

co-creation and by the fostering of a public discourse in which “experts” and “public” are united in the acknowledgment of a shared vulnerability before the unknown, whether this unknown takes a technological, societal, or ethical shape.

A robot may not injure a human being or, through inaction, allow a human being to come to harm.

– Isaac Asimov, *I, Robot*

Whenever there has been progress, there have been influential thinkers who denied that it was genuine, that it was desirable, or even that the concept was meaningful.

– David Deutsch, *The Beginning of Infinity*

The ongoing debate on the benefits, or potential benefits, and also the harms—or potential harms—of artificial intelligence and, more broadly, new emerging technologies, is well-known.^[1] When I first heard the term “new emerging technologies,” I thought: “What a strange expression! ‘Emerging’ already implies novelty, so why do we need to indicate the same thing twice?” While I did not delve into this linguistic nuance at the time, I now perceive it as one among the myriad intriguing questions in a new world we have stepped into without even noticing, a world in which things are unclear or not clear enough, but where the evolving context demands instant answers and projections.

A world with many unknowns, words with many unknowns. What can we expect from ourselves and each other if we cannot even agree on definitions? For instance, what exactly do we mean by “artificial intelligence?”^[2] Similarly, curricula are rushing to adjust to the demands of the “real world,” although one may question where the classroom ends and the “real world” begins, to what extent this “real world” is real, and even what is real in the first place.

Imagination plays a key role in understanding technological progress.^[3] It is also required for looking ahead: while we aim to catch up with the ongoing change, be it for the purposes of education or regulation, there is a sense of always lagging behind. In our quest to foresee what is coming, creative

abilities, whether (scientific) imagination or intuition, are crucial. Can they be taught? Should we, for instance, open book clubs at local libraries and cultural centres, especially in disadvantaged areas, and have them discuss, say, Asimov's books?^[4]

The technological divide is all too real: a new asymmetry of power, requiring action to bridge the widening gap. Knowledge and skills, starting from basic digital literacy, factor into power structures.^[5] Vulnerable and marginalised communities, be it the elderly, disabled, long-term unemployed, refugees, or low-income groups, are just as likely to become victims of rapid technological development as they are to be empowered by new emerging technologies.^[6] On both the individual and community levels, the changing context may make a positive difference but also worsen the situation.^[7]

In addition, there may be both “objective” and “subjective” parts to the outcome, with the latter reflecting how one “feels” about the new technologies. Hearing that algorithms discriminate against people like oneself may make one “feel” the victim of technological progress. Another, softer response may consist of feeling “lost”: not so much a victim but a person who cannot keep up with the change. Maybe all this is not for me but for other people, maybe I am too old, not as smart as *them*. This new world may seem fascinating, but I have no clue how to navigate it.

Such “feelings” are as important to note as researched, investigated cases of bias and discrimination. This is also part of the discussion on the imaginary versus the real, the line to be drawn between both, or the absence of such a line: from the “outside,” something intangible, such as a feeling, may be easily disregarded as subjective, as opposed to the “real”—but for the person who feels it, it is very much real.

Being a victim, or feeling like one, are unpleasant modalities. They relate to a lack of agency and action, and in this way, feeling lost is not that different. As described by Würtz Jensen and Thunberg, the concept of victimhood “refers to how individuals construct their own ways of being a victim.”^[8] This negative isolating distinction is common to both the “real” and “imaginary” dimensions

isolating distinction is common to both the real and imaginary dimensions of victimhood: it separates the “self” from the “others.”

Asymmetry of Information, Power Structures, and Agency

Victimhood, whether related to real-life circumstances or to “feelings,” implies opposing an active position. Feeling oneself a victim long enough may create a microcosm of non-action, as any action may lead to further disappointment. Not everything works out every time, but if one is already in a disadvantaged position, or perceives oneself as such, every “no” may have a different effect than it would in a “normal” situation. Victimhood invites an interpretation of the world as working “against” oneself. Every new disappointment—as inevitable as it may be in human life—is associated with the victim position, leading to further reinforcement of victimhood.

In the case of rapid technological developments, real or perceived powerlessness—a lack of agency supplemented by the asymmetry of information, that is, a lack of knowledge and skills on the one hand and the inability to distinguish what we may call “true” expertise among a conflicting noise of claims on the other—leads to a passive, non-productive stance.

The presence of broader societal anxieties that exceed the community level creates a situation where algorithms, these entities of the unknown, may be perceived as perpetrators or potential perpetrators. As with other sensitive issues, one may see pre-existing factors as confirmation of unjust treatment, now by algorithms. For example, a situation of not getting a job may be seen as “predetermined,” in which the applicant had no power or agency since the outcome may have been orchestrated by a powerful algorithm deciding one’s, in this context professional, fate, and by this algorithm working against you.

The situation is exacerbated by public scandals involving algorithmic decision making, such as the child benefit scandal in the Netherlands, which shed light on real-life cases of injustice and also allow one to presume the possibility of other cases.^[9] Oftentimes, it is hard to glean from the news coverage to what extent the algorithms are to blame, and to what extent the people involved in

the processes.

What can be done in such cases? The real or perceived lack of power may be confronted through empowerment activities, yet there are several hurdles. First, in practice, it is not easy to design a framework for education, collaboration, and co-creation in the technological context, there are no ready recipes, and current attempts uneasily combine an experimental character with the responsibility of engaging vulnerable groups. Second, empowerment actions are directed from the “outside” and hence imply the agency of those initiating the action in question.

A situation in which the resources—knowledge and skills—are re-allocated in a way so that a previously disadvantaged person or group is simply “given” these resources reinforces the feeling of “outside” agency. While an individual or group may thus be enabled to improve their position, this is not in consequence of taking action, overcoming passivity and victimhood.

The condition of victimhood, while being “internal,” is built on one’s (negative) experiences or perceptions of the “outside” world. And since passivity reinforces victimhood, the design of any empowerment project needs to encourage action on the “receiving” side. The overcoming of barriers, be it the digital divide or an individual sense of being lost and unequipped in the AI age, must be an individual or collective achievement and result from an action that changes the power distribution.

Reclaiming responsibility for one’s life and fate, taking action resulting in change, may be considered as “antidotes” to victimhood. How to encourage a reclaiming of agency with the initiative coming from a person or group, how to facilitate good conditions and build an “infrastructure” but avoid any imposition, direction, or other actions that undermine agency, is an open question.

The Co-creation Dilemma

A useful tool in this context is co-creation. The practice is not new: according to

A useful tool in this context is co-creation. The practice is not new. according to Ruess et al., “since the 1960s, policymakers and social scientists alike have framed public engagement as a new governance mode that invites different actors to consult about technology-grounded matters and thus a central element in decision-making processes in modern democracies.”^[10] Curiously, though, decades of practice have not resulted in tried-and-tested formulas, and co-creation efforts still take the form of experiments, especially in light of peculiarities regarding context and subject matter, such as the specificity of the technological domain and/or the particular vulnerabilities of the “target” community.^[11]

The co-creation of technological solutions with vulnerable communities, if well facilitated, can be an active experience of gaining and exercising agency in the technological world.^[12] One may speak of being given a voice,^[13] but this phrase only reinforces the passive condition mentioned above, of being on the “receiving” side, of lacking agency. “Gaining a voice” is different: it is an active process directed towards overcoming victimhood.

For a technological provider, whether a start-up or a well-established company, co-creation can serve as an instrument of bias mitigation.^[14] When you develop your algorithm-based product in tandem with “real” people coming from a “non-mainstream” background, people who authentically participate in the creative process, you mitigate bias by design and enable a joint effort toward algorithmic fairness.^[15]

Co-creation is not easy, and although multiple works have been dedicated to the issue, challenges abound. Whether they concern the initial design of the process or the experimental character of the whole endeavour, these challenges may easily leave participants disappointed. The lion’s share of responsibility lies with technological providers “translating” their knowledge and know-how for the members of “target” groups so these are enabled and motivated to participate in the process. Without assigning value to such processes, the technological providers’ involvement may stay formal, uninspired, and hence unsuccessful.

The attempt to co-create technological solutions with people who may see themselves as victims or potential victims of algorithms must address asymmetries of power built on information, knowledge, and skills. Equalizing the initially unequal conditions is not easy, but it may be facilitated by an understanding that the required “translation” needs to go in both directions: not just from technological providers to people but also from people to technological providers. This mutuality helps to reimagine the whole context, not as one of technological knowledge-haves and knowledge-have-nots but as one of many unknowns, with as much to be learned by the “experts” as by the “people.” Such a process has the potential to interlink technological and social innovations with the promise of mutual reward,^[16] and, more fundamentally, to overcome the silos of the technological and social domains.

Bubbles and Frauds

To some, artificial intelligence is a bubble,^[17] and when bubbles burst, it is usually loud and damaging. However, this view is complicated by the fact that AI is not simply one thing—as attested to by the lack of a satisfactory definition. Rather, AI is a framework encompassing many algorithm-based “things,” and indeed, some of the projects, people, and even research directions within this framework may turn out to be bubbles.

This ongoing, evolving, not-fully-clear or mostly unclear nature of the processes may actually be of use. The potential presence of what we may simply call “frauds” in the area of technological development can actually help to overcome the asymmetry of information, knowledge, and skills, and to confront victimhood. We simply need to acknowledge that there are many things we are still figuring out, whether in computing or in regulatory, ethical, and societal matters, and that further questions will arise. There are, and will be, wrong decisions, controversial approaches, and opportunistic people who present themselves as experts but lack insight, goodwill, and other merits. The authority and agency of experts, already widely discussed in the context of COVID pandemic, is overrated in AI and the broader debate about emerging technologies.

We Cannot Explain Everything, But We Are Trying

In turbulent times, whether of a global pandemic or of rapid technological development, vulnerable groups find themselves in an even more vulnerable position.^[18] Asymmetries of power and information, a lack of agency—worrisome elements are only exacerbated by the challenging context.^[19]

In the case of rapid technological development, the debates contain a lot of noise, and not all “experts” are equally trustworthy. With high societal anxiety about vulnerability and victimhood vis-à-vis algorithms, people naturally tend to look for a “saviour,” someone who promises to steer them through the unknown waters. The problem is that such an easy solution may well be the most fraudulent one. Attempts to oversimplify the complex reality and present unknowns as knowns may lead to bubbles, which, once burst, foment even more distrust in experts.

The best way to confront this scenario is to admit and continuously emphasise that there are many questions with unknown answers, many points that are debatable—in the computational, subfield-specific, legal, societal, and ethical domains. We do not know, we do not agree, and many things we attempt (for instance, the AI Act of the European Union) may be wrong at their core, in the underlying philosophy determining the action. Such an approach would not only be honest but also help us bridge the divide of “victims” and “perpetrators” by exposing our shared vulnerability. We are all vulnerable before the unknown, and our task is to co-create knowledge and solutions as we confront more and more unknowns.

KRISTINA KHUTSISHVILI is a postdoctoral researcher and published fiction author currently based in Amsterdam. She holds an interdisciplinary doctoral degree from Scuola Superiore Sant’Anna, Pisa. The above essay derives, to a significant extent, from her experience of leading the work package “Ethical and Inclusive Engagement in Practice” of the CommuniCity Horizon Europe project.^[20]

kristina.khutsishvili@alumni.sssup.it

dePICTions volume 4 (2024): *Victimhood*

[1] See, for example, Daron Acemoglu, “Harms of AI,” *NBER Working Papers* 29247, National Bureau of Economic Research, September 2021 [15 February 2024].

[2] A legal definition of an AI system, with a reference to EU law, can now be derived from the EU Artificial Intelligence Act, which adopts the OECD definition: “a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments” [1 March 2024].

[3] Wu Hsuan-Yi, and Vic Callaghan, “From Imagination to Innovation: A Creative Development Process,” in *Intelligent Environments 2016: Workshop Proceedings of the 12th International Conference on Intelligent Environments*, edited by Paulo Novais and Shin’ichi Konomi, Amsterdam: IOS Press, 2016, 514–523; Phil Turner, *Imagination+Technology*, Cham: Springer, 2020.

[4] See, for example, Lucy Gill-Simmen, “In the Age of AI, Teach Your Students How to be Human,” *Times Higher Education*, 9 January 2024 [2 March 2024]; Kelley Cotter and Bianca Reisdorf, “Algorithmic Knowledge Gaps: A New Horizon of (Digital) Inequality,” *International Journal of Communication* 14 (2020): 745-765 [3 March 2024].

[5] Jeremy Moss, “Power and the Digital Divide,” *Ethics and Information Technology* 4.2 (2002): 159–165.

[6] Kristina Khutsishvili, “Guidelines for Translating Frameworks, Methods, Tools and Principles of Local Innovations for Marginalised and Vulnerable Communities – 2023,” *Open Research Europe*, 22 February 2024 [2 March 2024].

[7] Kristalina Georgieva, “AI Will Transform Economy. Let’s Make Sure it Benefits Humanity,” *IMF Blog*, 14 January 2024 [3 March 2024].

[8] Julie Würtz Jensen and Sara Thunberg, “Navigating Professionals’ Conditions for Co-Production of Victim Support: A Conceptual Article,” *International Review of Victimology* 30.2(May 2024): 401-416 [14 January 2024].

[9] “The Dutch Child Benefit Scandal, Institutional Racism and Algorithms,” Parliamentary Question, European Parliament, 28 June 2022 [14 January 2024]; “Targeted Support Offered to Victims of the Childcare Benefits Scandal with Children Taken into Care,” Government of the Netherlands, 18 September 2023 [14 January 2024].

[10] Anja Ruess, Ruth Müller, and Sebastian Pfotenhauer, “Opportunity or Responsibility? Tracing Co-creation in the European Policy Discourse,” *Science and Public Policy* 50.3 (June 2023): 433–444 [4 March 2024].

[11] Co-creation helps to address the implicit bias problem (prejudices, negative stereotypes) also in healthcare. See, for example, Connie Yang et al., “Imagining Improved Interactions: Patients’ Designs To Address Implicit Bias,” *AMIA Annual Symposium Proceedings* (2023): 774-783 [20 April 2024].

[12] Kristina Khutsishvili, Neeltje Pavicic, and Machteld Combé, “The Challenge of Co-creation: How to Connect Technologies and Communities in an Ethical Way,” *Smart Ethics in the Digital World: Proceedings of the ETHICOMP 2024*, edited by Mario Arias Oliva et al., La Rioja: Universidad de La Rioja, 2024, 259-261 [4 March 2024].

[13] An interesting account on “giving voice” in the context of social representations research is supplied by Sophie Zadeh in “The Implications of Dialogicality for ‘Giving Voice’ in Social Representations Research,” *Journal for the Theory of Social Behaviour* 47.3 (2017): 263–278 [4 March 2024].

[14] In the EU report, “Bias in Algorithms: Artificial Intelligence and Discrimination” (Vienna European Union Agency for Fundamental Rights 2022) bias is defined as

(Vienna, European Union Agency for Fundamental Rights, 2022), bias is defined as “the tendency for algorithms to produce outputs that lead to a disadvantage for certain groups, such as women, ethnic minorities or people with a disability” and explained as being “pervasive in society, rooted in psychological, social and cultural dynamics, and hence reflected in the data and texts that are used for developing AI models” (17) [5 March 2024].

[15] As the concepts of fairness and justice exceed the scope of studies on artificial intelligence and machine learning, multiple angles can be relevant. The emotional component is highly relevant here: what is “fair” for one agent, not only on the level of definition, but in its essence, may not be viewed as “fair” by another. In “Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies,” Emilio Ferrara gives an overview on algorithmic fairness, defining it as “the absence of bias or discrimination in AI systems, which can be challenging to achieve due to the different types of bias that can arise in these system” (9) [5 March 2024].

[16] The scope of literature exploring connections between technological and social innovation is unfortunately limited. See, for example, Michael Kohlgrüber, Karina Maldonado-Mariscal, and Antonius Schröder, “Mutual Learning in Innovation and Co-creation Processes: Integrating Technological and Social Innovation,” *Frontiers in Education* 6 (2021) [5 March 2024].

[17] Christiaan Hetzner, “As Nvidia hits \$2 trillion, billionaire Marc Rowan’s asset manager Apollo calls AI a ‘bubble’ worse than even the dotcom era,” *Fortune*, 26 February 2024 [5 March 2024].

[18] Richard Gray, “How Vulnerable Groups were Left Behind in Pandemic Response,” *Horizon: The EU Research and Innovation Magazine*, 7 June 2021 [9 January 2024].

[19] Gavin Clarkson, Trond Jacobsen, and Archer Batcheller. “Information Asymmetry and Information Sharing,” *Government Information Quarterly* 24.4 (2007): 827–839.

[20] The author’s work has been supported by the European Union’s Horizon Europe research and innovation programme under grant agreement No 101070325

(Innovative Solutions Responding to the Needs of Cities & Communities – CommuniCity). See the [CommuniCity Project website](#).

Categories: [depictions iv](#)

Responses

You must be [logged in](#) to post a comment.

Recent Posts

[PICT Launches Senlis Writers' Retreat](#)

[Maria Matalaev Translates Daughter of the Revolution](#)

[Mortaza Behboudi Releases Book on Iran Protests](#)

[PICT Book Receives Local Editions in Bangladesh](#)

[PICT Book Tops Bestseller Lists, Receives Media Coverage](#)

Subscribe to the Paris Institute Newsletter

First name or full name

Email

Errors

☐ By continuing, you accept the

Subscribe

About PICT

Contact

Login

© 2023 Paris Institute For Critical
Thinking. All rights Reserved.

FR