The indirect rebound effects of AI as undone science: philosophical reflection on two structural causes

Damien Lacroux	Aurélie Bugeau	Anne-Laure Ligozat
Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, F-33400 Talence, France	Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, IUF F-33400 Talence, France	Univ. Paris-Saclay, CNRS, ENSIIE, LISN, Orsay, France

Biography : The first author is a philosopher of science specializing in the ethical and epistemological impacts of AI on the medical field. The second and third authors are computer scientists specializing in the study of the environmental impacts of AI.

Abstract : This contribution shows that ignorance of the indirect and structural rebound effects of AI is less a matter of a lack of knowledge than of institutionalised ignorance, produced by dominant epistemological and institutional frameworks. By drawing on Collingridge's dilemma and the notion of projective responsibility, it proposes a new way of integrating these effects into a digital ethic that is attentive to uncertainty and long timeframes.

The deployment of artificial intelligence (AI) in all sectors of society raises a crucial question that is largely ignored in measurement tools and institutional priorities: that of the indirect and structural rebound effects induced by these technologies. As shown by Gossart (2015) and Horner et al. (2016), environmental assessments of digital technologies tend to neglect indirect rebound effects such as increased demand or infrastructure reorganisation in favour of measurable and immediate criteria. However, these indirect effects also account for a significant part of the overall footprint of digital innovations (Freire-González 2017; Eckchajzer 2025; Bieser & Hilty 2018). We hypothesise that this absence constitutes an emblematic case of *undone science*, not because these effects are negligible, but because they challenge dominant interests and established epistemological frameworks (Bieser & Hilty 2018; Flipo 2021). Our contribution does not merely point out this deficit: it examines two structural causes in order to reveal the theoretical and methodological assumptions that explain the exclusion of certain indirect and structural rebound effects from the debates. We will first show how Collingridge's dilemma helps us understand the structural difficulty of anticipating, reacting to, and measuring the indirect and structural rebound effects of AI. We will then propose reformulating technological vigilance in terms of projective responsibility, placing it within an ethics by design framework that is attentive to uncertainty and long-term temporality. Finally, we will highlight the institutional and discursive mechanisms that transform this ignorance into genuine institutional ignorance, helping to keep these rebound effects outside the legitimate field of research. We will illustrate our point with examples of the rebound effects of AI in the healthcare sector.

1. Rethinking rebound effects in light of the Collingridge dilemma

This difficulty can be explained, at least in part, by Collingridge's dilemma (Collingridge 1980). Formulated in the early 1980s in a work that has become a classic in the philosophy of technology, it describes a recurring paradox in the governance of innovation. It consists of two parts. 1) Upstream: when a technology is being designed, its actual uses and major consequences are invisible. Without empirical hindsight, it is difficult to anticipate its impacts, even though this would be the moment when there is still room for manoeuvre to guide or correct the technology (Collingridge 1980). 2) Downstream: once the technology has been deployed, its uses stabilise, infrastructure becomes established, investment accumulates and practices adapt. The societal and organisational effects then become visible, but it is virtually impossible to turn back the clock because of the strong dependence on the initial choices (Collingridge 1980). In summary: while there is still time to act, the effects are not yet visible; and when they do become visible, it is too late to make any significant changes.

Considering the impact of AI in healthcare in light of Collingridge's dilemma highlights underestimated structural problems. Upstream, rebound effects are difficult to predict because they result from distributed socio-technical dynamics. For example, the spread of medical AI stimulates the growth of imaging centres, storage infrastructures and image analysis companies. These systemic transformations are impossible to anticipate precisely at the initial design stage because they emerge from complex socio-technical dynamics (Börjesson et al. 2014). Downstream, once the technology has been deployed, these rebound effects remain difficult to measure. Some studies also point out that rebound effects, which are distributed among different actors and delayed over time, cannot be captured by standardised LCA methodologies (Coroamă et al. 2020). They are (i) dispersed over time, sometimes manifesting years later through changes in clinical practices; (ii) distributed among multiple actors (patients, hospitals, cloud operators, local authorities); and (iii) amplified by adaptive feedback loops, where changes in protocol alter the data produced and therefore the performance of the models. In practice, no authority has an overview capable of tracking these externalities.

While Collingridge's "classic" dilemma invites us to reflect on the complex balance between anticipation and feedback, the rebound effects associated with AI in healthcare shift this framework and reveal its limitations. In their indirect and structural dimensions, these effects simultaneously escape anticipation and feedback (Börjesson et al. 2014; Coroamă et al. 2020). The binary anticipation/feedback framework that is supposed to guide technological vigilance is itself thrown into crisis by the non-linear dynamics of rebound effects. This double impossibility directly contributes to making rebound effects a paradigmatic case of undone science. These effects remain largely understudied because they elude dominant measurement paradigms (Garin, 2024), fall outside institutional priorities such as clinical performance or energy efficiency (Morand et al., 2025), and require complex, multi-stakeholder coordination to study (Lenoir & Parker, 2025; Mhlanga, 2025). Undone science is therefore not the result of simple oversight or empirical deficit, but of structural constraints; rebound effects, due to their systemic and delayed nature, are difficult to access using both dominant institutional methods and logic. For example, some AI systems produce or cross-reference information that was initially considered secondary (from images, sensors or patient records), but which, thanks to innovations or new clinical recommendations, may subsequently acquire the status of "medical data" in its own right. This reclassification then leads to a regulatory obligation for long-term storage. This mechanism generates an inflation in the volumes of data to be stored, implying an unforeseen expansion of data centre capacities and, in turn, an increase in energy consumption and hardware requirements (servers, networks, cooling). Such a structural rebound is virtually impossible to anticipate during the initial design of AI, as it results from distributed and time-delayed dynamics, combining technological developments, institutional changes and legal transformations.

2. Technological vigilance and projective responsibility as ethics by design

This requires us to move beyond a sequential conception of vigilance, based on the classic "predict and correct" model. Drawing inspiration from Hans Jonas's concept of extended responsibility, it becomes necessary to integrate long-term temporality and radical uncertainty into technological governance. Indeed, the knock-on effects of AI in healthcare concern not only the current generation of patients and practitioners, but also future infrastructure, the sustainability of hospital systems and the ecological resources on which future generations will depend (Jonas 1990). From this perspective, vigilance cannot be reduced to a logic of control, but must be understood as a projective responsibility that combines precaution, scenario planning and continuous adjustment. However, as Isabelle Stengers points out, uncertainty is not just a lack of knowledge but a constitutive condition of collective action: it calls for reflexive and participatory mechanisms capable of dealing with what we cannot control (Stengers 1997, 2009). Re-reading Collingridge's dilemma through the lens of rebound effects therefore shifts the question from technological control to the collective capacity to adapt, detect weak signals and reconfigure choices over time. Vigilance here becomes a systemic and distributed process, where anticipation and feedback are no longer two distinct moments but two dimensions constantly intertwined in adaptive loops of the same ethics by design (Nurock et al. 2021). A further step is to refine our categories of analysis by clearly distinguishing between expected effects (e.g. an increase in the number of scans prescribed or the possibility of treating more patients thanks to greater efficiency) and rebound effects proper (e.g., the massive reallocation of hospital resources, the explosion of data flows leading to new energy dependencies, or the acceleration of logistics cycles linked to digital infrastructures). As Morand *et al.* 2025 suggest, this distinction would already make it possible to construct more realistic prospective scenarios and to anchor anticipation in plausible horizons rather than in abstract, unrepresentative speculations.

3. From involuntary ignorance to institutionalised ignorance

Finally, another structural explanation for the invisibility of rebound effects lies in the configuration of actors and their dominant discourses. Some doctors and clinical communities are interested in the environmental impacts of AI systems and promote eco-design approaches or carbon footprint calculations (Ueda et al. 2024; Richie 2022; Richie et al. 2025; Alshqaqeeq et al 2020; Wahl et al 2018; Thiel, Richie 2022). However, their analyses are most often limited to direct and measurable dimensions (energy consumption, equipment life cycle, efficiency per action), leaving aside rebound, indirect or structural effects. This shortcoming can undoubtedly be explained by a lack of knowledge and methodological tools, but also by a proximity to industrial legitimisation discourses that tend to naturalise the idea of "green" digital technology. These causes are based on a qualitative survey: a dozen semi-structured interviews conducted with stakeholders from four spheres (regulation, technological innovation, communication, expertise and counter-expertise), selected for the diversity of their positions in the controversy. The interviews followed a common framework focused on rebound effects, measurement regimes, and political and ethical trade-offs related to the deployment of AI in healthcare.

On the part of manufacturers and tech giants, the issue of rebound effects is often minimised or excluded from assessments because they are too complex to quantify and exceed the usual reporting and standardisation capabilities (Garin 2024). This configuration produces a systemic blind spot: rebound effects remain largely ignored because they lie at the intersection of different disciplines and interests. Lack of knowledge and disciplinary compartmentalisation marginalise the analysis of rebound effects in clinical studies, while methodological complexity and industry competitiveness constraints limit their consideration. Combined with a funding policy that favours clinical performance and technological innovation, this combination is a classic example *of* undone science*,* where certain critical issues remain structurally neglected.

The concept *of* undone science is therefore not limited to a simple lack of data or empirical studies, but reflects implicit normative orientations that structure scientific and technological practices. Raising the question of the indirect and structural rebound effects of AI in healthcare means questioning not only what is measured, but also for whom these measurements matter and whose interests they serve. The absence of research on these dimensions is therefore not neutral: it reveals how evaluation systems themselves create blind spots. Thus, in the health sector, for example, by focusing the value of AI on diagnostic accuracy and operational efficiency, these frameworks disregard the systemic transformations (energetic, organisational or institutional) that they themselves help to bring about. Thus, institutionalised ignorance does not refer to a simple lack of knowledge, but to the active creation of grey areas within the very systems of scientific evidence and legitimisation.

Bibliography:

- **Alshqaqeeq, F., Amin Esmaeili, M., Overcash, M., & Twomey, J.** (2020). Quantifying hospital services by carbon footprint: A systematic literature review of patient care alternatives. *Resources, Conservation and Recycling, 154*, Article 104560.
- **Bieser, J. C. T., & Hilty, L. M.** (2018). Assessing indirect environmental effects of information and communication technology (ICT): A systematic literature review. *Sustainability, 10*(8), 2662.
- **Börjesson Rivera, M., Håkansson, C., Svenfelt, Å., & Finnveden, G.** (2014). Including second order effects in environmental assessments of ICT. *Environmental Modelling & Software, 56*, 105–115.
- **Collingridge, D.** (1980). *The social control of technology*. Pinter.
- **Coroamă, V., Bergmark, P., Höjer, M., & Malmodin, J.** (2020). A methodology for assessing the environmental effects induced by ICT services. *Proceedings of the 7th International Conference on ICT for Sustainability*.
- **Ekchajzer, D., Gossart, C., Tezenas du Montcel, B., & Combaz, J.** (2025). Taking into account the systemic environmental effects of digital technology in organisations: a dual logical and organisational challenge. Terminal, (140)
- **Flipo, F.** (2021). The imperative of digital sobriety. *Cahiers Droit, Sciences & Technologies, 13*, 29–47.

- **Freire-González, J.** (2017). A new way to estimate the direct and indirect rebound effect and other rebound indicators. Energy, 128(C), 394-402
- **Garin, M.** (2024). *From federated learning to a critical theory of heteromatic learning* [Doctoral thesis, Paris-Saclay University].
- **Gossart, C.** (2015). Rebound effects and ICT: A review of the literature. In L. M. Hilty & B. Aebischer (Eds.), *ICT innovations for sustainability* (pp. 435–448). Springer.
- **Horner, N., Shehabi, A., & Azevedo, I.** (2016). Known unknowns: Indirect energy effects of information and communication technology. *Environmental Research Letters, 11*(10).
- **Jonas, H.** (1990). *The Imperative of Responsibility: An Ethics for Technological Civilisation*. Éditions du Cerf.
- **Lamnatou, C., Cristofari, C., & Chemisana, D.** (2024). Artificial intelligence (AI) in relation to environmental life-cycle assessment, photovoltaics, smart grids and small-island economies. *Sustainable Energy Technologies and Assessments, 71*, Article 104005.
- **Lenoir, T., & Parker, C.** (2025). Statactivism up and down the stack: Dilemmas in the estimation of AI's environmental footprint. *Digital Society, 4*, Article 47.
- **Ligozat, A.-L., Lefèvre, J., Bugeau, A., & Combaz, J.** (2022). Unravelling the hidden environmental impacts of AI solutions for the environment: Life cycle assessment of AI solutions. *Sustainability, 14*(9), 5172.
- **Mhlanga, D.** (2025). AI beyond efficiency: Navigating the rebound effect in AI-driven sustainable development. *Frontiers in Energy Research*.
- **Morand, C., Ligozat, A.-L., & Névéol, A.** (2025). Characterising goals and impacts of digitalisation: The case of promises in French healthcare policies. *Preprint*.
- **Nurock, V., Chatila, R., & Parizeau, M.-H.** (2021). What does "ethical by design" mean? In B. Braunschweig & M. Ghallab (Eds.), *Reflections on artificial intelligence for humanity* (pp. 171–190). Springer.
- **Richie, C.** (2022). Environmentally sustainable development and use of artificial intelligence in healthcare. *Bioethics, 36*(5), 547–555.
- **Richie, C. S., Hinrichs-Krapels, S., Dobbe, R., French, P. J., Wei, J. C. J., Diehl, J. C., & Kong, R.** (2025). Environmental impacts of artificial intelligence in health care: Considerations and recommendations. *Health and Technology*.
- **Stengers, I.** (1997). *Cosmopolitics* (Volumes 1–7). La Découverte / Les Empêcheurs de penser en rond.
- **Stengers, I.** (2009). *Au temps des catastrophes: Résister à la barbarie qui vient*. La Découverte.
- **Thiel, C., & Richie, C.** (2022). Carbon emissions from overuse of U.S. health care: Medical and ethical problems. *The Hastings Centre Report, 52*(4), 10–16.
- **Ueda, D., Walston, S. L., Fujita, S., Fushimi, Y., Tsuboyama, T., Kamagata, K., Yamada, A., Yanagawa, M., Ito, R., Fujima, N., Kawamura, M., Nakaura, T., Matsui, Y., Tatsugami, F., Fujioka, T., Nozaki, T., Hirata, K., & Naganawa, S.** (2024). Climate change and artificial intelligence in healthcare: Review and recommendations towards a sustainable future. *Diagnostic and Interventional Imaging, 105(*11)
- **Wahl, B., Cossy-Gantner, A., Germann, S., & Schwalbe, N. R.** (2018). Artificial intelligence and global health: How can AI contribute to health in resource-poor settings? *BMJ Global Health, 3*(4).
- **Wright, D., Igel, C., Samuel, G., & Selvan, R.** (2023). Efficiency is not enough: A critical perspective of environmentally sustainable AI. *arXiv*.

Acknowledgements :

This research was conducted with the support of the Observatory on AI's Environmental Footprint led by the AI and Society Institute at ENS.