

EXAMPLE

AI-powered chatbots and virtual assistants can enhance digital inclusion by providing personalized support for users in multiple languages, including simplified language options for those with limited literacy. For instance, in rural areas, AI tools can guide individuals in accessing healthcare, education, or government services, breaking down barriers caused by lack of resources or technical skills.

QUESTIONS

- How can we design AI to ensure everyone, no matter their location or background, has access to technology, skills, and support?
- How can we manage and protect digital commons, like shared online resources of data, information, and knowledge, to ensure everyone can access and use them fairly?

Sources:

Shams, R., Zowghi, D., & Bano, M. (2023). AI and the quest for diversity and inclusion: a systematic literature review. *AI and Ethics*, 1–28. https://doi.org/10.1007/s43681-023-00362-w

Alhassan, M. D., & Adam, I. O. (2021). The effects of digital inclusion and ICT access on the quality of life: A global perspective. Technology in Society, 64, 101511. https://doi.org/10.1016/J.TECHSOC.2020.101511

Dual Use

The dual-use nature of AI refers to its ability to serve both beneficial and harmful purposes, depending on how it is applied. It is almost impossible to know in advance all the possible uses of AI once they are deployed. This duality makes regulation challenging, as the same technology can be used for vastly different goals, often influenced by intent, context, and access, raising ethical and governance concerns.

EXAMPLE

Generative AI illustrates dual-use risks, particularly with disinformation. On one hand, it creates valuable content, such as realistic simulations, creative works, or language tools enhancing communication. On the other, it can produce highly convincing fake news, deepfake videos, or propaganda, manipulating public opinion and undermining trust in information. For instance, AI-generated deepfakes of public figures have spread false narratives.

QUESTIONS

- How to regulate dual-use technologies?
- How can we ensure that benefits are equally shared among populations and risks effectively mitigated?

Sources:

Urbina, F., Lentzos, F., Invernizzi, C., & Ekins, S. (2022). Dual use of artificial-intelligence-powered drug discovery. *Nature machine intelligence,* 4(3), 189-191.

Vaynman, J., & Volpe, T. A. (2023). Dual use deception: How technology shapes cooperation in international relations. *International Organization*, 77(3), 599-632.

Gullibility Gap

The Gullibility gap refers to the human tendency to overestimate the intelligence of machines. It's a type of illusion that bears similarities with pareidolia, the phenomenon by which the human mind constructs meaningful interpretations for random patterns. Just as we see faces in clouds or sleeping giants in rock formations, we tend to attribute intelligence to machines that don't actually display any.

EXAMPLE

Eliza is probably the most telling example of the "Gullibility Gap". Created by Joseph Weinzebaum in 1966, Eliza was a rudimentary chatbot that merely restated what the user had said in the form of a question. Despite the simplicity of the software, users engaged in deep and intimate conversations with Eliza, so much so that they would often ask to be left alone with the bot to preserve the confidentiality of their conversations with it. Generative AI offers new illustrations of the same phenomenon. In 2022, an ex-engineer from Google alerted the media Google's latest LLM which he believed had become sentient.

The expression "Gullibility Gap" can also apply to more mundane overestimations of intelligence such as believing that empty statements generated by ChatGPT are deeply insightful.

QUESTIONS

- When you come across AI-generated content, ask yourself: what does this sentence actually mean? Try rephrasing it in your own words. Is the content genuinely smart, or does it only sound smart?
- Why do we want machines to be intelligent? How are intelligent machines supposed to make society better? Are current AI systems going in the right direction to do so?

Sources:
Marcus, G & Davis E. (2019). *Rebooting AI - Building Artificial Intelligence we can trust*, New York, Pantheon Books at p. 14.
Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. San Francisco.
Nitasha Tiku, (June 2022), "The Google engineer who thinks the company's AI has come to life", *The Washington Post.*

Human-AI Symbiosis

Human-AI symbiosis refers to the collaboration between humans and artificial intelligence systems to enhance decision-making, creativity, and problem-solving. In this relationship, humans and AI complement each other's strengths: humans bring intuition, ethics, and emotional intelligence, while AI contributes speed, data analysis, and precision. The goal of this symbiosis is not to replace human agency but to empower individuals and organizations to achieve outcomes that neither could accomplish alone.

EXAMPLE

AI-powered tools like decision-support systems in healthcare exemplify human-AI symbiosis. For instance, AI algorithms can analyze vast datasets to identify patterns in medical imaging, but doctors make the final decisions, contextualizing these insights within a patient's unique circumstances. Similarly, in industries like design or architecture, AI can generate creative suggestions or simulate solutions, which humans refine with their expertise and innovation.

QUESTIONS

- How can we foster a collaborative framework between humans and AI that ensures mutual enhancement, ethical outcomes, and equitable access to the benefits of this partnership?

Sources:

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. https://doi.org/10.1016/J.BUSHOR.2018.03.007

Almeida, F., & Senapati, B. (2024). Striving for Symbiosis: Human-Machine Relations in the AI Era. *2024 IEEE Integrated STEM Education Conference (ISEC)*, 1–4. https://doi.org/10.1109/ISEC61299.2024.10664823

Pluriversal Literacies

The pluriverse is a concept that allows and recognizes many different cultures, value systems, ways of being, of living and of thinking as multiple, diverse and coexisting realities around the world.

Literacies describe the capacity to engage with the semiotic practices that allow us to “read”, de-code, make meaning, and communicate in the world.

Pluriversal literacies then describe the many and different ways needed for people in different places and contexts to read and engage in their realities.

EXAMPLE

For many central powers and influences (e.g. UNESCO, OECD, World Bank), linguistic (printed and spoken language) literacy is considered the primary and most important literacy for all. For those that don’t benefit from the market economy including agrarian and land-dependent populations, land and environmental literacies are paramount. For physical scientists, material literacies (e.g. rocks, critical earth elements, etc.) are critical. For all humans, regardless of culture, a literacy of the body is central to the capacity to care for oneself.

QUESTIONS

- What semiotic (or de-coding and meaning making) practices are needed for you (or the people you serve) to flourish while ensuring the sustainability of lands, water, and food security in the current global context?
- What knowledge cannot be communicated through words and numbers alone?
  - What are the literacies that we therefore need to develop to consciously and ethically engage in that knowledge production?

Sources:

Perry, M. (2023). P*luriversal Literacies for Sustainable Futures: When Words are not Enough*. Expanding Literacies in Education Series, Routledge.

Perry, M. (2021). Pluriversal literacies: affect and relationality in vulnerable times. *Reading Research Quarterly*. 56 (2). 293-309.

https://sustainablefuturesglobal.org/pluriversal-literacies-hub/

Social Robots

Social robots are physically embodied autonomous agents capable of interaction and communication with humans or other autonomous agents. Their actions, roles, and interactions can be guided by social behavior patterns and integrated rules. The term “robot” comes from the Czech word “robota” meaning “work, chore” recalling its initial origin of being destined for servile tasks.

EXAMPLE

This term first appeared in a science fiction play written by Karel Čapek: R.U.R. (Rossum’s Universal Robots) in 1920. The play tells the story of biological machines designed for forced labor revolting and replacing humanity.

Today, social robots can be integrated into residences for the elderly to keep them company, assist with medication management, or help with physical exercises.

QUESTIONS

- Does the introduction of social robots alter our understanding of what it means to be human?
- Limits of human-robot interactions: Can a meaningful relationship truly be established with an artificial being?
- Will these robots be accessible to everyone, or will they reinforce economic and social inequalities?
- Do social robots perpetuate gender, cultural, or other social biases?
- What will be the impact of social robots on employment in the sectors where they are deployed, such as healthcare or service industries?

Sources:

Henschel, A., Laban, G. & Cross, E.S. (2021). What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You. *Curr Robot* Rep 2, 9–19. https://doi.org/10.1007/s43154-020-00035-0

Sweeney, P. (2021). A fictional dualism model of social robots. *Ethics Inf Technol* 23, 465–472. https://doi.org/10.1007/s10676-021-09589-9

Techno-solutionism

Reality is complex and requires equally complex answers. Yet technology is often presented as a quick and safe solution with tangible results. Evgeny Morozov coined the term “techno-solutionism” to describe our tendency to turn to technology as an easy solution to real-world challenges. It is related to our tendency to see data sets as objective and true representations of the world. But datafication, the quantification of human life through digital information, is subjective and reflects societal biases.

EXAMPLE

AI-driven predictive policing can be presented as an efficient way to reduce crime by algorithmically directing patrols. While it appeals to policymakers and funders for its tangible results, the reality is complex. These systems often rely on biased historical data, disproportionately target marginalized communities, and perpetuate systemic inequalities. By focusing on symptoms rather than root causes such as poverty and education, predictive policing oversimplifies crime prevention.

QUESTIONS

- When is a societal problem susceptible to being addressed by AI, and when is it not?
- Is it possible to translate the complexity and diversity of the world into data?

Sources:
Mejias, U. A. & Couldry, N. (2019). Datafication. *Internet Policy Review,* 8(4). https://doi.org/10.14763/2019.4.1428
Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism.* PublicAffairs.

Transhumanism

Transhumanism is the idea that humans should transcend their current natural state and limitations through the use of technology, that we should embrace self-directed human evolution. Based on Moore's Law, which states that technological capabilities double every 18 months, transhumanists believe that this law can apply to human evolution. For some authors, this belief falls under science fiction. Transhumanism is therefore an intellectual and cultural movement asserting that it is desirable to use technologies and biomedical sciences to enhance our health, abilities, and lifespan, even if it leads us to evolve into a new species called post-humans.

EXAMPLE

Advances in genetic engineering allow us to correct genetic disorders, enabling individuals to become healthier and potentially better versions of themselves.

Connected brain implants can communicate directly with computers, allowing individuals to control devices with their thoughts and potentially enhance their cognitive functions. Technologies are being developed to improve memory, speed up information processing, and enable accelerated learning, significantly boosting our cognitive capabilities.

The concept of mind uploading involves transferring our consciousness to a digital format, potentially allowing us to live beyond our biological limitations.

QUESTIONS

- Can a machine become conscious?
- Can human relationships be controlled and used for any purpose?
- Will AI put an end to what is human and non-performing?
- Will humans always remain human or become something else?

Sources:

Monin, P. (2019). Ganascia J.G., Le mythe de la singularité. Faut-il craindre l'intelligence artificielle ?, Le Seuil, Coll. Science ouverte, 2017. *RIMHE : Revue Interdisciplinaire Management, Homme & Entreprise*, n° 36, vol. 8(3), 121-128. https://doi.org/10.3917/rimhe.036.0121.

Hottois, G. (2014). *Le transhumanisme est-il un humanisme ?* Paris : Vrin.

Sandel, M. (2007). *The Case Against Perfection : Ethics in the Age of Genetic Engineering*. Cambridge, Mass. : Belknap Press of Harvard University Press.