AI and environmental sustainability: how to govern an ambivalent relationship

FEDERICA LUCIVERO

Abstract

While AITs hold promise in optimizing supply chains, circular economies, and renewable energy, they also contribute to significant environmental costs, often overlooked in the policy debate. The concept of "digital pollution" emphasizes the physical and ecological impacts of AI infrastructures, data storage, resource consumption, and toxic emissions.

The chapter underscores the limitations of conventional cost-benefit analyses in assessing AI's environmental effects and calls for a more value-based and political approach. It emphasizes the need to transparently evaluate who benefits from or is harmed by AI's environmental impact and allocate responsibilities accordingly. It concludes that while AI can contribute to sustainable development, its environmental costs and impacts must be addressed. It argues for a global governance project that considers a multitude of values, stakeholders, and regulatory mechanisms to ensure the sustainability of AI technologies, in contrast to the current fragmented and often market-driven approaches. Critical reflections on these issues are essential for guiding this global governance project effectively.

Introduction

AI enabled technologies are often presented as tools to address global societal challenges, including the climate crisis. Machine learning enhances the capacity of discovering patterns and predicting outcomes that can be used to optimize supply chains or circular economies or to make renewable energy more efficient. What gets lost in these visions of "AI for Good", "AI for Sustainability" is that AI is not simply a technical solution, but it is also part of the problem: storing large amounts of data, training models, running complex software, manufacturing and disposing hardware have significant environmental and social costs. This dual role of AI, as solution as well as problem for environmental sustainability is nearly absent in documents on AI policy and regulation. As this chapter will show and argue, the invisibility of the environmental sustainability issue in the policy and regulatory context is a problem: first of all because environmental costs always have a social and ethical dimension and secondly because ignoring such costs in debates of AI regulation creates a responsibility gap.

To develop this argument, the chapter starts with an overview of the relationship between AI and sustainability. It will then introduce the concept of digital pollution, through examples that draw attention to the physical configuration of the infrastructures that enable AITs, the space occupied for data storage, the natural resources used for hardware manufacturing and energy production, and the toxic emissions discharged in the air and soil in informal recycling sites for electronic waste. These examples will lead over to a discussion of two sets of issues that are crucial for governing AITs in more sustainable ways, namely: the limits of cost/benefit analysis in this area and the allocation of responsibility for the harmful effects of digital pollution. As I will argue, the search for solutions to the digital pollution of AI is not only a technological issue, but also a political and value-based one. It requires articulation of the plurality of values involved in decisions, a transparent evaluation of who benefits from and who is harmed by AIT development in environmental terms, and consideration of who has the responsibility to intervene[i].

Environmental sustainability, sustainable development and AI

In modern parlance, when we speak of "sustainable" practices, we refer to their ability to be sustained, or maintained over time. Usually the term has a positive connotation, but in its literal meaning it could also refer to practices, such as slavery, that are sustainable over time but morally problematic. In the context of environmental movements and ethics, the concept of sustainability refers to behavioral practices that last over time because they respect ecological limits, such as to bring benefits to the present and future generations (Attfield 2013). In sustainable fishing and agriculture, for example, fish or land resources are not treated as mere products to be exploited, but as ecosystems made to thrive. Although sustainable practices are considered as virtuous examples from an environmental point of view, it should be emphasized that the term is increasingly used in a generic way to refer to "good" commercial policies, practices or products, often ending up being used improperly and loosely.

The concept of sustainability is often associated with that of "development". In the definition that can be read in the Brundtland Report, published in 1987 by the World Commission on Environment and Development (WCED), sustainable development is "development that meets the needs of the present without compromising the possibility that future generations meet their own" (WCED, 1987). Related to this definition of sustainable development is that of "needs", represented in the report as a three-dimensional space which contemplates needs of an economic, social and environmental nature. Policies aimed at meeting these needs must therefore address not only problems of poverty and malnutrition – for example, through sustainable farming and fishing practices or through the promotion of renewable energies – but also issues of gender equality, access to primary education and basic health care, essential to ensure the well-being of individuals in the present and in the future. The Millenium Sustainable Goals (MSG, 2000-2015) and the Sustainable Development Goals (SDG, 2015-2030) are objectives set by the United Nations so that international policies move towards the realization of an ideal sustainable development.

It is important to note that the definition of sustainable development is grounded on a very broad concept of justice, which promotes a fair distribution of resources and wellbeing not only within the same society but also between different populations, species and ecosystems globally. A definition which, moreover, does not stop at assigning a certain moral status to individuals, animals and ecological systems in the present, but also considers the responsibilities of the present generation to ensure conditions of livability and prosperity for their descendants. This widening of the scope of sustainable practices (in space, time and dimensions) inevitably creates moral conflicts: for example, guaranteeing energy access to even the most vulnerable groups in poorer countries could imply an increase in greenhouse gas emissions globally and consequently harm future generations. On the other hand, there may be economists or politicians who believe that the responsibility for ensuring the wellbeing of the countries they represent (and the vulnerable groups within them) is more important than worrying about the wellbeing of other populations or even future generations. Although the concept of sustainable development suggests that the wellbeing of future and present generations, ecosystems and populations on a global scale are interconnected and should be promoted in a harmonized way, the delivery of this vision is not only practically difficult, but also raises ethical questions.

Sustainability and AI

AIT have the potential to contribute to sustainable development in general and environmental sustainability in particular. Sensors for environmental data collection, machine learning algorithms for processing energy consumptions, models to support decisions to optimize cycles of consumption and waste recycling are all examples of how AIT can help us address environmental concerns and "save the planet" (Dauvergne 2020). In parallel to these visions there has been a growing academic interest in the relationship between AI and sustainability in the context of AI governance. An exploratory search on Scopus for keywords such as AI, sustainability and governance[ii] shows a steep increase in publications: the hits in 2022 have tripled compared to 2020 and grown even by the factor ten when compared to 2019. The majority of these contributions explore the ways in which artificial intelligence offers solutions to problems related to climate change, biodiversity and environmental degradation. However, increasingly scholars have pointed out the need for a critical understanding of the relationship between AI and sustainable development goals (Jobin 2019, Dauvergne 2020, Mazza and Floridi 2022, Coeckelbergh 2021). As van Wynsberghe puts it, there is a broad corpus of resources discussing "AI for sustainability" (where AIT are seen as tools to achieve the goals of sustainable development), with a comparatively limited amount of contributions that address the issue of the "sustainability of AI" (showing how AI exacerbates unsustainable practices).

While the societal and ethical risks of AI have been largely discussed, the environmental "dark side" of AI has been only recently acknowledged in policy and regulatory contexts, and often in a very timid way. The draft of the European AI Act, for example, originally only mentioned environmental sustainability as a value or principle innovators are encouraged to thrive towards. One of the amendments adopted by the European Parliament in June 2023 more explicitly refer to the need to establish recommendations or even targets for sustainability to enable a comparison between AI environmental efficiency and resource concerns (Amendment 83. Proposal for a regulation
Recital 46 b (new))[iii] (although in the mandate document from 2023 the environmental sustainability becomes more visible and more of a requirement). In June 2019, the Chinese National Governance Professional Committee on New Generation AI released the "Governance Principles for the New Generation Artificial Intelligence–Developing Responsible Artificial Intelligence". Although this document states that the principles aim to better 'promote sustainable development of economy, society and ecology', they are very abstract and come without a specific recommendation in this direction[iv]. In order to understand why the invisibility of environmental sustainability issues in the AI policy and regulatory context is a problem, to the next section articulates AIT's detrimental impacts for the environment.

Digital Pollution: What is it?

The terminology used to refer to AIT often refers to speed, to fluidity and the presumed immateriality of the network and IT systems underpinning AI solutions. The metaphor of the "cloud" – the ethereal and limitless cloud, which can be extended according to customer needs for a few Euros per month, and where data can be stored and made accessible from any location or networked device – is indicative of this misleading perspective on the digital as something immaterial. In reality, these systems require real industrial infrastructures which extract minerals, occupy land, alter urban and suburban landscapes, consume energy, release greenhouse gases into the atmosphere, end their life cycle in landfills.

The term "digital pollution" refers to the fact that the production and maintenance of AI technologies and their underpinning energy and data infrastructures have a significant environmental impact. This involves the production of greenhouse gases with negative consequences on the climate crisis, the depletion of natural resources used for AIT production, but also the release of toxic substances when they are decommissioned. The concept "digital pollution" shifts our attention towards the material infrastructure enabling AIT and its extractive nature (Crawford 2021). For example, data centers, the core infrastructure of AI systems housing servers, storing data, and supporting high-performance computing *processes* consume an increasing amount of energy to run their operations and cool the servers (Avgerinou et al., 2017). They also need diesel generators to keep the servers running in the event of a power outage, thus producing additional greenhouse gas emissions. The materiality of infrastructures and objects that make AITs clearly appear in the (dirty) mining of lithium and other minerals that are needed for their manufacturing as well as in the disposal of these materials through polluting practices such as incineration (Williams 2011).

Another source of digital pollution is directly related to the development of AI systems. A group of researchers from the University of Massachusetts has calculated the energy consumption linked to the training of some deep neural networks used for natural language processing (Natural Language Processing, NLP). The study showed that training the GPT-2 system resulted in the emission of greenhouse gases equal to almost five times the emissions of a medium-sized car throughout its life cycle (including the production of the car itself) (Strubell et al., 2019). As ChatGPT has become available to the broader public in 2022, it has become clear that we not only need to account for the energy consumption of training language models with large amounts of textual data, but also the environmental costs of using them (in addition to its social and political implications).

In summary, Information and Communication Technologies in general, and AI more specifically, involve high raw material and energy consumption and have a significant environmental impact. It must be emphasized that these environmental impact estimates still entail elements of uncertainty and incompleteness. And yet, they give an idea of the extent of a less discussed problem in a context of constant growth in digital demand and supply (Freitag et al 2020). These elements of uncertainty are not sufficient reason to overlook the problem. In fact, according to the Paris Agreements, to contain the rise in global temperature below 1.5°C, every productive sector must adapt and make significant changes. Even taking the most optimistic forecasts as a reference, it has been shown that the ICT and AI sectors cannot continue to produce the same amount of emissions, but must commit to a reduction to avoid an excessive increase in global warming (Blair 2020). So, despite the lack of scientific consensus on the extent of digital pollution, there are legitimate reasons to predict irreversible damage in the event of inaction, especially given the growth rate of the industry[v]. As we will see later, this is only partly a technical-scientific problem (linked, for example, to improving the efficiency of the systems or to a more precise calculation of their environmental impact), but also (and perhaps above all) a problem of moral and political choices.

Digital pollution and moral values

In acknowledging and tackling the problem of digital pollution, we cannot rely on a mechanistic calculation of data, but we must 1) consider value questions and political choices in regulating AI, and 2) reflect on how to best assign responsibilities for compensating for environmental costs of AI production and use to the various interest groups. Let us examine these two aspects in turn.

Facts and values

Building houses and heating them, eating, moving: it is reasonable to think that all human activities have an impact on the environment and on the climate crisis [vi], but this does not make us give up these activities. Only some of the most radical activists of the environmental cause recommend going to live in huts without heating, picking fruits spontaneously offered by nature and moving only on foot [vii]. A more moderate (and more widespread) environmentalist approach consists in *evaluating the costs and benefits* of each of these activities, promoting the construction of houses with low energy consumption (never zero!), a diet based on products with a reduced carbon footprint (for example, consuming less products of animal origin and more local products) and traveling by public transport or lower emission transport, such as hybrid cars or trains. A similar approach regarding digital pollution would involve seeking a reduction in the environmental impact of digital technologies and AI, but not a renunciation of their use. This approach is common in policy contexts where technologies' social costs are weighted against their potential benefits. While adopting a cost-benefit analysis (CBA) approach seems to be a more reasonable option than a complete renunciation of technology, it is not as straightforward as often implied (on risk analysis and AI, also see: Krutzinna, this volume). CBA is a valuable tool in contexts where the relationship between benefits and costs is calculated through the assignment of a monetary value. However, it has been debated whether monetizing options or assigning a quantifiable value to positive or negative consequences is a good approach in political decision making (Turner 1979). First of all, quantifying costs or benefits can be difficult when the impact of a certain policy or technology cannot be determined with certainty. Secondly, CBA assumes a consensus on how to define a benefit (a consequence with a positive connotation) or a cost (a consequence with a negative connotation). This consensus is relatively easy to reach when considering monetary values that are easy to quantify, but becomes more complex when different evaluation criteria and perspectives come into play. Let's scrutinize the limits of cost-benefit analyses in the concrete case of digital pollution in more depth.

To manage and reduce digital pollution it is important to know its extent, both in absolute terms (how much information technologies and AI pollute) and in relative terms (how much they pollute compared to the resulting environmental and social benefits). These two seemingly simple questions require complex answers. First of all, it is not easy to establish the extent of digital pollution in absolute terms. Let's take the example of data centers. Analyses of current energy consumption and forecasts of future development vary greatly from each other. They primarily depend on data that is not always easy to acquire because it is often not publicly available and the metrics used to acquire it vary across contexts (Whitehead et al., 2014). Secondly, although the technologies used in data centers are constantly evolving and tending to become more efficient, the demand for access, sharing, storage and processing of data is increasing simultaneously. The picture becomes even more varied and complicated when other factors are considered. Energy consumption and greenhouse gas emissions from data centers are a *direct environmental impact* and relatively easier to measure once standard parameters are decided and data is made available. But a more realistic assessment of the environmental impacts of AI (or other digital)applications must also take into consideration their *indirect* effects on users' behaviors and practices (Berkhout and Hertin, 2004). For example, increases in efficiency of AI models, combined with the promise of reduction of costs associated to experiments with animals, drives digital twins research. Investments in this field (if not monitored with respect to their environmental impacts) could induce some energy consumption behaviors that have a negative impact on the environment: researchers could in fact run more experiments, repeat the same experiments and train their models for longer times with an overall back fire to the initial efficiency gain. Thus, even if the higher performance of a technology makes it more efficient and allows for a decrease in energy consumption and unit price, consumers could increase its use or use the savings achieved to use other technologies and develop new high-cost behaviors. Systemic effects of this type are also called "rebound effects" (Kelly et al 2022).

Indirect effects are quite difficult to evaluate because they require consideration of social and cultural behaviors and norms. The daily practices of Internet use, for example the times of day people stream video (or GPT applications), are an important variable to evaluate the demand for data and therefore to evaluate the corresponding energy consumption (Morley et al., 2018). Too often, however, effects related to social and cultural characteristics of the user and rebound effects are excluded from assessments of the environmental impacts of digital technologies (Pohl et al., 2019). This analysis becomes even more complex when we try to consider the entire life cycle of AI hardware and software systems, from the extraction of minerals that are used to build computers, to the development and use of software systems, to the disposing of hardware in landfills [viii]. The limitations of existing methodologies for assessing all relevant variables and managing the corresponding factors of uncertainty require constant monitoring of AI developments and complicate cost-benefit analyses.

It is also crucial to consider that a CBA not only requires having available numerical values that can be compared, but also requires having a reference system to evaluate the social significance of these values. As we have seen above, the concept of sustainable development refers not only to the environmental dimension, but also to the economic and social one. Therefore, to establish the sustainability of AIT we must not only be able to measure the numerical value of their environmental impact[ix], but we must also ask ourselves about their "social value", for example the way in which they promote the satisfaction of needs such as access to education and basic health care, or foster a society with less inequality and discrimination. Based on a CBA, if the AIT in question has a higher social value, we may be more willing to justify a higher environmental impact. But from what (or whose) perspective is the social value of an AI service measured? How, for example, are we to quantify the value of communication services compared to services for online commerce or for the storage of electronic health records? Should we accept the environmental costs of producing life-improving AIs in a Global South context more readily than those developed to patrol Global North borders or even kill people in warzones (cf. Bode and Qiao-Franco, this volume; Molnar, this volume). And how do we evaluate the social value of different types of research or applications in AI and other fields of computing?

Social value is not only not easily measurable, but it is not even determinable uniformly. It changes, for example, according to the population in reference to which the benefits are evaluated. For example, health data collection and processing services may be more important in a highly computerized country than they are in a less computerized country, where paper records are less used. Furthermore, the populations most vulnerable to the climate crisis are also those who are currently least benefitting from the digital revolution. Or, if they are, like in the case of digital farming in Africa, the economic benefits of "digital inclusion" might still be accumulated in the Global North (see discussion of colonial "digital intrusion" by Brooks, this volume; and of data colonialism by Gray, this volume, as well as Omotubora and Basu, this volume). The assessment becomes even more complicated if the needs and interests of future generations are taken into consideration or if a less anthropocentric perspective with consideration of the needs of different ecosystems is assumed. . An environmental assessment based on CBA is not easy to implement in this field, since the values of both environmental and social impact are difficult to measure and vary according to context.

These various elements of uncertainty and complexity suggest that current environmental impact assessment measures may not be sufficient to provide definitive answers or a clear weighting of benefits and costs. It is important to note that such conceptual and empirical difficulties do not provide a justification for neglecting the problem, since the climate crisis requires urgent action and measurements are important for this. Rather, they show that decisions and moral value judgments need to be made in a context of uncertainty and incomplete information. Not trying to tackle these issues equals an acceptance of the environmental (and social) costs of AIT production and deployment. Furthermore, a critical reflection needs to consider not only quantitative aspects, but also more qualitative, social and moral dimensions of AI's sustainability. This is not a *one-off* exercise, but requires a continuous reflection that runs in parallel with the development of AIT.

Individual or institutional responsibility?

There are many related political, social and ethical issues that need to be addressed for the sustainable design, deployment and governance of AI technologies. But by whom? In their article on energy consumption of model training for NLP, researchers from the University of Massachusetts in Amherst question the responsibilities of the industrial, academic and public sectors with respect to the environmental pollution of AI research. They suggest a more equal distribution of the limited AI resources (which are currently largely in private hands in the Global North) and the development and use of more computationally efficient algorithms and devices with lower energy consumption. The recommendations go in the same direction as those of the European Commission's High-Level Expert Group on Artificial Intelligence, appointed to develop guidelines for a "Trustworthy AI"[x]. Among the criteria identified by the group they mention the environmental sustainability of AI: according to the group monitoring the environmental impact resulting from the development, distribution and use of AI systems must be implemented together with appropriate measures to reduce their environmental impact throughout the life cycle. These requirements remain at such a high level of abstraction that they are difficult to implement. What measures of the lifecycle environmental impact of an AI system are to be trusted? Who bears the costs of such monitoring, of the creation of more efficient systems and with a better overall environmental footprint? How to ensure that these costs do not further penalize small research centers, small and medium enterprises and users in favor of the already powerful multinationals in the sector? What incentive schemes – like in carbon emission certificate trading – should be designed to make AI developers and users consider and offset the environmental costs of their activities?

The fact that these ethically motivated requests in essence still translate into a self-regulatory system which is entirely voluntary, means that various questions regarding the distribution of responsibilities and compliance with certain environmental sustainability standards remain unanswered so far. If the moral and legal responsibilities for AI technologies' environmental impact are not well identified, any regulatory attempt to mitigate this impact loses strength and remains at the mercy of individual (profit-seeking) actors' decisions. For example, some large technology companies such as Microsoft, Google, Amazon and Apple have often made public commitments to reduce greenhouse gas emissions. While this commitment is commendable in principle, it is unclear in practice what these companies promise. Some propose to achieve a *carbon neutral objective*, which consists in offsetting *their* greenhouse gas emissions into the atmosphere with actions to eliminate these gases (such as reforestation activities, for example). At the moment, however, there is no common standard to determine: 1) which elements of the supply chain are considered in the calculation of the emissions produced by a company (for example if only emissions produced by fossil fuels *in situ* are considered or if the measures include emissions produced by use of electricity and gas purchased from third parties, or if we also calculate the indirect emissions produced at different stages of the supply chain, such as extraction, manufacturing, transport, use, disposal); 2) how these emissions should be offset (for example through the use of renewable energies or the purchase of carbon credits) and 3) which compensation strategies are most effective (for example buying and conserving an existing forest or financing reforestation projects).

These ambiguities lead some activists and researchers to argue that a carbon neutral goal for AIT and digital technologies is not ambitious enough, because it merely offsets, but does not *reduce in absolute terms*, all emissions (Freitag et al 2020). According to this perspective, aiming for *zero net energy* is a more desirable goal, since it implies not only a compensation between gases emitted and removed, but also an absolute reduction of emissions throughout the supply chain. In theory, a company aiming for "carbon neutrality" could increase total emissions and thus not contribute to the effort to limit global warming to 1.5°C, a goal which requires reducing annual emissions by 50% by 2030 and the achievement of "zero net emissions" (*net zero)* by 2050 (https://www.ipcc.ch/sr15 /). Given these premises, companies that are committed to achieving *carbon negative targets* are even more commendable, because they intend to eliminate more greenhouse gases than they emit throughout the supply chain. Yet these goals are often presented without precise information on their implementation and have therefore led some commentators to speak of *greenwashing*, problematizing that companies may adopt a facade of ecologism to appease public opinion and disseminate a virtuous image of their environmental commitment. For example, some companies claim they want to use only renewable energy (solar, wind, hydro), but at the same time they do not specify if and how the transition to one hundred per cent renewables will decrease their overall emissions and from which these energy sources *come* from. These examples show the limits of leaving issues of environmental impact reduction – in AIT development just as elsewhere – to the self-regulation of individual organizations.

Consensus on terminology, standards and indicators will be crucial for ensuring that policymakers can rely on robust measurements (OECD 2022). The development of a shared framework for addressing AI environmental costs must be an international governance effort and cannot be left to the market. At the same time, such international efforts must also recognize that many aspects of AIT's environmental impacts are either difficult to measure (e.g. biodiversity assessments) or not easily comparable with other benefits and costs (e.g. health impacts on workers in informal e-waste sites). Public discussions of who benefits from AIT and who bears its environmental (but also wider social) costs thus remains crucial.

Conclusions

The use of AI-fueled services and devices, but also the collection and analysis of large amounts of data and the training of AI systems, certainly contribute to social and economic development and the search for solutions to environmental problems. In this sense, these are technologies that can help achieve sustainable development and meet the needs of the global population, animal species and ecosystems. At the same time, the relationship between AIT and sustainability is highly ambiguous due to the technologies inherent environmental costs and impacts.

The chapter reviewed the concept of digital pollution to capture different types of negative effects of AIT production and deployment for the environment. This includes energy consumption, the use of natural resources and greenhouse gas emissions by data centers, the emission of toxic substances during the disposal of electronic devices and the emissions produced by training neural networks in research on AI. Although data on the lifecycle of these processes and products are difficult to collect and analyze and forecasts of future emissions are conflicting , it is clear that the issue of environmental sustainability of these technologies must be addressed.

As this chapter argued, the search for solutions to digital pollution is not an exclusively technological issue, but also a political and value-based one. In this sense, CBA aiming to quantify the costs and benefits of AIT have limited use in a context where knowledge on the very costs and benefits are uncertain because they strongly depend on behavioral and social aspects. Rather than (merely) attempting to quantify the environmental implications of the AI production process and life cycle, it is important to articulate the plurality of values involved and to discuss relevant stakeholders' perspectives in such assessments (beyond a focus on corporate actors in the Global North). Also, the issue of digital pollution cannot be separated from the discussion on the assignment of responsibilities between the parties involved. At the moment, solutions are often left to the market and the voluntary adoption by the private sector of targets to reduce its emissions (for a wider debate of regulatory capture in AI, see Paul, this volume). This leads to a lack of clear and effective standards on AIT's environmental impact. To address this responsibility gap public, national and supranational organizations will have a more decisive role to play in setting up monitoring, control and verification mechanisms and in enforcing them.

At the moment there are several codes of self-regulation by the private sector and some legislation that specifies certain standards (for example regarding the improvement of the energy efficiency of data centers or the inclusion of electronic waste in special recycling and disposal plans), [xi] but the approach of public institutions is fragmented and there is a lack of policies and regulations that address the issue of digital sustainability effectively, through incentives or sanctions. This is partially due to the lack of accurate data on AITs environmental footprint. In line with these considerations, in a 2022 report about the AI footprint, the OECD recommends "the establishment of measurement standards, expanding data collection, identifying AI-specific impacts, looking beyond operational energy use and emissions, and improving transparency and equity to help policy makers make AI part of the solution to sustainability challenges" (OECD 2022, p.2). As the report explains, accurate measurements of AI environmental impacts are key to guide sustainable policy decision making. Although in agreement with this statement, this chapter argues that achieving these measurements may take time and may still be insufficient as it leaves aside societal and non-quantifiable values. Sustainable policy making is bound to act in this context of numeric uncertainty and request more transparency from relevant actors, while at the same acknowledging the inadequacy of measurements and ensuring that complex and non-measurable aspect are taken into account into the decisional process. a Questions of accountability need to be addressed in this complex situation and cannot be postponed to the moment that we get the number right. Overall, we need a critical approach to guide a global AIT governance project capable of meaningfully addressing the sustainability of these technologies and defining the criteria that must form the basis of regulatory instruments. Critical reflections as those shared in this chapter can help clarify the underlying concepts, diagnose problems, identify contradictions, explain conflicts between different values and needs, and offer a moral basis to justify such a global governance project.

Notes

[i] This chapter is a revised version of a text originally published (in Italian) in the book Fossa F., Schiaffonati, V., Tamburrini, G. (eds) *Automi e persone: Introduzione all'etica dell'Intelligenza Artificiale e della robotica* (, Carrocci Editore, 112-127

[ii] Exact
search:(AI AND sustainab* AND (governance OR ethics OR regulation)) AND PUBYEAR > 2006 A ND PUBYEAR < 2024

[iii] Artificial Intelligence Act. Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))

[iv]https://ai-ethics-and-governance.institute/translation-series-on-ai-ethics-governance-and-sustainable-development/#:~:text=And%20in%20June%202019%2C%20the,is%20safe%2C%20controllable%20and%20reliable%2C
https://www.loc.gov/item/global-legal-monitor/2019-09-09/china-ai-governance-principles-released/
https://perma.cc/V9FL-H6J7

[v]This consideration refers to the precautionary principle, which is central to environmental reflection. According to the formulation of the Rio Declaration (1992), this approach aims to protect the environment and consists in adopting, in the event of risk of serious or irreversible damage, precautionary measures to prevent environmental degradation, even in the absence of a consensus of the scientific community on the extent of the risks: "Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation". The principle is often criticized because in its more rigid formulation it would require certain proof of the absence of risk before the approval of any innovation, thus causing a conspicuous slowdown in scientific-technological progress; however, in its softer formulation, it underlines the responsibility of States and interested parties to assess the possible risks and costs for the environment and future generations. For further information on the precautionary principle, see Hanson (2018).

[vi] Although there is no consensus that human activity is the sole cause of global warming, the report published by the Intergovernmental Panel on Climate Change (IPPC) in 2013 establishes it is highly probable that climate change is anthropogenic (caused from human activity).

[vii] An example of this approach is the one adopted by some political movements in the 60s and 70s, in which environmentalist motivations were associated with the rejection of the capitalist ideology and the excesses of the industrial revolution (Smith, 2003).

[viii] Environmental scientists refer to methodologies that evaluate the environmental impacts associated with the different stages of the life cycle of a service, process or product as "Life Cycle Assessment" (LCA) or "cradle-to-grave".

[ix]It should be emphasized that there are different parameters for assessing the environmental impact. For example, together with the *carbon footprint* (discussed in footnote 1), the *ecological footprint* is another parameter to measure the impact of human activities on the environment. If the *carbon footprint* is a measure of the weight (in tons or kilograms) of the gas emissions responsible for global warming, the *carbon footprint ecological* (or environmental) has a more literal meaning, as it is an indicator of space that converts the consumption of natural resources and the production of resources into units of measure of area. Global hectares (gha) express the amount of environmental space (biologically productive area of sea and land) that would be necessary for the entire world population to maintain a given lifestyle. It has been calculated that it would take five "planet Earths" to support the average US individual's lifestyle, three and a half to support the Italian lifestyle, and a half to support the average Indian lifestyle (https:// data . footprintnetwork.org/#/ ?; accessed February 18, 2021).

[x]High-Level Expert Group on Artificial Intelligence (https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence ; accessed 19 February 2021)

[xi]As regards self-regulatory codes, it is worth noting that of the International Telecommunication Union (ITU) (https://www.itu.int/en/mediacentre/Pages/PR04-2020-ICT-industry- to -reduce- greenhouse-gas-emissions-by-45-percent-by-2030.aspx). As far as European regulations are concerned, the *European Green Deal* partly discusses the need to reduce emissions from data centers (https://ec.europa.eu/info/sites/info/files/european-

green-deal-communication_en. pdf). Circular economy policy issues for e-waste are addressed in the document: *European Commission, 2020d. Changing how we produce and consume: New Circular Economy Action Plan shows the way to a climate-neutral, competitive economy of empowered consumers.*
(https://ec.europa.eu/commission/presscorner/detail/en/ip_20_420).

<b>References<b>

Attfield, R. (2013), *Sustainability*, in H. Lafollette (ed.) *International Encyclopedia of Ethics.*

Avgerinou, M., Bertoldi, P., Castellazzi, L. (2017), *Trends in Data Center Energy Consumption under the European Code of Conduct for Data Center Energy Efficiency*, in "Energies", 10(10), 1470.

Berkhout, F., & Hertin, J. (2004), *De-materialising and re-materialising: Digital technologies and the environment*, in "Futures".

Blair, G. (2020), *A tale of two cities. Reflections on digital technologies and the natural environment*, in "Patterns", 1(5).

Coeckelbergh, M. (2021). AI for climate: freedom, justice, and other ethical and political challenges. *AI and Ethics*, *1*, 67–72. https://doi.org/10.1007/s43681-020-00007-2

Crawford, Kate. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021. https://doi.org/10.2307/j.ctv1ghv45t.

Dauvergne, Peter. *AI in the Wild: Sustainability in the Age of Artificial Intelligence*. MIT Press, 2020.

Freitag, C., Berners-Lee, M., Widdicks, K., Knowles, B., Blair, G., & Friday, A. (2020). *The climate impact of ICT: A review of estimates, trends and regulations* (available at https://arxiv.org/abs/2102.02622 , accessed February 5, 2021).

Gabrys, J. (2013), *Digital Rubbish: A Natural History of Electronics*, University of Michigan Press.

Hanson, J. (2018), *Precautionary Principle: Current Understandings in Law and Society* , in DA Dellasala & A. Goldstein (eds.) *Encyclopedia of the Anthropocene*, pp. 361–6. Elsevier.

Herweijer, C & Waughray, D. (2018), *Harnessing Artificial Intelligence for the Earth* (https://www.pwc.com/gx/en/services/sustainability/publications/ai-for-the-earth.html , consulted on February 2, 2021).

Hilty, Lm, Aebischer, B., Andersson, G., & Lohmann, W., (2013). *ICT4S 2013* , Proceedings of the First International Conference on Information and Communication Technologies for Sustainability, ETH Zurich, February 14-16, 2013 (https://doi.org/10.3929/ETHZ-A-007337628 , accessed February 2, 2021).

Jobin, A., Ienca, M., And Vayena, E. (2019), *The global landscape of AI ethics guidelines*, in "Nature Machine Intelligence", 1(9), pp. 389-99.

Lucivero, F. (2019), *Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives*, in " Science and Engineering Ethics".

Maxwell, R., & Miller, T. (2012), *Greening the Media*, Oxford University Press, USA.

Morley, J., Widdicks, K., & Hazas, M. (2018), *Digitalisation, energy and data demand: The impact of Internet traffic on overall and peak electricity consumption* , in "Energy Research & Social Science", 38, pp128 –37.

OECD (2022), "Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint", *OECD Digital Economy Papers*, No. 341, OECD Publishing, Paris, https://doi.org/10.1787/7babf571-en.

Schlosberg, D. (2009), *Defining environmental justice: theories, movements, and nature* , Oxford University Press, Oxford.

Smith, K. (2003), *Wendell Berry and the Agrarian Tradition: A Common Grace*, University of Kansas Press, Lawrence.

Strubell, E., Ganesh A., And Mccallum A. (2019), *Energy and policy considerations for deep learning in NLP* , arXiv preprint *(arXiv:1906.02243)* .

Taylor, L. (2017), *What is data justice? The case for connecting digital rights and freedoms globally* , in "Big Data & Society", *4* (2), 205395171773633.

Turner, Rk (1979), *Cost-benefit analysis—a critique,* in " Omega ", *7* (5), pp. 411–9.

van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics 2021 1:3*, *1*(3), 213–218. https://doi.org/10.1007/S43681-021-00043-6

Vonderau, A. (2018), *Scaling the Cloud: Making State and Infrastructure in Sweden* in "Ethnos Journal of Anthropology"

WCED (World commission on environment and development) (1987), *Our common future* (https://www.un.org/ga/search/view_doc.asp?symbol=A/42/427&Lang=E , accessed 3 February 2021).

Whitehead, B., Andrews, D., Shah, A., & Maidment, G. (2014), *Assessing the environmental impact of data centers part 1: Background, energy use and metrics*, in "Building and Environment", *82*, pp.151–9.

Widdicks K, Lucivero F, Samuel G, Croxatto LS, Smith MT, Holter CT, Berners-Lee M, Blair GS, Jirotka M, Knowles B, Sorrell S, Rivera MB, Cook C, Coroamă VC, Foxon TJ, Hardy J, Hilty LM, Hinterholzer S, Penzenstadler B. *Systems thinking and efficiency under emissions constraints: Addressing rebound effects in digital innovation and policy.* Patterns (N Y). 2023 Feb 10;4(2):100679.

Williams, E. (2011), *Environmental effects of information and communications technologies* , in “ Nature”, *479* (7373), pp. 354–8.