



# Rethinking priorities in AI ethics guidelines in today's era of humanity under threat

Bako Rajaonah, Enrico Zio

## ► To cite this version:

Bako Rajaonah, Enrico Zio. Rethinking priorities in AI ethics guidelines in today's era of humanity under threat. 2022. hal-03585261

**HAL Id: hal-03585261**

**<https://hal.science/hal-03585261>**

Preprint submitted on 23 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Title

Rethinking priorities in AI ethics guidelines in today's era of humanity under threat

## Authors

Bako Rajaonah<sup>a</sup>, Enrico Zio<sup>bc</sup>

## Affiliations

<sup>a</sup>Laboratory of Industrial and Human Automation control, Mechanical engineering and Computer Science (LAMIH UMR CNRS, 8201), Université Polytechnique Hauts-de-France, Valenciennes, France

<sup>b</sup>Centre de Recherche sur les Risques et les Crises (CRC), MINES ParisTech/PSL Université Paris, Sophia Antipolis, France.

<sup>c</sup>Department of Energy, Politecnico di Milano, Milan, Italy.

**Abstract:** This paper does not call into question the existing ethical guidelines for artificial intelligence, but suggests rethinking their priorities in today's era of humanity under threat. In December 2020, the United Nations' Secretary General described the anthropogenic degradation of the planet, claiming that "the planet is broken." Adopting the approach of disaster risk reduction, we assert that the ongoing destruction of the planet is a disaster, which leads us to think in terms of resilience. We have examined existing works on ethical guidelines for AI through the lens of philosophy, namely, the imperative of responsibility toward the distant future of nature, including humanity, and an ethics of care articulated around maintenance, continuance and repairs. We have identified five ethical principles: respect for nature, respect for human rights, AI usefulness, AI transparency, and AI trustworthiness, which are explained through 19 subprinciples. We conclude with the difficulty of being concretely nature-friendly in today's era of humanity under threat.

**Keywords:** artificial intelligence; nature; future; resilience; ethical principles; care ethics

# **Rethinking priorities in AI ethics guidelines in today's era of humanity under threat**

## **1. Motivation**

The landscape of artificial intelligence (AI) ethics guidelines developed by Jobin et al. (2019) reveals a plethora of works about creating systems based on ethical AI. These authors have thoroughly analyzed the ethical guidelines for AI contained in 84 documents from national, European, and international organizations. Jobin et al.'s goal was to identify principles that would transcend the diverse interests of AI-related stakeholders. They concluded that these principles could be connected to the following values: transparency; justice and fairness; non-maleficence; responsibility; privacy; beneficence; freedom and autonomy; trust; sustainability; dignity; and solidarity.

Respect for all these values should lead to ideal AI for humanity's and society's well-being. However, is this ideal high enough when it comes time "to address our planet's emergency" in transforming humankind's relationship with nature to restore the "broken" planet and to ensure humanity's survival?<sup>1</sup> Although the natural environment is taken into account in ethical AI guidelines, it is often a mere point among many others and does not deserve to be seen as a full-fledged principle, except in the 2018 Montréal Declaration, as outlined below.

- Well-being is one of the 8 general principles of the IEEE Global Initiative (2019), which considers human well-being to be inseparable from the well-being of the natural environment.
- At the European level, the principle of preventing harm—one of the 4 principles of the EU's ethics guidelines for trustworthy AI—considers both people and the natural

---

<sup>1</sup> UN's Secretary General at Columbia University on 02 December 2020: [https://www.un.org/sg/en/content/sg/statement/2020-12-](https://www.un.org/sg/en/content/sg/statement/2020-12-02/secretary-generals-address-columbia-university-the-state-of-the-planet-scroll-downfor-language-versions)

[02/secretary-generals-address-columbia-university-the-state-of-the-planet-scroll-downfor-language-versions](https://www.un.org/sg/en/content/sg/statement/2020-12-02/secretary-generals-address-columbia-university-the-state-of-the-planet-scroll-downfor-language-versions) (accessed September 22, 2021)

environment, especially by requiring societal and environmental well-being (HLEG AI 2019).

- One of the 10 principles of the 2018 Montréal Declaration for the responsible development of AI is entirely devoted to the environmental sustainability of the planet, including energy efficiency, electric and electronic waste, maintenance, and repair (Dilhac et al. 2018.)

The fact that the protection and repair of nature are not at the heart of most ethical AI guidelines could be explained by the need to first ensure a good use of AI when its potentiality became evident; for example, with the rise of big data (see Mayer-Schönberger & Cukier 2013) or regarding AI's uncontrollability (see Yampolskiy 2021).

However, the context has changed: We assert that, given the ongoing anthropogenic destruction of the planet thanks to increasingly powerful technologies, considering resilient AI might be more fruitful than focusing on ethical AI alone: we argue that in today's era of humanity under threat, the goal of ethics for AI should shift from "good, ethical AI for people" to an AI that adapts its values to meet nature's requirements. The aim of the present work is to suggest a new order of priorities in ethical AI guidelines which moves in this direction.

In the next section, we try to justify this change in perspective by blending philosophies, institutional frameworks, and theoretical approaches such as resilience, disaster risk reduction, care, and life-centered design. Section 3 rethinks ethical AI guidelines to match the new purpose of acceptable AI for nature. The final section concludes with the urgency of being concrete regarding nature-friendly attitudes and behaviors in today's time of humanity under threat.

## **2. Corroborating the need for resilient AI in the era of humanity under threat**

This section introduces frameworks and approaches to argue about the need to rethink priorities in ethical AI guidelines.

### **2.1 Some facts to start with**

The 1972 Declaration of the Stockholm Conference on the Environment, organized by the United Nations (UN), called upon “Governments and peoples to exert common efforts for the preservation and improvement of the human environment,” after having acknowledged “the dangerous levels of pollution in water, air, earth and live beings; major and undesirable disturbances to the ecological balance of the biosphere; destruction and depletion of irreplaceable resources.”<sup>2</sup>

However, in addition to the anthropogenic, direct degradation of the environment, there is an anthropogenic part of global warming: The 1990 Report of the Intergovernmental Panel on Climate Change (IPCC)<sup>3</sup> acknowledged the role of greenhouse gases (GHG) in natural and anthropogenic warming, as well as the effects of warming on ice volume, sea level, and ecosystems, and of ecosystem changes on climate (e.g., reduced rainfall related to forest cover modification; shifts in ocean-atmosphere exchange due to changes in ocean plankton).

We have thus known for half a century that our activities damage the environment. However, the IPCC<sup>4</sup> reported in 2018 that global warming continued to rise and warned us of climate-related risks for natural and human systems. In December 2020, the UN Secretary General released a report on the state of the planet<sup>1</sup>: in other words, the facts reported in 1972<sup>2</sup> and 1990<sup>3</sup> have worsened, and “the planet is broken.”

---

<sup>2</sup> UN Declaration of Stockholm, 1972: <https://www.un.org/en/conferences/environment/stockholm1972> (accessed September 9, 2021)

<sup>3</sup> IPCC, 1990: [https://archive.ipcc.ch/ipccreports/far/wg\\_I/ipcc\\_far\\_wg\\_I\\_full\\_report.pdf](https://archive.ipcc.ch/ipccreports/far/wg_I/ipcc_far_wg_I_full_report.pdf) (accessed September 9, 2021)

<sup>4</sup> IPCC, 2018: <https://www.ipcc.ch/sr15/> (accessed September 9, 2021)

Thus, what can—and should—we do to keep on designing, developing and using AI without worsening the state of the planet? One answer could be to consider that acceptable AI is not just ethical AI, but above all, resilient; that is, AI that expresses concern about the preservation and repair of the planet, even if it means upsetting priorities and ethical values.

## **2.2 A bit of philosophy**

It is necessary to question moral values in new contexts, upstream of elaborating norms that constitute morality (Ricoeur 1990 2000).

The philosopher Hans Jonas has argued since the 1970s for a new conception of duties and rights that would span beyond our time and space to the consequences of our use of *techne*: To ensure the existence of future generations, Jonas suggested a new ethics based on the imperative of responsibility toward the distant future of nature, including humanity (Jonas 1973 1984). Similarly, the philosopher Paul Taylor proposed a life-centered system of environmental ethics centered not on humans, but on respect for nature (Taylor 1981). Such ethics rely on beliefs such as “the human species, along with all other species, are integral elements in a system of interdependence such that the survival of each living thing [...] is determined not only by the physical conditions of its environments, but also by its relations to other living things” (Taylor 2011: 99).

Importantly, in many cultural groups and cultures, gratitude toward and respect for nature were (or are still) considered duties; for example, the generalized traditional American Indian (Callicott 1982) or, more specifically, the Tibetan community in Jisha (Jianchu et al. 2004) and the Santhals of Eastern India (Dutta 2020).

For the ethics of care, the psychologist Carol Gilligan contended that it was time to consider a different (i.e. feminine) voice regarding morality since women generally have greater sensibilities than men in terms of responsibility and care within the relationship between the self and the other (Gilligan 1977). The questioning of care ethicists is first about

the concrete needs of others, and next about whether their actions would truly be a response to these needs (Mizzoni 2017). Joan Tronto proposed an ethics of care in which she described care as “a species activity that includes everything that we do to maintain, continue, and repair our ‘world’ so that we can live in it as well as possible” (Tronto 1988: 16).

### **2.3 Theoretical approaches**

Four approaches underlie our view of resilient AI: disaster risk reduction, caretaking, green ICT, and life-centered AI design.

Given the alarming state of the planet (§ 2.1), working on AI-related guidelines should be placed in the context of disasters. A disaster is defined by the UN Office for Disaster Risk Reduction (UNDRR) as “a serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability and capacity, leading to one or more of the following: human, material, economic and environmental losses and impacts.”<sup>5</sup> However, a disaster for those who suffer from the losses and impacts may simply be a social problem for those who are safe and privileged (Zack 2010), and yet the ongoing destruction of the planet concerns us all. Adopting the DRR approach enables us to benefit based on existing knowledge, know-how and, above all, the fundamental idea of reinforcing resilience; that is, “the ability of a system, community or society exposed to hazards to resist, absorb, accommodate, adapt to, transform and recover from the effects of a hazard in a timely and efficient manner.”<sup>5</sup>

If the destruction of nature concerns all of us, then we should show solidarity, even though doing so may upset our priorities and values. AI should be life-centered; that is, the priority would no longer be human well-being, but instead the well-being of each living thing on the planet since the survival of some depends on the survival of others (Taylor 1981 2011). Hence, it is imperative to care about nature in the sense of the ethics of care.

---

<sup>5</sup> UNDRR online glossary: <https://www.undrr.org/terminology> (accessed September 10, 2021)

Steven Jackson extended the ethics of care from interpersonal ties to relationships with technology; he argued for new attitudes and practices “vis-à-vis the world of media and technology today,” [...] “new forms of solidarity with our objects (and they with us),” and new views of responsibility and sustainability regarding technological breakdown, obsolescence and reuse (Jackson 2014; Jackson & Kang 2014).

Finally, we cannot overlook the approach of green information and communication technologies (green ICT). The concept of green ICT refers to “technologies and processes that are environmentally friendly, i.e., which have a lower negative impact on the natural environment than conventional ones” (Loeser 2015: 5). The holistic approach of green ICT, described by Murugesan (2008), has four dimensions of greenness that should be taken into account regarding ethical principles for resilient AI: green use (low energy consumption technology), green disposal (refurbishment, reuse and recycling), green design (environmentally friendly technology), and green manufacturing (manufacturing technology and components with minimal or no impact on the environment).

To conclude Section 2, we place care for nature at the heart of ethical AI guidelines, with the primary ambition to repair and protect the natural environment, and to ensure the continuance of humanity. This is in line with the UN document published in 2021: “Making peace with nature,” which concerns 10 of the UN 17 Sustainable Development Goals (SDGs)<sup>6</sup>.

### **3. Our view of ethics guidelines for resilient AI in the era of humanity under threat**

Following § 2.2 and § 2.3, AI is not only a powerful *technique* that must be designed, developed and used in a way that does not destroy nature, but also a nature-friendly technology that takes care of nature through maintenance, continuance and repair.

---

<sup>6</sup> UN (2021) Making peace with nature: <https://sdgs.un.org/publications/publication-making-peace-nature-scientific-blueprint-tackle-climate-biodiversity-and> (accessed: 10/21/2021)



First, we assume that resilient AI means that acceptable AI is not simply ethical AI: We assume that there is an “if and only if” relationship between the two properties of ethicalness and environmental friendliness, and it is around this fundamental relation of equivalence that ethical principles are connected.

Under the assumption that all authors had roughly the same definition of “principle” in mind (i.e., “a moral rule or belief that helps you know what is right and wrong and influences your actions”<sup>7</sup>), we start from Table 1, which outlines the “principles” in four works (numbered [1], [2], [3] and [4] in Table 1): Dilhac et al. (2018), HLEG AI (2019), the IEEE Global Initiative (2019), and Jobin et al. (2019). After examining their interpretation of their principles, we concluded that 5 principles could uphold our view of resilient AI with respect to the care of nature and the continuance of humanity (Figure 1). The next subsections describe these general principles through the 19 related subprinciples.

### 3.1 Respect for nature

This principle is at the heart of our view of ethics for resilient AI. It includes at least the following five subprinciples.

- **Non-maleficence:** AI systems should not even be considered if their development, usage, or both threaten the existence and future of living beings and/or their ecosystems. The subprinciple of non-maleficence is part of the principle of preventing harm in [2] and a full principle in [4]. Non-maleficence, seen through the lens of life, is particularly involved in UN SDG No. 3 (Good health and well-being), 14 (Life below water), and 15 (Life on land).

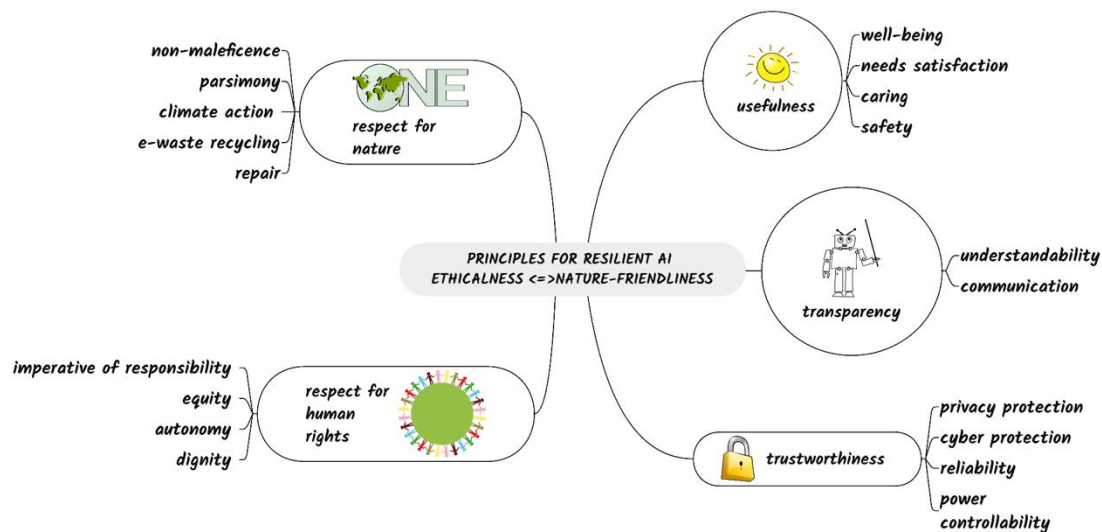
Table 1. Ethical principles in existing works on ethical AI guidelines.

---

<sup>7</sup> <https://www.merriam-webster.com>

Dilhac et al. (2018): 10 principles [1]	HLEG AI (2019): 4 principles [2]	The Global Initiative (2019): 8 general principles [3]	Jobin et al. (2019): 11 principles [4]
<ul style="list-style-type: none"> <li>- well-being</li> <li>- respect for autonomy</li> <li>- protection of privacy and intimacy</li> <li>- solidarity</li> <li>- democratic participation</li> <li>- equity</li> <li>- diversity and inclusion</li> <li>- caution</li> <li>- responsibility</li> <li>- sustainable development</li> </ul>	<ul style="list-style-type: none"> <li>- respect for human autonomy</li> <li>- prevention of harm</li> <li>- fairness</li> <li>- explicability</li> </ul>	<ul style="list-style-type: none"> <li>- human rights</li> <li>- well-being</li> <li>- data agency</li> <li>- effectiveness</li> <li>- transparency</li> <li>- accountability</li> <li>- awareness of misuse</li> <li>- competence</li> </ul>	<ul style="list-style-type: none"> <li>- transparency</li> <li>- justice, fairness and equity</li> <li>- non-maleficence</li> <li>- responsibility and accountability</li> <li>- privacy</li> <li>- beneficence</li> <li>- freedom and autonomy</li> <li>- trust</li> <li>- sustainability</li> <li>- dignity</li> <li>- solidarity</li> </ul>

Figure 1. Ethical principles and subprinciples of resilient AI in today's era of humanity under threat (created with XMind 2021, pictures from pixabay.com).



- **Parsimony:** The development and use of future AI systems should not be considered if they involve the use of natural resources that are excessive with respect to their reserves for subsequent generations. Sparing resources

necessarily implies sacrifices and trade-offs between necessity and costs, hence the need to change not only individual desires and habits, but also society and economic goals. Modifying our priorities and values is not easy at all; moreover, the subprinciple of parsimony is a big challenge for equity between people, countries and generations; it may even be irrational, as Nuta (2015) noticed indignantly.

- **Climate action:** At the time of the COP26,<sup>8</sup> actions against anthropogenic GHG emissions entailed, unsurprisingly, the development and use of AI systems, notably regarding the production and use of fossil-fuel electricity, and thus the impacts of resulting CO<sub>2</sub> emissions on the natural environment and society. Concretely, and in accordance with the green ICT approach, the subprinciple of climate action aims to reduce the use of energy-consuming AI systems, to eliminate their use, or to only employ green AI systems.

On the other hand, AI systems are useful for climate action, and it is unthinkable to do without them. For example, when applied to Copernicus satellite data, AI techniques contribute to providing information about the health of the planet (e.g., Bereta et al. 2018) and enhance predictive models (for instance, of wildfire<sup>9</sup>). AI techniques provide opportunities in the sustainable energy industry (Ahmad et al. 2021) and for decreasing the energy consumed in buildings (Kaack et al. 2020). However, as seen earlier, this would be genuinely beneficial only if AI systems and applications were themselves green.

- **E-waste recycling:** Electric and electronic waste (e-waste) impacts not only the environment, but also people's health<sup>10</sup> (Noel-Brune et al. 2013)—not even sparing children (Lebbie et al. 2021), especially in developed nations (Osibanjo & Nnorom

---

<sup>8</sup> UN COP26: <https://www.un.org/en/unclimate-change-conference-cop-26>

<sup>9</sup>European Forest Fire Information System: <https://effis.jrc.ec.europa.eu> (accessed: 21 October 2021)

<sup>10</sup> <https://news.un.org/en/story/2019/01/1031242>

2007). Moreover, e-waste contains a large quantity of metals whose recycling remains a challenge, and existing techniques are not all environmentally friendly (Zhang & Xu 2016).

The challenge of e-waste is addressed only in [1] through the principle of sustainable development: “AIS hardware, its digital infrastructure and the relevant objects on which it relies, must aim to generate the least amount of electric and electronic waste and to provide for maintenance, repair, and recycling procedures according to the principles of the circular economy” (Dilhac et al. 2018: 17). This is completely in line with the ethics of care.

**Repair:** The first imperative of this subprinciple would be to break with the planned obsolescence of high technology at the manufacturer level. However, prohibiting it would pose serious societal and economic problems such as unemployment and economic gloom, not to mention the possibly insoluble dilemma raised by the compromise between technological progress and protecting the environment (e.g., Ali & Sinha 2016; Nowviskie 2015; Nuta 2015). As a consequence of dismissing the strategy of planned obsolescence, the repair subprinciple includes the design of AI systems that are not only long lasting, but also repairable as much as possible.

### 3.2 Respect for human rights

We consider four subprinciples that may be either congruent or, unfortunately, at variance with the principle of respect for nature.

- **Imperative of responsibility:** Both Article 2 of the European Union Charter of fundamental rights<sup>11</sup> and Article 3 of the Universal Declaration of Human Rights<sup>12</sup> concern the right to life for everyone. Logically, this applies to future generations,

---

<sup>11</sup> Article 2: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT>

<sup>12</sup> Article 3: <https://www.un.org/en/about-us/universal-declaration-of-human-rights>

which explains why we assert that the imperative of responsibility (Jonas 1984) is part of the principle of respect for human rights.

- **Equity:** This concept is explicitly a principle in the form of equity in [1], fairness in [2], and justice, fairness and equity in [4]. Equity has to do with equal treatment of everyone, fairness with impartiality, and justice with rightness<sup>7</sup>. The common denominator may amount to the three following guidelines: (i) AI systems should be designed for the common good, but without *a priori* discrimination about who the “common” are; (ii) the benefits of an AI system should be accessible to everyone concerned by the system; and (iii) inconvenience and disadvantages generated by the development and use of AI systems (such as natural resources use, e-waste, carbon footprints, energy costs, etc.) should be shared with discrimination; that is, distinguishing AI system producers and end users from those who have nothing to do with these systems (not a question of equality, but of rightness).
- **Autonomy:** AI systems should be designed to facilitate users’ decisional autonomy, freedom of choice, free will, and their control over their privacy and personal data ([1] [2][3] [4]). They should enhance human autonomy in augmenting, complementing and empowering human cognitive, social and cultural skills [2], as well as people’s knowledge of AI [1] [3] [4] and critical thinking [1]. The decisional autonomy of AI systems should be designed following the human-centered approach; that is, in putting humans’ values and interests above all and in preserving the human privilege of intervention [2][3].
- **Dignity** refers to the “intrinsic worth” of human beings, which includes identity integrity as well as physical, mental, and cognitive integrity [2]; in a word, one’s value as a human being. Therefore, the subprinciple of dignity implies that AI systems should be designed so that their development and use do not threaten integrity. For

example, using AI systems whose components contain raw materials for extraction involves forcing children to work in artisanal mining (see Schwartz et al. 2021); this violates—even indirectly—the subprinciple of dignity because those children are treated as objects, not as human subjects (e.g., HLEG AI 2019: 10).

### 3.3 Usefulness

The usefulness of AI systems could be considered a condition for their existence (e.g., Ennals, 1987); “useful” means advantageous, beneficial, helpful or serviceable<sup>7</sup>. We consider the principle of usefulness to be related to the following subprinciples: well-being, the satisfaction of needs, caring, and safety.

- **Well-being:** The “omnipresence” of well-being in [1], [2], [3] and [4] explains why we maintain it as an ethical principle for resilient AI. However, the concept is quite abstract, and well-being has several dimensions (individual, social, economic, environmental, etc.), which makes it difficult to assess. The OECD (2013) published guidelines on measuring subjective well-being, but subjective well-being is only one of the 11 dimensions of well-being at the country level<sup>13</sup>.
- **The satisfaction of needs:** The usefulness of AI systems can be gauged through the satisfaction and trust of target users’ needs and expectations (e.g., Lankton et al. 2014). Analyzing the needs of all stakeholders, including end users, is a prerequisite from the perspective of human-centered systems (e.g., Maguire 2001), which is explicitly recommended in [2] and [3]. Acceptability and acceptance may also be related to the subprinciple of the satisfaction of needs. Acceptability is the target users’ attitude toward a system before it is introduced and acceptance of the system

---

<sup>13</sup> Measuring well-being and progress: <https://www.oecd.org/statistics/measuring-well-being-and-progress.htm>

after its introduction, but both are related to willingness to use the system (Adell et al. 2016).

- **Caring:** AI systems should be designed with the sole purpose of protecting, taking care of, and repairing nature, including humans and related activities; that is, not only medical and educational AI systems but also agricultural, commercial and industrial systems, provided that they are “green-based” and truly useful to the common good.
- **Safety:** We have grouped together under this subprinciple some parts of existing principles related to unintentional harm. The principle of prudence [1] states that “every person involved in AI development must exercise caution by anticipating, as far as possible, the adverse consequences of AIS use.” The prevention of harm principle [2] stipulates that “AI systems should neither cause nor exacerbate nor otherwise adversely affect human beings.” The safety of systems is a condition for the principles of human rights and well-being in [3]; it occurs under the responsibility of creators and operators through the principle of competence, and could be taught to the public via the principle of awareness of misuse. Safety is broached in the principle of non-maleficence in [4].

### 3.5 Transparency

Transparency is predominant in 73 of the 84 documents that Jobin et al. (2019) examined. We consider two dimensions that are well documented in [1], [2], [3], and [4]; hence the brevity of our explanations.

- **Understandability:** The transparency of AI systems is linked to explainability, explicability, understandability, and interpretability in [4]. Explicability includes traceability of the datasets and the processes underlying systems’ decisions, as well as the explainability of systems’ decisions ([2] [3]). The principle of democratic

participation [1] states that the code for decision-making algorithms used for public decisions should be accessible to all unless there is a danger of misuse.

- **Communication:** This subprinciple is closely tied to “understandability”: Not only should information related to AI systems’ functioning be understandable and explainable; it should also be discoverable and shared, and, as emphasized in [3], it is important to adapt the level of transparency to each stakeholder (see also Buiten 2019).

### 3.6 Trustworthiness

We restrict trustworthiness to its definition related to the reliability<sup>7</sup> of AI systems in the sense that people should be able to depend on AI systems without worrying about them. From this angle, we consider four subprinciples.

- **Privacy protection:** This subprinciple refers to the design and use of AI systems that protect privacy, intimacy, and personal information without violating them [4].
- **Cyber protection:** The cyber security of AI systems, protecting them from cyber-attacks and malicious use (e.g., Craglia 2018).
- **Reliability:** AI systems should be dependable, even when unexpected threats of harm or danger arise; fallback plans are required for such situations [2].
- **Controllable power:** This subprinciple relates not only to the extent of an AI system’s decisional autonomy, but also to its power regarding its capacities: Humans should be kept in the loop of the AI system’s decision-making process ([2] [3]) and the systems that promote human agency ([1] [3]). Humans should always be able to understand the system’s purpose and functioning, and should never be in a position in which they cannot comprehend the system (see Buiten 2019; Yampolskiy 2021; and sci-fi stories and films...).



## 4. Conclusions

This paper does not call into question the existing AI ethical guidelines, but suggests rethinking their priorities in today's era of humanity under threat. However, ethical guidelines remain empty words if they are not applied. Fortunately, technological systems have to be certified by independent bodies before use and/or marketing, certification being “a useful tool to add credibility, by demonstrating that your product or service meets the expectations of your customers.”<sup>14</sup> For example, the International Organization for Standardization (ISO) has published more than 24,000 standards, including those devoted to environmental protection (ISO 13020), human-centered design (e.g., ISO 9241-210:2019 and ISO/TR 16982:2002), medical devices (ISO 13485), or mining and quarrying, including the traceability of rare earths (ISO 23664:2021). ISO is currently working on standards, such as guidelines for AI applications (ISO/IEC AWI 5339), and for functional safety and AI systems (ISO/IEC AWI TR 5469)<sup>15</sup>. The IEEE Standard Association has published a standard that addresses ethical concerns during systems design, including autonomous intelligent systems (IEEE 7000™-2021).

However, even if all objective technological, social, and environmental dimensions of the design and use of AI systems were formally framed by standards, there would still be ethical concepts that are difficult to apply, among other reasons because they are difficult to measure and will therefore remain empty words. Notwithstanding, these concepts are crucial; for example, those of well-being and equity.

Thus, the solution lies in a deep transformation in our values and ways of life. Further, “solidarity is humanity, humanity is survival,” the UN Secretary General said<sup>1</sup>. As such, it is urgent to be concrete and stick together in today's era of humanity under threat.

---

<sup>14</sup> International Organization for Standardization: <https://www.iso.org/home.html> (accessed November 5, 2021)

<sup>15</sup> ISO Standards for AI: <https://www.iso.org/committee/6794475/x/catalogue/> (accessed November 5, 2021)



## References

- Adell, Emeli, András Várhelyi, and Lena Nilsson. 2014. "The Definition of Acceptance and Acceptability." In *Driver Acceptance of New Technology*. Boca Raton, FL: CRC Press.
- Ahmad, Tanveer, Dongdong Zhang, Chao Huang, Hongcai Zhang, Ningyi Dai, Yonghua Song, and Huanxin Chen. 2021. "Artificial Intelligence in Sustainable Energy Industry: Status Quo, Challenges and Opportunities." *Journal of Cleaner Production* 289 (March): 125834. <https://doi.org/10.1016/j.jclepro.2021.125834>.
- Ali, Zeeshan, and Ashish Ranjan Sinha. 2016. "Integrating Ethics in Technical Education for Sustainable Development." *PURUSHARTHA - A Journal of Management, Ethics and Spirituality* 9 (1): 85–97.
- Bereta, Konstantina, Hervé Caumont, Erwin Goor, Manolis Koubarakis, Despina-Athanasia Pantazi, George Stamoulis, Sam Ubels, Valentijn Venus, and Firman Wahyudi. 2018. "From Copernicus Big Data to Big Information and Big Knowledge: A Demo from the Copernicus App Lab Project." In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 1911–14. CIKM '18. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3269206.3269232>.
- Buiten, Miriam C. 2019. "Towards Intelligent Regulation of Artificial Intelligence." *European Journal of Risk Regulation* 10 (1): 41–59. <https://doi.org/10.1017/err.2019.8>.
- Callicott, J. Baird. 1982. "Traditional American Indian and Western European Attitudes toward Nature: An Overview." *Environmental Ethics* 4 (4): 293–318. <https://doi.org/10.5840/enviroethics1982443>.
- Craglia, Max. 2018. "Artificial Intelligence: A European Perspective." Luxembourg: Publications Office of the European Union. <https://publications.jrc.ec.europa.eu/repository/handle/JRC113826>.
- Dutta, Uttaran. 2020. "Protecting Sacred-Groves: Community-Led Environmental Organizing by Santhals of Eastern India." *Environmental Communication* 14 (1): 36–51. <https://doi.org/10.1080/17524032.2019.1585895>.
- Ennals, Richard. 1987. "Socially Useful Artificial Intelligence." *AI & SOCIETY* 1 (1): 5–15. <https://doi.org/10.1007/BF01905885>.
- Fisher, Berenice, and Joan Tronto. 1990. "Toward a Feminist Theory of Caring." In *Circles of Care*, edited by Emily K Abel and Margaret K. Nelson, 36–54. Albany, NY: SUNY Press.
- Gilligan, Carol. 1977. "In a Different Voice: Women's Conceptions of Self and of Morality." *Harvard Educational Review* 47 (4): 481–517. <https://doi.org/10.17763/haer.47.4.g6167429416hg510>.
- HLEG AI. 2019. "Ethics Guidelines for Trustworthy AI." Luxembourg: Publications Office of the European Union. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- Jackson, Steven J. 2014. "Rethinking Repair." In *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: The MIT Press. <https://doi.org/10.7551/mitpress/9780262525374.003.0011>.
- Jackson, Steven J., and Laewoo Kang. 2014. "Breakdown, Obsolescence and Reuse: HCI and the Art of Repair." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 449–58. CHI '14. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/2556288.2557332>.
- Jianchu, Xu, Li Bo, and David Waltner-Toews. 2004. "Habitat of Tibetan Nature and

- Culture.” *EcoHealth* 1 (4): 327–29. <https://doi.org/10.1007/s10393-004-0147-x>.
- Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. “The Global Landscape of AI Ethics Guidelines.” *Nature Machine Intelligence* 1 (9): 389–99. <https://doi.org/10.1038/s42256-019-0088-2>.
- Jonas, Hans. 1973. “Technology and Responsibility: Reflections on the New Task of Ethics.” *Social Research* 40 (1): 31–54.
- . 1984. *The Imperative of Responsibility. In Search on an Ethics for the Technology Age*. Chicago, IL; London: The University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/I/bo5953283.html>.
- Kaack, Lynn, Priya Donti, Emma Strubell, and David Rolnick. 2020. “Artificial Intelligence and Climate Change: Opportunities, Considerations, and Policy Levers to Align AI with Climate Change Goals.” *Heinrich-Böll-Stiftung*. <https://opus4.kobv.de/opus4-hsog/frontdoor/index/index/docId/4129>.
- Lankton, Nancy, D. Harrison McKnight, and Jason Bennett Thatcher. 2014. “Incorporating Trust-in-Technology into Expectation Disconfirmation Theory.” *The Journal of Strategic Information Systems* 23 (2): 128–45. <https://doi.org/10.1016/j.jsis.2013.09.001>.
- Lebbie, Tamba S., Omosehin D. Moyebi, Kwadwo Ansong Asante, Julius Fobil, Marie Noel Brune-Drisse, William A. Suk, Peter D. Sly, Julia Gorman, and David O. Carpenter. 2021. “E-Waste in Africa: A Serious Threat to the Health of Children.” *International Journal of Environmental Research and Public Health* 18 (16): 8488. <https://doi.org/10.3390/ijerph18168488>.
- Loeser, Fabian. 2013. “Green IT and Green IS: Definition of Constructs and Overview of Current Practices.” *AMCIS 2013 Proceedings*. <https://aisel.aisnet.org/amcis2013/GreenIS/GeneralPresentations/4>.
- Maguire, Martin. 2001. “Methods to Support Human-Centred Design.” *International Journal of Human-Computer Studies* 55 (4): 587–634. <https://doi.org/10.1006/ijhc.2001.0503>.
- Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston, MA: Houghton Mifflin Harcourt.
- Mizzoni, John. 2017. *Ethics. The Basics*. Second edition. Malden, MA: Wiley Blackwell. <https://www.wiley.com/en-am/Ethics%3A+The+Basics%2C+2nd+Edition-p-9781119150701>.
- “Montréal Declaration for a Responsible Development of Artificial Intelligence.” 2018a. Montréal, QC: Université de Montréal. <https://www.montrealdeclaration-responsibleai.com/reports-of-montreal-declaration>.
- “———.” 2018b. Montréal, QC: Université de Montréal.
- Murugesan, San. 2008. “Harnessing Green IT: Principles and Practices.” *IT Professional* 10 (1): 24–33. <https://doi.org/10.1109/MITP.2008.10>.
- Noel-Brune, Marie, Fiona C. Goldizen, Maria Neira, Martin van den Berg, Nancy Lewis, Malcolm King, William A. Suk, David O. Carpenter, Robert G. Arnold, and Peter D. Sly. 2013. “Health Effects of Exposure to E-Waste.” *The Lancet Global Health* 1 (2): e70. [https://doi.org/10.1016/S2214-109X\(13\)70020-2](https://doi.org/10.1016/S2214-109X(13)70020-2).
- Nowviskie, Bethany. 2015. “Digital Humanities in the Anthropocene.” *Digital Scholarship in the Humanities* 30 (suppl\_1): i4–15. <https://doi.org/10.1093/llc/fqv015>.
- Nuta, Florian Marcel. 2015. “An Essay on Sustainable Development Moral.” *The Journal of Accounting and Management*, no. 3: 95–96.
- OECD. 2013. *OECD Guidelines on Measuring Subjective Well-Being*. Paris: Organisation for Economic Co-operation and Development. <https://www.oecd-ilibrary.org/economics/oecd-guidelines-on-measuring-subjective-well->

- being\_9789264191655-en.
- Osibanjo, O., and I.C. Nnorom. 2007. "The Challenge of Electronic Waste (e-Waste) Management in Developing Countries." *Waste Management & Research* 25 (6): 489–501. <https://doi.org/10.1177/0734242X07082028>.
- Ricoeur, Paul. 1990. "Éthique et Morale [Ethics and Morality]." *Revista Portuguesa de Filosofia* 46 (1): 5–17.
- Ricœur, Paul. 2000. "De La Morale à l'éthique et Aux Éthiques [From Morality to Ethics]." In *Un Siècle de Philosophie. 1900-2000*, 101–20. Paris: Gallimard.
- Schwartz, Franklin W., Sangsuk Lee, and Thomas H. Darrah. 2021. "A Review of Health Issues Related to Child Labor and Violence within Artisanal and Small-Scale Mining." *GeoHealth* 5 (2): e2020GH000326. <https://doi.org/10.1029/2020GH000326>.
- Taylor, Paul W. 1981. "The Ethics of Respect for Nature." *Environmental Ethics* 3 (3): 197–218. <https://doi.org/10.5840/enviroethics19813321>.
- . 2011. *Respect for Nature. A Theory of Environmental Ethics*. 25th-anniversary edition ed. Princeton, NJ; Oxford, UK: Princeton University Press. <https://press.princeton.edu/books/paperback/9780691150246/respect-for-nature>.
- The IEEE Global Initiative. 2019. "Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems." Piscataway, NJ: IEEE.
- Tronto, Joan C. 1998. "An Ethic of Care." *Generations: Journal of the American Society on Aging* 22 (3): 15–20.
- Yampolskiy, Roman V. 2021. "The Uncontrollability of Artificial Intelligence." IAI TV - Changing How the World Thinks. March 18, 2021. <https://iai.tv/articles/the-hard-problem-of-ai-safety-auid-1773>.
- Zack, Naomi. 2010. *Ethics for Disaster*. Lanham, MA: Rowman & Littlefield Publishers. <https://rowman.com/ISBN/9780742564954/Ethics-for-Disaster>.
- Zhang, Lingen, and Zhenming Xu. 2016. "A Review of Current Progress of Recycling Technologies for Metals from Waste Electrical and Electronic Equipment." *Journal of Cleaner Production* 127 (July): 19–36. <https://doi.org/10.1016/j.jclepro.2016.04.004>.