UNIVERSITY OF PIRAEUS

DEPARTMENT OF DIGITAL SYSTEMS

Postgraduate Programme

CLIMATE CRISIS AND INFORMATION AND COMMUNICATION TECHNOLOGIES

2021-2023

MSc Thesis

Ifigeneia Tsakalogianni (ID MKK2147)

THE TWIN TRANSITION: MAPPING THE USE OF ARTIFICIAL INTELLIGENCE IN SUSTAINABILITY

The role of Law and Ethics

<u>**Supervised by:**</u>

Professor Stefanos Gritzalis

Piraeus, July 2023

Δήλωση Πνευματικών Δικαιωμάτων

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα Διπλωματική Εργασία με τίτλο:

"THE TWIN TRANSITION: MAPPING THE USE OF ARTIFICIAL INTELLIGENCE IN SUSTAINABILITY - The role of Law and Ethics (Η Δίδυμη Μετάβαση: Χαρτογράφηση της χρήσης της Τεχνητής Νοημοσύνης στη Βιωσιμότητα - Ο ρόλος της Νομοθεσίας και της Δεοντολογίας)"

καθώς και τα ηλεκτρονικά αρχεία και οι πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στο πλαίσιο αυτής της εργασίας και αναφέρονται ρητώς μέσα στο κείμενο που συνοδεύουν και η οποία έχει εκπονηθεί στο Τμήμα Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς την συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν την συγγραφέα και μόνο.

Abstract

The purpose of the present thesis is to map the most important AI governance and policy implications along with the legislation produced, in order to provide a thorough analysis of the ethical and legal concerns that arise with mass production and utilisation of AI, and, finally, to examine the interplay between AI and environmental sustainability. Section 1 serves as an introduction to AI itself and an overview of the branches and approaches of AI, while Section 2 focuses on the regulatory and ethical implications of AI use. In Section 3, the risks and the benefits of harnessing the power of AI towards green transition are highlighted and practical examples of environmentally sustainable AI are given. It is concluded that if properly governed, digitalisation through AI could make a decisive contribution to the establishment of a sustainable, climate neutral economy and society. To this end, certain recommendations are made in Section 4 towards an ethical and legal nexus between AI and sustainability, which include harnessing the potential of regulatory sandboxes, ensuring climate mainstreaming into AI, promoting cross-sectoral consistency of policies and fostering collaboration between all actors of economy and society.

Acknowledgements

I would like to express my gratitude towards the individuals who have significantly contributed to the successful completion of my thesis. Firstly, I want to extend my heartfelt appreciation to my supervisor Mr. Gritzalis for his guidance and insightful input, as well as to the professors who inspired me every step of the way. I would like to express my sincere gratitude especially towards Mr. Maniatis, the Director of the Programme, for his unwavering enthusiasm and passion for the first Postgraduate Programme in Greece focusing on Climate Change and ICT, the two challenges of modern society.

Equally deserving of recognition is my mother, whose unwavering support and endless encouragement played an integral role in enabling me to pursue this academic endeavour. Furthermore, I am indebted to my partner, whose faith in my capabilities and constant reassurance cannot be overstated. As I wish to embark on more educational endeavors, I request their continued patience and understanding.

Table of Contents

Acronyms and abbreviations

AI	Artificial Intelligence
AVs	Autonomous Vehicles
DP	Deep Learning
EU	European Union
GAI	Generative Artificial Intelligence
GPAI	General Purpose Artificial Intelligence
ICT	Information and Communication Technologies
LLM	Large Language Model
ML	Machine Learning
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
TFEU	Treaty on the Functioning of the European Union

"…one cannot resist the invasion of ideas"

- Victor Hugo, The History of a Crime, 1877

Introduction

Climate change - also widely referred to as "***climate crisis***" or ***"climate breakdown"*** - could potentially lead to the collapse of our civilisations and the extinction of "*much of the natural world*"[1]. The UNFCCC gives us the definition of climate change as "*a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable time periods*" [2]. Anthropogenic emissions of Greenhouse Gas (GHG) that have been detected during the most recent decades are the highest in history, and it is now scientifically proven that "*the atmosphere and ocean have warmed, the amounts of snow and ice have diminished, and sea level has risen*"[3]. Continued emission of greenhouse gases will cause further warming and long-lasting changes in all components of the climate system, increasing the likelihood of severe, pervasive and irreversible impacts for people and ecosystems[4].

As of 2023, global GHG emissions have continued to increase, with unequal historical and ongoing contributions arising from unsustainable energy use, land use and land-use change, lifestyles and patterns of consumption and production across regions, between and within countries, and among individuals[5]. Ecosystems and biodiversity are at risk as heat waves will occur more often and last longer, extreme precipitation events will become more intense and frequent, phenomena like ocean acidification, sea level rise, floods, landslides, air pollution, drought, desertification, water scarcity, salinization of groundwater and soil, changes in ocean currents and the emergence of a huge variety of tropical diseases are among the most serious consequences, in high scientific certainty.

In 1979, the First World Climate Conference issued a declaration calling on the world's nations "*to foresee and prevent potential man-made changes in climate that might be adverse to the well-being of*

[1] Sir David Attenborough (2018, December 3) during COP24 in Katowice, Poland.
[2] United Nations, UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE [1992] art 1, para 2.
[3] IPCC, 'Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change' [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)] (IPCC, 2014) Geneva, Switzerland, p. 40.
[4] *Ibid*, p. 56.
[5] IPCC (2023) Sixth Assessment Report, Synthesis Report (SYR), Summary for Policymakers, p. 4.

humanity".[6] Since then, similar alarms have been made from the 1992 Rio Summit[7] to the 2015 Paris Agreement[8]; and yet, greenhouse gas (GHG) emissions have been steadily increasing[9]. Concurrently with the adoption of binding, hybrid and soft law legal instruments with the aim to crystallize the "legalities" of climate policy[10], talk on accelerating climate change mitigation, i.e., reduction of GHG emissions, and adaptation, i.e., proper preparation for unavoidable consequences, has increased exponentially in recent years. As a pioneering actor in the global sphere, the EU announced the EU Green Deal in late 2019[11], which sought to accelerate and underpin the transition needed in all sectors, namely: energy, mobility, industry (steel, chemicals, cement), agriculture, construction, biodiversity protection and restoration. Afterwards, the 2021 EU Climate Law[12] wrote into law the goal for Europe's economy and society to become climate-neutral by 2050 and to reduce net GHG emissions by at least 55% by 2030, compared to 1990 levels.

In this vein, harnessing the climate change-deceleration and -adaptation potential of **Artificial Intelligence (AI) technologies** has been suggested. AI-based technologies are among the most transformative tools that exist today; as climate change progresses, green AI alternatives to replace technologies with high environmental impact are demanded, while AI has been hailed as one of the solutions to solve major global environmental crises. Machine Learning (ML) models have been developed that forecast floods and allow farmers to detect, identify and predict agricultural pests

---

[6] World Meteorological Organisation, World Climate Conference-1 (12-23 February 1979; Geneva, Switzerland), https://library.wmo.int/doc_num.php?explnum_id=3778, 12.

[7] Rio Declaration on Environment and Development, Jun. 13, 1992. 31 ILM 874 (1992)

[8] Conference of the Parties, Adoption of the Paris Agreement, Dec. 12, 2015. U.N. Doc. FCCC/CP/2015/L.9/Rev/1 (Dec. 12, 2015).

[9] See William F Lamb et al. (2021) Environmental Research Letters, 16 073005, DOI: 10.1088/1748-9326/abee4e

[10] Namely, the UNFCCC contains an ambiguous commitment to return to 1990 emissions levels by the year 2000, whereas the Paris Agreement provides that nations submit non-binding Nationally Determined Contributions to state their national climate ambition. See Urs Luterbacher and Detlef Sprinz (eds.), *International Relations And Climate Change* (PIK Report 21, 1996) https://www.pik-potsdam.de/en/output/publications/pikreports/.files/pr21.pdf, 22 and Daniel Bodansky, "The Legal Character Of The Paris Agreement" *RECIEL*, 25 (2016): 142-150, https://doi.org/10.1111/reel.12154

[11] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS The European Green Deal, COM/2019/640 final

[12] Regulation (EU) 2021/1119 of the European Parliament and of the Council of 30 June 2021 establishing the framework for achieving climate neutrality and amending Regulations (EC) No 401/2009 and (EU) 2018/1999 ('European Climate Law').

and diseases[13], sound recognition models identify potential poachers and illegal deforestation in tropical forests[14], autonomous electrified vehicles are put on the market, building climate-proofing and AI urban optimization is studied, AI is used to discover new efficient ways to store and use renewable energy and low-carbon electricity systems are enabled *en masse.*

While originally confined to theoretical, technical and academic debates, the issue of governing AI has recently entered the mainstream with both governments and private companies formulating statements and policies regarding AI and ethics[15]. Without prejudice to the true promise of AI for climate and society, there are certain risks and challenges AI poses. They include, *inter alia*, breaches of privacy[16], cases of AI-bias, spreading of fake news, undermining basic human rights through facial recognition technology and general ethical concerns on controlling AI. As regards climate change, AI is no silver bullet; for instance, AI is used extensively within the fossil fuel industry such as during exploration and extraction[17], some ML models are energy-intensive to train and run[18] and "green" autonomous vehicles are the subject of various moral dilemmas regarding safety and their operation independently of human intervention.

However, the so far untapped opportunity for technology, and AI specifically, to drive environmental and climate sustainability has generated talks regarding a "**twin transition**"; The term refers not only to the concurrent digital and green transitions, but also to uniting the two, which could accelerate necessary changes and bring societies closer to the level of transformation needed[19].

---

[13] Domingues, T., Brandão, T., & Ferreira, J. C. (2022). Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey. Agriculture, 12(9), 1350. https://doi.org/10.3390/agriculture12091350

[14] Organizations such as the "Rainforest Connection" at: https://rfcx.org/guardian.

[15] Daly, A., Hagendorff, T., Hui, L., Mann, M., Marda, V., Wagner, B., & Wang, W. (2021). AI, Governance and Ethics: Global Perspectives. In H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, & G. De Gregorio (Eds.), Constitutional Challenges in the Algorithmic Society (pp. 182-201). Cambridge: Cambridge University Press. doi:10.1017/9781108914857.010, p. 183.

[16] A recent ban of advanced AI chatbot ChatGPT was ordered in Italy, at: https://www.bbc.com/news/technology-65139406.

[17] See Gupta, D., & Shah, M. (2022). A comprehensive study on artificial intelligence in oil and gas sector. Environmental science and pollution research international, 29(34), 50984–50997. https://doi.org/10.1007/s11356-021-15379-z.

[18] See Lacoste, A., Luccioni, A., Schmidt, V., and Dandres, T. (2019). *Quantifying the carbon emissions of machine learning.* Preprint arXiv:1910.09700 (2019), https://doi.org/10.48550/arXiv.1910.09700.

[19] Muench, S., Stoermer, E., Jensen, K., Asikainen, T., Salvi, M. and Scapolo, F. (2022). Towards a green and digital future, EUR 31075 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-52451-9, doi:10.2760/977331, JRC129319.

While the digital and green transitions are different in nature and each subject to specific dynamics, their twinning – i.e., their capacity to mutually reinforce each other – deserves closer scrutiny[20]. Even more so, if properly governed, digitalisation through AI could make a decisive contribution to the establishment of a climate neutral economy and society; hence, the "sweet spot" where digitalisation amplifies sustainability should be sought. But that is no easy feat, as the training and use of large-scale AI systems can have significant environmental footprints from energy and water use, GHG emissions and end-of-life considerations (non-recyclability of materials) and can also raise issues of ethical and regulatory, otherwise legal, nature.

The purpose of the present thesis is to map the most important AI governance and policy initiatives at an international and European level along with the legislation produced, to provide a thorough analysis of the ethical concerns that arise with mass production and utilisation of AI, and, finally, to examine the interplay between AI and environmental policy. More specifically, Section 1 serves as an introduction to AI and a thorough overview of the branches and approaches of AI, while Section 2 focuses on the regulatory and ethical implications of AI use. In Section 3, the risks and the benefits of harnessing the power of AI towards green transition are highlighted, while practical examples of environmentally sustainable AI utilisation are given. Finally, in Section 4 certain observations and recommendations are made towards ensuring an ethical, fair and legal nexus between AI and the "greening" of policy, society and economy.

[20] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL, 2022 Strategic Foresight Report, Twinning the green and digital transitions in the new geopolitical context, COM (2022) 289 final.

SECTION 1: Artificial Intelligence: a very real challenge of our time

1. Artificial Intelligence in a changing world

1.1. Introduction to AI

In the year 2023, the development of algorithms and AI technologies is becoming ubiquitous and omnipresent[21]; from Apple's Siri Virtual Assistant to Tesla cars' autosteer function, algorithms are now used extensively in various areas of modern life; without, of course, ceasing to be a source of various concerns and controversy, from plagiarism[22], unemployment boost due to automation and safety concerns on AI weapon systems[23], to the everyday constant smartphone use that kills the quality of the "art of human interaction".

In 1950, Alan Turing posed the question on whether machines can really think[24], and more specifically, whether they are able to conceal their artificial nature from a human through conversation. Turing's "Imitation Game", was the first test of a machine's ability to exhibit intelligent behaviour equivalent or indistinguishable from that of a human[25]. Such a test would then mean, for example, that a human could no longer determine whether a conversation partner on the telephone, on a virtual platform or during an online chat[26] is a human or software. Even though Turing did not have access to the resources needed

---

[21] H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, & G. De Gregorio (Eds.), Constitutional Challenges in the Algorithmic Society, Cambridge: Cambridge University Press. doi:10.1017/9781108914857.010, p. 28.
[22] Fazackerley, A. (2023) "AI makes plagiarism harder to detect, argue academics – in paper written by chatbot",), The Guardian, at: https://www.theguardian.com/technology/2023/mar/19/ai-makes-plagiarism-harder-to-detect-argue-academics-in-paper-written-by-chatbot.
[23] On Lethal Autonomous Weapons (LAWs) see Coyne, C., Yahya A. (2021). "Perverse Consequences of Lethal Autonomous Weapons Systems". Peace Review. 33 (2): 190–198. doi:10.1080/10402659.2021.1998747.
[24] Turing, A. (1950) "Computing machinery and intelligence", in Parsing the Turing Test, Springer, Dordrecht.
[25] Turing asks: "*Let us fix our attention on one particular digital computer C. Is it true that by modifying this computer to have an adequate storage, suitably increasing its speed of action, and providing it with an appropriate programme, C can be made to play satisfactorily the part of A in the imitation game, the part of B being taken by a man*?", A. M. Turing .(1950) "I.—Computing Machinery and Intelligence", Mind, Volume LIX, Issue 236, 1950, Pages 433–460, https://doi.org/10.1093/mind/LIX.236.433, p. 442.
[26] In November 2022 OpenAI released an AI-powered chatbot named ChatGPT "*which interacts* [with the human user] *in a conversational way*". See https://openai.com/blog/chatgpt.

to materialize his ideas, his essay "*Computing Machinery and Intelligence*" paved the way towards the evolution of AI and its integration to the entire spectrum of human activity.

Not long after Turing's death, the ELIZA natural language processing (NLP) chatbot was created by MIT's computer scientist Joseph Weizenbaum, stunning the world with its convincing ability to stimulate human conversation[27]. Since then, AI has been utilized in the most important sectors of human society, including healthcare, policing, public surveillance, infrastructure, manufacturing, transport and marketing.

AI-based systems can be software-only operating in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and facial recognition systems) or AI can be embedded in software devices (e.g., advanced robots, autonomous cars or Internet of Things applications). Generally, AI deals with the construction of intelligent behaviour systems that have the ability to analyse their environment and act - with varying degrees of autonomy - to achieve specific goals. They are machines classified as "intelligent", otherwise capable of performing "anthropomorphic" tasks after receiving specific training and learning:

(a) to adapt to new inputs, and

(b) to perform tasks usually assigned and attributed to humans.

The term "Artificial Intelligence" was coined at a workshop at the Dartmouth Summer Research Project on AI organized by computer and cognitive scientist John McCarthy in 1956, who is credited as the first user of the term in the proposal he co-authored with Marvin Minsky, Nathaniel Rochester and Claude Shannon[28], defining AI as "*the science and engineering of making intelligent machines, especially intelligent computer programs*". Thus, "AI" could be perceived as an umbrella term for a range of algorithm-based technologies and approaches that attempt to mimic human way of thinking, and, more importantly, the human ability to take decisions and to act on them *autonomously*.

---

[27] Weizenbaum, J. (1966) "ELIZA—a computer program for the study of natural language communication between man and machine". Communications of the ACM, 9, 1 (Jan. 1966), 36–45. https://doi.org/10.1145/365153.365168.
[28] McCarthy, J., Minsky L., M., Rochester, N., and Shannon E., C., (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," August 31, 1955, at: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

In recent years AI has become a key technology, offering unprecedented opportunities to increase the efficiency of conventional processes and optimize their quality through intelligent analysis and processing of large amounts of data. Therefore, it is not surprising that even early on, certain regulating or stricter voices were heard that asked for the setting of rules regarding AI use or for, at the very least, the observation of certain axioms. Science fiction author Isaac Asimov was a pioneer on the subject of AI principles, providing basic "guidelines" for robots in his 1942 short story "Runaround" (included in the 1950 collection I, Robot) and creating the so-called "The Three Laws of Robotics" or Asimov's Laws:

The Three Laws of Robotics		
1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.	**2.** A robot must obey orders given it by human beings except where such orders would conflict with the First Law.	**3.** A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Figure 1

Asimov, I. (1950). I, Robot. Garden City, N.Y., Doubleday

The most fundamental first Law enshrines the *sine qua non* of AI utilization: that no harm may come to a human being by an artificial entity by either intervention or withholding of care. Without delving into the ethical implications of omission bias[29], Asimov aimed to establish that AI should be incompatible with the occurrence of damage by its own "hand" or by its inability or "unwillingness" (?) to help a human. It follows that AI agents must obey the commands of their operator (2nd Law), except when they concern physical or other harm towards another human, in breach of the 1st Law. Lastly, the acquisition of a "survival instinct" by an AI entity is allowed; however, that is no horizontal rule, but is limited to the compliance with the 1st and 2nd Laws.

Hence, the protection from the possibility of the evolvement of AI in human harm is, potentially, the priority of all AI regulation attempts. Essentially, the basic steps are:

a) ensuring that AI does not turn against humanity or, by extension to the above, against universal goods like the environment and all living beings,

b) formulating a set of rules of a certain degree of ethical and legal responsibility and transparency, and

c) installing a "safety net", otherwise a trustworthy system that allows a "disconnection" or a shutdown of an AI system in case of emergency.

This "emergency shutdown" option has been fuelled by various dystopian scenarios involving AI and robots both in literature and films. Utopian and dystopian scenarios regarding AI in the future appear in opposition, but one thing is certain: they share a foundation in the idea that AI will have a substantial impact on society[30].

[29] See Yeung, Siu Kit; Yay, Tijen; Feldman, Gilad (2021). "Action and Inaction in Moral Judgments and Decisions: Meta-Analysis of Omission Bias Omission-Commission Asymmetries". Personality and Social Psychology Bulletin. 48 (10): 1499–1515. doi:10.1177/01461672211042315 and Woollard, Fiona and Frances Howard-Snyder, "Doing vs. Allowing Harm", The Stanford Encyclopedia of Philosophy (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.).
[30] Boucher, P. (2020) "*Artificial intelligence: How does it work, why does it matter, and what can we do about it*?", Panel for the Future of Science and Technology EPRS, European Parliamentary Research Service, p. 29.

1.2. A definition for AI?

In essence, AI is a branch of computer science that deals with the simulation of intelligent behaviour. In particular, AI aims to create systems that think or behave like humans or with the maximum possible reasoning ability; thus, it could be defined as the result of scientists' curiosity to decode the human brain and explore the concept of consciousness by eventually "copying" or mapping it into an artificial, otherwise man-made, system. In the words of John McCarthy, [AI] "*is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable*"[31]. In contrast to the congenital intellectual mechanisms of humans, the abilities of AI are dependent on the intellectual mechanisms created by programmers[32].

One common thread in definitions of AI is that an autonomous agent executes or recommends actions[33,34]. Many definitions in literature consider AI's *autonomy* as the unique feature that separates it from other technologies; a feature which leads to a considerable number of legal implications, as discussed in Section 2. Other approaches tend to attribute human or human-like characteristics to AI in order to define it, like the ability to learn and to adapt as well as varying degrees of self-awareness and consciousness.

It is appropriate to understand AI as a collective term that conveys a variety of methods using statistical and mathematical models that mimic intelligent or cognitive functions. AI has been named as "*the ability of a system to identify, interpret, make inferences, and learn from data to achieve predetermined organizational and societal goals*[35]". Cambridge Dictionary takes a rather AI-as-a-scientific-discipline approach, as AI is "*the study of how to produce machines that have some of the qualities that the human mind has, such as the ability*

---

[31] McCarthy, J., "What is Artificial Intelligence?" (12 November 2007), at: www.formal.stanford.edu/jmc/whatisai.pdf, p. 1.
[32] *Id*., p. 4.
[33] Computational Intelligence, A Logical Approach, David Poole Alan Mackworth Randy Goebel New York Oxford Oxford University Press, 1998.
[34] Artificial Intelligence, A Modern Approach Fourth Edition, Stuart J. Russell and Peter Norvig, Pearson, [2021] Series: Pearson series in artificial intelligence.
[35] Mikalef, P., Gupta, M., (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance, Information & Management, Volume 58, Issue 3, 2021, https://doi.org/10.1016/j.im.2021.103434.

to understand language, recognize pictures, solve problems, and learn", while Merriam-Webster Dictionary provides two definitions of AI as "*1. a branch of computer science dealing with the simulation of intelligent behaviour in computers,* and "*2. the capability of a machine to imitate intelligent human behaviour*".

And as the term "intelligence" becomes very difficult to define, the actual definition of AI is, naturally, correspondingly difficult and numerous definitions can be found in literature[36], some of which use the notion of "rationality" instead. As Russell and Norvig put it, "*intelligence is concerned mainly with rational action. Ideally, an intelligent agent takes the best possible action in a situation*[37]"[38]. They provided a categorisation of a human-centred and a rationalist approach as regards AI definitions, as follows:

[36] See Burgard, W. (2022). Artificial Intelligence: Key Technologies and Opportunities. In S. Voeneky, P. Kellmeyer, O. Mueller, & W. Burgard (Eds.), The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives (Cambridge Law Handbooks, pp. 11-18). Cambridge: Cambridge University Press. doi:10.1017/9781009207898.003, p. 11.
[37] S. Russell, S., Norvig, P. (2009). "Artificial Intelligence: A Modern Approach", Prentice Hall, 3rd edition, 2009, p. 30.
[38] Russell and Norvig also provided a definition of a "rational agent": "*For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.*", *id.*, on p. 37.

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985)
"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998)
"The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Figure 2

S. Russell, S., Norvig, P. (2009). "Artificial Intelligence: A Modern Approach", Prentice Hall, 3rd edition, 2009, p. 2.

It is no coincidence that, so far, there is neither a binding common AI definition at international level, nor legally binding (regulatory) rules on AI use.

In 2018, the possibility of utilization of AI systems by the judiciary led to the European Ethical Charter on the Use of AI in Judicial Systems and their environment by the European Commission for the Efficiency of Justice, which defined AI as "*a set of scientific methods, theories and techniques whose aim is to reproduce, by a machine, the cognitive abilities of human beings. Current developments seek to have machines perform complex tasks previously carried out by humans*[39]".

[39] Council of Europe, European Commission for the Efficiency of Justice (2018) "European Ethical Charter on the Use of AI in Judicial Systems and their environment", as adopted at the 31st plenary meeting of the Commission, p. 70.

AI applications often impact human behaviour and evolve dynamically in ways that are at times unforeseen by the systems' designers[40]. Therefore, an "umbrella" term could prove to be more useful than definitions that seem too restrictive or narrowly interpreted. A solid counterpoint though is that circular definitions provide little to no functional use; as per the common quote by Albert Einstein, "*the formulation of a problem is often more essential than its solution*"; through an analogy and definitely without identifying AI as a "problem", omitting to properly define AI or failing to provide a sufficient definition could be perceived as a serious indication of regulatory uncertainty as regards AI; which in turn could further make the task of clarifying the grey areas of AI more daunting.

1.3. The AI technical landscape

1.3.1. How does AI work?

As AI permeates societies, it becomes all the more imperative to consider how and why AI systems work. Not from a computer scientist's but rather a legal scholar's perspective such as the writer, AI works with sensors, much like Wi-Fi or a motion detection camera. Those sensors are programmed through algorithms to detect various data to model some aspect of the world, which are then processed through said algorithm.

As to what regards the collected data, it is often useful to distinguish between structured and unstructured data. Structured data is data that has been predefined and formatted to a set structure before being placed in data storage. The best example of management of structured data is Structured Query Language (SQL), a language originally built by IBM in 1974 to manage structured data within a relational database and one of the world's most popular programming languages. Structured data can

[40] Gasser, U., Virgilio, A. (2017). "A Layered Model for AI Governance." IEEE Internet Computing 21 (6) (November): 58–62. doi:10.1109/mic.2017.4180835.

be defined as data that is simple enough for human understanding both in volume and structure[41]. The advantage of it is obvious: the organised and quantitative nature of standardised data can easily be analysed by AI; however, its inflexibility and non-adaptability renders it unsuitable for extended or more complicated usage than its intended purpose. On the other hand, unstructured, otherwise qualitative data is stored in its native format, e.g., video files, surveillance imagery, social media activity and engagement etc., and can often provide a deeper understanding of the environment in which they were generated.

In an AI system, an algorithm could be perceived as a finite sequence of operations, rules and instructions, making it possible to obtain a result from an initial input of data, which should be relevant to the goal given to the AI system by its programmer. The algorithmic steps frequently involve repetition of an operation[42]. The data collected by the sensors are transformed into information that is processed by an assigned AI module, which operates according to the assigned algorithm. At the final stage, an external output is produced, which translates into:

- **(a)** a certain action, as the automatic steering of a car wheel through the AI system's actuators or effectors,
- **(b)** a classification, like flagging an email as spam, or
- **(c)** a recommendation, used in cases of medical diagnoses or suggested prognoses[43],
- **(d)** a prioritisation, which could prove useful in managing workflow, and
- **(e)** an association, used in market analysis in order to identify the relations between supply and demand through promotional pricing or product placement[44].

---

[41] Hekler EB, Klasnja P, Chevance G, Golaszewski NM, Lewis D, Sim I. (2019). Why we need a structured data paradigm. *BMC Med.* 17:133. doi: 10.1186/s12916-019-1366-x

[42] As per the definition of the term "algorithm" by the Merriam-Webster dictionary.

[43] See Rajpurkar, P., Chen, E., Banerjee, O. et al. (2022). AI in health and medicine. Nat Med 28, 31–38, https://doi.org/10.1038/s41591-021-01614-0, and Kather, J.N. (2023). Artificial intelligence in oncology: chances and pitfalls, J Cancer Res Clin Oncol, https://doi.org/10.1007/s00432-023-04666-6.

[44] See Kumar, B., Roy, S., Sinha, A., Iwendi, C., & Strážovská, Ľ. (2022). E-Commerce Website Usability Analysis Using the Association Rule Mining and Machine Learning Algorithm. Mathematics, 11(1), 25. https://doi.org/10.3390/math11010025

Therefore, the way an AI system works could, roughly and, maybe, quite simplistically, be schematically depicted as follows:

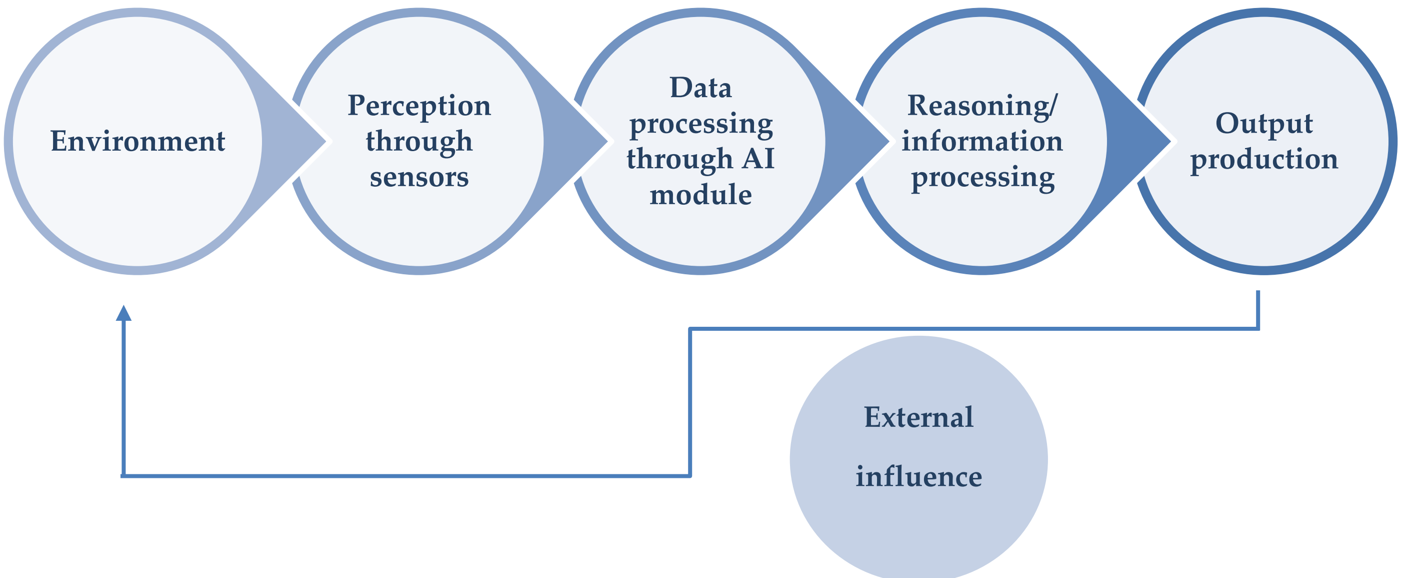


Figure 3

Created by the author

1.3.2. *The branches of AI*

In 1965, Gordon Moore, the co-founder of Intel, observed that the number of transistors on computer chips doubles approximately every two years. This was then known as "Moore's Law"[45]. In essence, Moore's Law was a scientific prediction or admission that technological growth would continue to grow rapidly or at least exponentially. Indeed, and without strictly confining Moore's Law within the electronic transistors' limits, the ever-growing computational capacity of computers was the contributing factor which led to the boom of AI technologies.

[45] Moore, G. (1965). Cramming more components onto integrated circuits, *Electronics*, Volume 38, Number 8.

As previously defined, AI is the man-made system which, based on the data and goals defined by a human, produces a specific result that is identifiable or observable by the human senses. By using an "umbrella" term, **AI is computer software that mimics human cognition through differing levels of autonomy in order to perform certain tasks of varying complexity**. As labyrinthine as AI may seem to a non-computer scientist, one may find that a categorisation of AI is a way to "unravel" this intriguing technical mystery.

1.3.2.1. Symbolic AI

Symbolic AI was the primary form of AI applications in the previous century. It is an approach that "trains" AI the same way human brain learns; that is, by forming internal symbolic representations, as symbols play a vital role in the human thought and reasoning process. In symbolic AI the "world" is presented through symbols and the processes or algorithms that programmers have created lead the system towards an "intelligent" output in a given situation. Symbolic AI's potential is quite limited, as it restricts itself between controlled environments that allow autonomy only in the way(s) indicated by the system's specific instructions; in other words, the rules used by the system are created through human intervention.

A restricted kind of symbolic AI is "Good Old-Fashioned Artificial Intelligence" (GOFAI) invented by the philosopher John Haugeland in 1985[46], which includes reasoning-based and logical agents. In GOFAI, the overall intelligence is explained by analysing the system into smaller (less intelligent) components, whose external symbolic moves and interactions are the larger system's internal reasonable cognitions[47]. Even though GOFAI is still used, the developments in AI in the recent years have halted its use on a larger scale.

[46] Haugeland, J. (1985). Artificial Intelligence: The Very Idea, Cambridge, Mass: MIT Press.
[47] *Id.*, p. 117.

1.3.2.2. Weak or Narrow AI

In the beginning of the third decade of the 21st century, the currently deployed AI systems are what is described as "narrow" or "weak" AI; that is, systems that can perform one or few of their assigned tasks, like a self-driving car or a chess-playing robot. Narrow AI constitutes a tool for the system to perform specific and limited calculations and is operated through Neural Networks (NN) and Machine Learning (ML), between others.

1.3.2.3. Strong or General AI

A more powerful form of AI is Strong or General AI, which resembles human intelligence by referring to algorithms in non-domain-specific problems. In other words, that kind of AI is stronger, more intelligent and significantly closer to the abilities of the human mind. The tasks it can learn to perform include perceptual and cognitive problems such as language processing and "thinking". However, as it does not yet exist, it belongs to the realm of speculative AI[48].

1.3.2.4. Artificial Superintelligence

A superior form of intelligence that surpasses even the human cognitive abilities is known as "Artificial Superintelligence". A concept mostly depicted in utopian/dystopian sci-fi scenarios, the moment when AI develops thinking on its own and exceeds the limitations of the human mind will most definitely be a turning point in human history. Artificial Superintelligence goes beyond just mimicking human behaviour; it involves developing actual consciousness, beliefs, desires and emotions. It is the result of an explosion of intelligence and technological progress that does not cease to spiral. Sceptics of that superior AI cite the loss of control over it as a reason for concern, as well as ethical problems like the

[48] Boucher, P. (2021) *Artificial intelligence: How does it work, why does it matter, and what can we do about it?*, Panel for the Future of Science and Technology EPRS, European Parliamentary Research Service, p. 13.

instilling of human personality in man-made creations, while others warn that AI might eventually decide it doesn't need humans at all[49].

By using so-called "recursive self-improvement" learning technique, which allows for autonomous repeated cycles of the AI system's processes in order for it to maintain continuous and rapid improvement of its intelligence (the so-called "Self-Improving AI"), Artificial Superintelligence will be continuously evolving and may further distance itself from its human creator – and from human control.

1.3.2.5. Singularity

Stephen Hawking, the renowned British physicist, once warned that "*the development of full artificial intelligence could spell the end of the human race*[50]". This prediction, albeit of a certain pessimistic nature, showcases the perhaps limitless potential of AI in the hands of the ambitious human race. The hypothesis that the development of robotics, AI and technology will become uncontrollable and surpass human abilities is called "Singularity". A concept of the distant future (?), Technological Singularity is the point in history when AI becomes *sentient* and breaks free of human control; at that point, AI could also be able to generate even more intelligent AI on each own, in a way, to reproduce and evolve into the next generations of AI. This has already happened on a much smaller and less significant scale, as many AI applications are developed and tuned by other AI applications[51] today. Yet, it seems that the myth behind "Singularity" is very far from materialising.

---

[49] Heaven D. W. (2021) *AI is learning how to create itself*, MIT Technology Review, at: https://www.technologyreview.com/2021/05/27/1025453/artificial-intelligence-learning-create-itself-agi/.

[50] Rory Cellan-Jones for BBC News (2014) *Stephen Hawking warns artificial intelligence could end mankind*, at: https://www.bbc.com/news/technology-30290540.

[51] Mikalef, P., Gupta, M., (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance, *Information & Management*, Volume 58, Issue 3, 2021, https://doi.org/10.1016/j.im.2021.103434, p. 3.

1.3.3. *Machine Learning and Deep Learning approaches*

AI, as we know it today, covers a game-changing set of capabilities as new technologies are developed. Technologies that come under the "umbrella" of AI include Machine Learning (ML) and Deep Learning (DL), which operate through various approaches. Machine Learning (ML), and Deep Learning (DL) are sometimes used interchangeably, but they are distinct terms. By using the definition adopted by the EU Draft AI Act (as it stood up until the European Parliament's final resolution in summer 2023), an AI system is developed with one or more of the techniques and approaches listed in its Annex I, which include "*Machine Learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning*". In general, ML is perceived as a sub-area of AI and DP is a subset of ML.

1.3.3.1. *Machine Learning*

ML is a set of techniques for automatically extracting patterns from data[52]. These patterns, or more precisely, mathematical models or statistical tools, incorporate a large number of variables not known in advance, as the algorithm learns rules as it establishes correlations between inputs and predicted output values. The parameters of the process are established during the "learning" phase of the system, which uses training data to be able to observe and classify the required links between the data, in order to be able to solve complex problems without human direction.

ML allows for specific techniques that automate the process of algorithmic training. By applying ML algorithms, a computer, with the help of training data, can learn rules and build a decision model. ML

[52] Lynn H. Kaack, Priya L. Donti, Emma Strubell, and David Rolnick. 2020. Artificial intelligence and climate change: Opportunities, considerations, and policy levers to align AI with climate change goals. Retrived from https://eu.boell.org/en/2020/12/03/artificial-intelligence-and-climate-change.

is thus an attractive alternative to manually constructing these programs[53]. The most prominent ML approaches are (human) supervised learning, unsupervised learning and reinforcement learning, which group together different methods including Neural Networks, DL, etc.

- *Supervised Learning*: Supervised learning models are trained on a provided dataset which contains labelled data, i.e., a large set of answers or solutions to a desired goal. Therefore, the system "learns" to map the variables onto the desired outputs and to produce new inputs based on the learned patterns. This kind of ML is used for data classification and regression, i.e., forecasting and predictive modelling[54].
- *Unsupervised Learning*: Unsupervised learning models are trained on a dataset without explicit external instructions or labelled data by identifying their similarities and incorporating them into patterns. This kind of ML approach is mainly used for clustering data – i.e., grouping together a set of elements that have relative similarities or connections (e.g., when classifying texts or terms of a related meaning) or for detecting glitches.
- *Reinforcement Learning*: Reinforcement learning models learn on the basis of their interactions with a dynamic environment and search for an optimal way to complete a task by taking a series of steps that maximise the probability of success. In the case of reinforcement learning, the system learns from the results of its actions through rewards (positive feedback or points) or "punishments" (negative feedback or loss of points), otherwise through trial and error. This kind of learning enhances systemic adaptability, as the system learns and improves itself while operating in an evolving environment.

[53] Buiten, M. (2019). Towards Intelligent Regulation of Artificial Intelligence. *European Journal of Risk Regulation*, 10(1), 41-59. doi:10.1017/err.2019.8, p. 49.
[54] See Shen Rong, Zhang Bao-wen (2018). The research of regression model in machine learning field, MATEC Web, Conf. 176 01033 (2018), doi: 10.1051/matecconf/201817601033

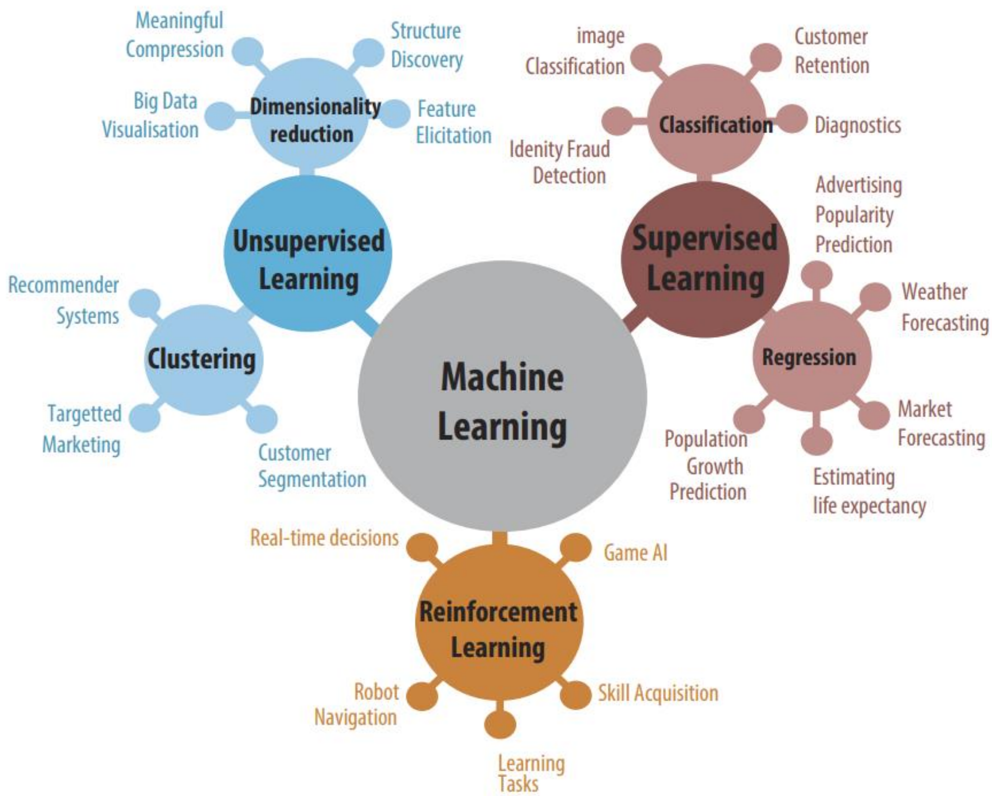


Figure 4

Council of Europe, European Commission for the Efficiency of Justice (2018) "European Ethical Charter on the Use of AI in Judicial Systems and their environment", as adopted at the 31st plenary meeting of the Commission, p. 72.

1.3.3.2. Deep Learning

DL, a subset of ML, is based on human understanding of how the brain is structured. DL could be defined as method of ML that uses several layers within a Multi Neural Network System to do some of the most complex ML tasks without human input; the aforementioned system is structured with Artificial Neural Networks (ANNs).

- *Artificial Neural Networks (ANNs):* In essence, ANNs are frameworks for many different ML algorithms to work together and process complex data inputs[55]. They are inspired by the functionality of the brain. Essentially, a neural network captures the relation between case features and responses through a set of nodes (called neurons) and weighted connections between them. Under some approaches, the system's responses can be evaluated, and based on this evaluation the system can self-update and gradually optimize itself[56]. In DL, the ANNs have several layers, through which the data is translated into signals in order to generate responses, or outputs, towards solving complex problems.

Application areas of ANNs include system identification and control (vehicle control, process control), game-playing and decision making (backgammon, chess, racing), pattern recognition (radar systems, face identification, object recognition, etc.), sequence recognition (gesture, speech, handwritten text recognition), medical diagnosis, drug detection, financial applications (loan risk estimation), quality analysis etc.[57].

[55] Buiten, M. (2019). Towards Intelligent Regulation of Artificial Intelligence. European Journal of Risk Regulation, 10(1), 41-59. doi:10.1017/err.2019.8, p. 49.

[56] Reichman, A., & Sartor, G. (2021). Algorithms and Regulation. In H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, & G. De Gregorio (Eds.), Constitutional Challenges in the Algorithmic Society (pp. 131-181). Cambridge: Cambridge University Press. doi:10.1017/9781108914857.009, p. 145.

[57] Maind, S., Wankar, P. (2014). Research Paper on Basic of Artificial Neural Network, International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 2 Issue: 1 96 – 100, p. 99.

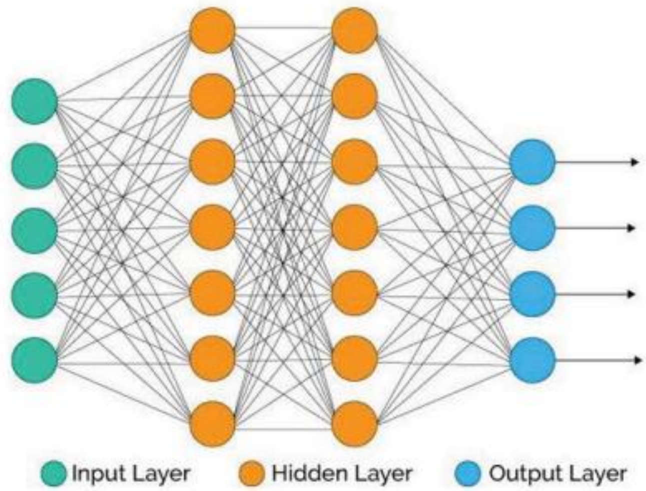


Figure 5 - A DL Artificial Neural Network

Gimenez, J.M., Bre, F., Nigro, N.M. et al. (2018). Computational modeling of natural ventilation in low-rise non-rectangular floor-plan buildings. *Build. Simul.* 11, 1255–1271. https://doi.org/10.1007/s12273-018-0461-9

1.3.4. Generative AI, General-Purpose AI and Large Language Models

Generative AI (GAI) refers to a form of AI that can generate, otherwise create, content in the form of images, text, video that is purely a new, AI-made creation. Rather than recognising patterns and providing outputs like classifications, predictions or prioritisations, this kind of AI becomes, in a way, the "creator" of new content that may be indistinguishable from the real data on which its learning model is based on. GAI uses a type of deep learning called generative adversarial networks (GANs) to create new content. A GAN consists of a generator NN that creates new data and a discriminator NN that evaluates the data and provides feedback in order for the former to improve itself over time.

GAI can create new images, write poetry and essays as well as generate new music, sound effects and even voice acting. The most famous example of GAI is ChatGPT, the AI chatbot launched by OpenAI (GPT stands for generative pretrained transformer)[58]. Similarly, AI-generated art models like DALL-E

[58] See https://openai.com/blog/chatgpt.

and DALL-E 2 (its name is a mash-up of the artist Salvador Dalí and the Pixar robot WALL-E) can create images, or art, on demand[59].

Generative AI belongs to a greater group of AI called **AI General-Purpose Artificial Intelligence** (GPAI) or a general class of models called *foundation models*. A foundation model, like the GPT chatbot, is any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks[60]; it is based on standard DP through training and fine-tuning and is enabled by *transfer learning*. The idea of transfer learning is to take the "knowledge" learned from one task (e.g., object recognition in images) and apply it to another task (e.g., activity recognition in videos)[61]. Up until early 2023, the Draft AI Act provided no binding provisions on GPAI systems, despite their continuous evolution on capabilities and size; scholars have characterized these technologies as having enough self-sufficiency to adapt to unknown environments and even can perform tasks that they were not originally trained for (due to a combination of factors such as the quantity of input data or model structure[62]), make decisions with limited resources, and function in complex domains that are currently reserved for humans due to the need for contextualisation[63].

GPAI Systems like ChatGPT operate on the so-called "**Large Language Models**" (LLMs); that kind of models consist of NN of billion parameters and are trained on billions of words in order to excel at not just one, but many tasks. Their skill, their range of tasks and their efficiency seems to be dependent on the amount and quality of the sources used to train them. Their training usually starts with unsupervised learning, as distributed training algorithms learn the network parameters of LLMs and various parallel strategies are jointly utilised to train the system on large-scale data (otherwise called "corpora")[64]. The

---

[59] See https://openai.com/product/dall-e-2.

[60] Bommasani, R. et al. (2021) *On the Opportunities and Risks of Foundation Models*, arXiv:2108.07258v3, https://doi.org/10.48550/arXiv.2108.07258, p. 3.

[61] *Id*., p. 4.

[62] Gutierrez, C., Aguirre, A., Uuk, R., Boine, C., and Franklin M., *A Proposal for a Definition of General Purpose Artificial Intelligence Systems*, Future of Life Institute Working Paper, at: https://futureoflife.org/wp-content/uploads/2022/11/SSRN-id4238951-1.pdf

[63] See Bieger, J., Thórisson, K. R., Steunebrink, B. R., Thorarensen, T., and Sigurdardottir, J. S. (2016). *Evaluation of general-purpose artificial intelligence: why, what & how*, Evaluating General-Purpose AI.

[64] Zhao, Z. W, Zhou, K. et. al (2023, ongoing work). A Survey of Large Language Models, arXiv:2303.18223v4, doi: 10.48550/arXiv.2303.18223 p. 3 – 4.

system's fine-tuning is a significant step, as human feedback is provided to align the LLM with ethical instructions[65].

When trained on massive datasets like nearly a terabyte of English text, their ability to generate fluent natural language is increased, which also allows them to be applied to a plethora of tasks like translation, question answering, text summarising and classification[66].

1.3.5. Explainable AI: The Black Box problem

Some AI approaches are very opaque in terms of understanding how they make decisions. Hence, the notion of "Black-Box AI" refers to scenarios when it is not possible to identify the reason for certain decisions made by AI. As explained by Riccardo Guidotti, "*black-box models are tools used by AI to accomplish a task for which either the logic of the decision process is not accessible, or it is accessible but not human-understandable*". In other words, it is a machine taking decisions over humans' lives without human oversight or awareness of the reasons behind those decisions[67].

In that regard, the need for AI **explainability**, i.e., the provision of or the ability to explain and retrace how an AI algorithm came to a result has emerged, creating the "**Explainable AI**" (XAI) notion. That is, an AI system that can provide adequate reasoning for its processes, acts and outputs, which, according to some, may ultimately be instrumental to achieving greater AI performance and trust through performing pre-hoc over post-hoc analysis[68]. A counter-argument of the necessity of XAI is that too

---

[65] In order to avoid AI "rudeness", sexist or racist comments. For example, see Wolf, B. Z. for CNN (2023), *AI can be racist, sexist and creepy. What should we do about it?* at: https://edition.cnn.com/2023/03/18/politics/ai-chatgpt-racist-what-matters/index.html

[66] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu J. P. (2019). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, *Journal of Machine Learning Research*, 21 (2020) 1-67, arXiv:1910.10683v3, doi: 10.48550/arXiv.1910.10683.

[67] Cappello M. (2020), *Artificial intelligence in the audiovisual sector*, IRIS Special, European Audiovisual Observatory, Strasbourg, p.1.

[68] Bordt, S., Finck, M., Raidl, E., von Luxburg, U. (2022) Post-Hoc Explanations Fail to Achieve their Purpose in Adversarial Contexts, FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, p. 891–905, doi: 10.1145/3531146.3533153 and McCoy, G., L., Brenna, C., Stacy S. C., Vold, K., Das, S., (2022). Believing in black boxes: machine learning for healthcare does not need explainability to be evidence-based, *Journal of Clinical Epidemiology*, Volume 142, p. 252-257, doi: /10.1016/j.jclinepi.2021.11.001.

much focus on explainability may hinder the improvement of AI's effectiveness, or that one should not prioritise the reason behind an AI's output over the output itself.

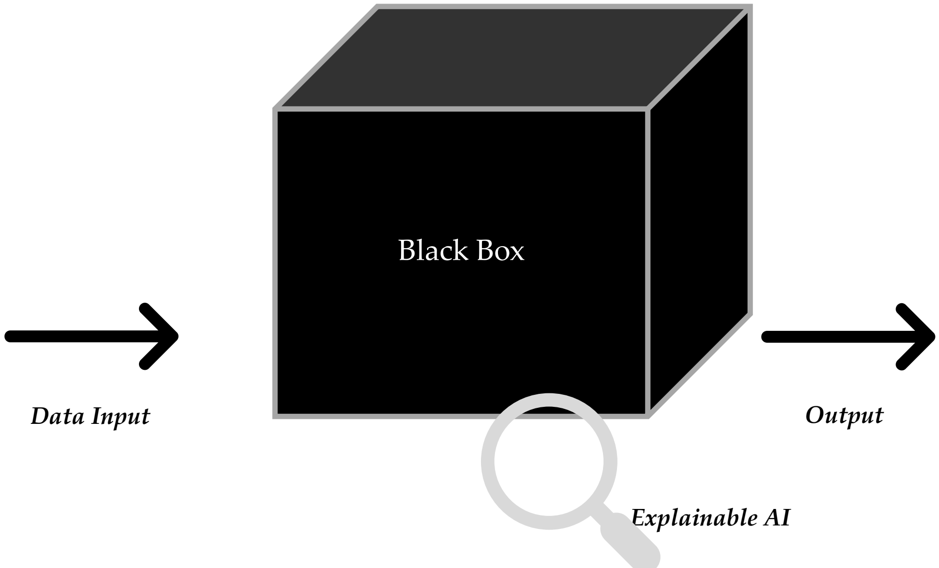


Figure 6

Created by the author

As scholars have put it, a socially [and legally] acceptable AI system would be one which would execute the algorithms in the best possible way (effectiveness), but will be able to explain its decision (explainability) by achieving absolute clarity (transparency) so that it can potentially be questioned (accountability). Further analysis on this complex issue is provided in Section 2.

1.3.6. The AI levels of automation

Decisions made using AI are either fully automated, or with a "human in the loop"; that is, the selective inclusion of human participation through the incorporation of human judgment into the AI system[69].

[69] See Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the People: The Role of Humans in Interactive Machine Learning. AI Magazine, 35(4), 105-120. https://doi.org/10.1609/aimag.v35i4.2513 and Mosqueira-Rey, E., Hernández-

Varying degrees of automation exist; each level corresponds to a gradually increasing degree of AI autonomy.

For example, autonomy in vehicles is often categorised into six levels, according to a system developed by SAE International[70], as revised periodically:

- *Level 0*: No driving automation. The driver is needed to control the speed and direction of the vehicle.
- *Level 1*: Driver support features provided (e.g., steering or acceleration/deceleration support or adaptive cruise control etc.).
- *Level 2*: Driver support features provided which may operate simultaneously (e.g., lane centring and adaptive cruise control at the same time).
- *Level 3*: The vehicle drives itself under limited conditions and the driver is expected to intervene if requested (e.g., in case of emergency).
- *Level 4:* The vehicle drives itself under limited conditions and the driver is not expected to intervene.
- *Level 5:* The vehicle drives itself and the driver is not expected to intervene.

Nowadays, most vehicles feature limited levels of autonomy corresponding to levels 1-2 above. However, automotive industry has announced the development of SAE Level 3 vehicles, while Mercedes-Benz recently got approval to deploy Level 3 Automated Driver Assist System (ADAS) in

Pereira, E., Alonso-Ríos, D. et al. Human-in-the-loop machine learning: a state of the art. Artif Intell Rev 56, 3005–3054 (2023). https://doi.org/10.1007/s10462-022-10246-w.

[70] SAE International J3016, International Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016:Sept2016), doi: https://doi.org/10.4271/J3016_202104. SAE International (Society of Automotive Engineers) is a global association of engineers and related technical experts in the aerospace, automotive and commercial-vehicle industries. The SAE system on autonomy has been also adopted by the USA Department of Transportation.

Nevada, USA, after ensuring compliance with the statewide legislation on autonomous vehicle technology[71].

Using the automotive autonomous SAE levels on a wider scale, **AI automation levels** could be depicted as follows:

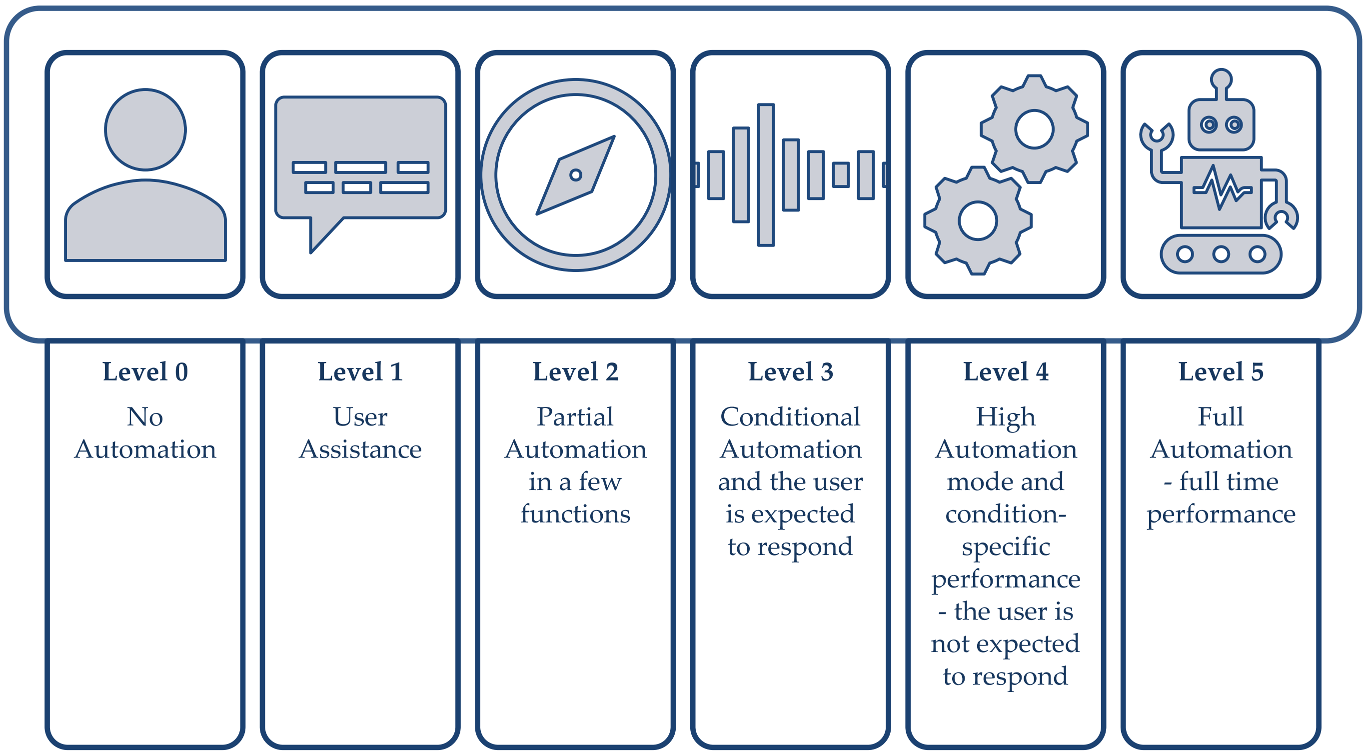


Figure 7

Created by the author

As it is expected, the more AI's autonomy increases, the more regulatory and ethical issues arise as regards the safety of the user and third parties, the reliability of the autonomous system and the liability

[71] Mercedes-Benz Group (2023). *Certification for SAE Level 3 system for U.S. market*, at: https://group.mercedes-benz.com/innovation/product-innovation/autonomous-driving/drive-pilot-nevada.html.

in case of accident, miscalculation or malfunction. The next Section puts AI through these exact regulatory/legal and ethical lenses.

SECTION 2: AI – is it legal? Can it be ethical?

2. Building a legal AI ecosystem

2.1. The technological pacing problem and the need to regulate AI

While some piecemeal regulation of specific AI applications and risks using traditional regulatory approaches may be feasible and even called for, AI has many of the characteristics of other emerging technologies that make them refractory to comprehensive regulatory solutions[72]; in the case of AI, "traditional" Law seems insufficient, as AI rises concerns of human-machine relationships, biased algorithms and pure human existential risks[73].

A decisive parameter is the constant evolution of AI technologies -a challenge known as the "pacing problem[74]". As G. Marchant put it, "*our traditional government oversight systems are mired in stagnation, ossification and bureaucratic inertia, and are seriously and increasingly lagging behind the new technologies accelerating into the future*"[75]. The pacing problem, as first described by L. Downes[76] refers to the difference between legal and technological progress; that is, the incremental growth of the former and the exponential evolution of the latter. As technology never stagnates, it manages to constantly outpace the ability of the Law to keep up. That means that the moment a certain piece of technological regulation comes into force, the technology attempted to be regulated will potentially have already further evolved, and thus, will be outside of the scope of said regulation. A similar notion is that of the Collingridge Dilemma. D. Collingridge found that technology is easy to influence and, thus, easy to control, when its consequences are not yet manifest; but at that stage, the technology's effect on society is widely

---

[72] Marchant E. G., Wallach, W. (2016). Introduction, in: *Emerging Technologies: Ethics, Law and Governance*, 1-12.
[73] Marchant, G. (2019). "Soft *Law" Governance of Artificial Intelligence*. UCLA: The Program on Understanding Law, Science, and Evidence (PULSE), p. 2.
[74] Marchant, G. (2011). Addressing the Pacing Problem, in: *The Growing Gap Between Emerging Technologies* and Legal-Ethical Oversight, doi: 10.1007/978-94-007-1356-7_13.
[75] *Id*.
[76] Downes, L. (2009). *The Laws of Disruption: Harnessing the New Forces that Govern Life and Business in the Digital Age*, Basic Books.

unknown. Vice versa, when the technology has become societally embedded, its implications are known but the development of said technology is difficult to change, or control[77]. As Collingridge put it: "*When change is easy, the need for it cannot be foreseen; when the need for change is apparent, change has become expensive, difficult and time consuming*"[78].

AI as a quasi-autopoietic tool[79] creates ethical, societal and political implications. This is precisely where the role of Law comes in. The great regulatory challenge lies in how to put AI at the service of humanity by regulating it without halting or discouraging technological progress and by creating a "regulatory toolbox" framework within which the development of AI can be, in a way, "value-driven". **That is, an amalgam of technological progress, social contribution, financial stability and optimised ethical and regulatory practice; and, as analysed in Section 3, of environmental sustainability** as well. At its heart, one approach of "AI and law" involves the application of computer and mathematical techniques to make law more understandable, manageable, useful, accessible, or predictable[80], as is the case when AI techniques are employed by the judiciary or administrators in decision-making, or even in cases where a "robot-judge" delivers justice[81]. Our focus is, however, is on the very regulatory core of Technology Law; more precisely, on the genuine need to regulate AI and the implications of this endeavour in order to tackle the contemporary quandaries that stem from AI evolution.

Some of that implications of the regulation of AI include the need to provide a horizontal, binding AI definition and a new approach to assigning a quasi-personhood to advanced AI systems, the choice

---

[77] See Genus, A., Stirling, A. (2018). Collingridge and the dilemma of control: Towards responsible and accountable innovation. *Research Policy*, Volume 47, Issue 1, doi: 10.1016/j.respol.2017.09.012, and Kudina, O., Verbeek, P.-P. (2019). Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy. *Science, Technology, & Human Values*, 44(2), 291–314. doi: 10.1177/0162243918793711.

[78] Collingridge, D. (1980). *The Social Control of Technology*, New York: St. Martin's Press; London: Pinter.

[79] Pampoukis, H., (2022) *Legal aspects of AI*, at "UNESCO: Open Science - Artificial Intelligence. Young New Horizons, Visions and Frontiers" Conference. Followers on Law as an autopoietic system assign on Law the ability of self-production and self-reproduction of its own elements and structures, i.e., its legal norms, after assessing the exterior environmental influences. On Law as an autopoietic system see Teubner, G., (1994) *Law as an Autopoietic System*, Oxford/Cambridge, Blackwell Publishers, 1993, and a review in Beck, A., (1994) Re*view:* Is Law an Autopoietic System? Reviewed Work: Law as an Autopoietic System by Gunther Teubner, *Oxford Journal of Legal Studies*, Vol. 14, No. 3 (Autumn, 1994), pp. 401-418 (18 pages), Oxford University Press, https://www.jstor.org/stable/764738.

[80] Surden, H. (2019), Artificial Intelligence and Law: An Overview, *Georgia State University Law Review*, Vol. 35, 2019, p. 1327.

[81] Wang, N., Tian, M.Y. (2022). 'Intelligent Justice': AI Implementations in China's Legal Systems. In: Hanemaayer, A. (eds) *Artificial Intelligence and Its Discontents*. Social and Cultural Studies of Robots and AI. Palgrave Macmillan, Cham. doi: /10.1007/978-3-030-88615-8_10.

between soft law, i.e., not strictly binding, or hard law, i.e., binding "AI Law", as well as specific concerns like AI liability in the case of autonomous systems, general AI bias, and the application of ethical guidelines for the fair, transparent and controlled production, distribution and utilisation of AI.

2.2. A legal definition for AI: the international examples and an attempt to define it

As previously analysed in Section 1, agreeing on a common technical definition of AI seems to be a difficult task; the same applies to Law. The complexity of the system and its innumerable -and perhaps yet known- potential, combined with the "pacing problem" make it exceptionally hard to provide a fitting and complete definition. However, everything starts at and falls upon that very definition; too broad and potentially technology is overregulated and progress is halted or discouraged - too narrow and AI is underregulated.

OECD defined an AI system as "*a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments*"[82]. Subsequently, in 2019, the Independent High-Level Expert Group (HLEG) on AI, established by the European Commission since 2018, proposed an updated definition, according to which: "*Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans [humans design AI systems directly, but they may also use AI techniques to optimise their design] that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions*"[83].

[82] The Recommendation on AI was adopted by the OECD Council on 22 May 2019 on the proposal of the Committee on Digital Economy Policy (CDEP).
[83] AI HLEG (2019) "A Definition of AI: Main Capabilities and Disciplines", p. 6.

The historic draft EU AI Act, the first attempt at EU AI regulation, as first published in April 2021, defined AI as a "*software that is developed with one or more of the techniques and approaches listed in Annex I* [ML and statistical approaches, logic- and knowledge-based approaches and more], *and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with*"[84]. As mentioned in the AI Act explanatory memorandum, "*the definition of AI system in the legal framework aims to be as technology neutral and future proof as possible, taking into account the fast technological and market developments related to AI*" – and that fluidity is exactly one of the difficulties that emerge when attempting to delineate a field as unique and complex as AI.

In the updated AI Act of 2023, the Internal Market Committee and the Civil Liberties Committee of the European Parliament adopted a draft negotiating mandate and according to EURACTIV's Luca Bertuzzi[85], the European Parliament reached a political agreement on 3 March to adopt an AI definition similar to the one used by the Organisation for Economic Cooperation and Development (OECD) for the EU AI Act. This definition is important as it determines an AI system as "*a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments*"[86], which ensures harmonisation with the OECD's definition.

In the USA, the definition of AI from the National Defense Authorization Act for Fiscal Year 2019, which is also referenced in the Executive Order on Maintaining American Leadership in Artificial Intelligence, provides five different versions of AI definitions: "*1. Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets. 2. An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning,*

---

[84] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, COM/2021/206 final.

[85] Bertucci, L. for EURACTIV (March 7, 2023). *EU lawmakers set to settle on OECD definition for Artificial Intelligence*, at: https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/?utm_source=substack&utm_medium=email

[86] See the compromise text as of 11/05/2023 at: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf, p. 137.

communication, or physical action. 3. An artificial system designed to think or act like a human, including cognitive architectures and neural networks. 4. A set of techniques, including machine learning, that is designed to approximate a cognitive task. 5. An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision making, and acting[87]." A very broad approach as this showcases the issue described above; that the legal answer to the "What is AI?" question is definitely not easy to provide. An interesting case of AI system definition by regulation is that of New York, where local Law 144, which took effect on the 1st of January 2023, defined automated [employment] decision tools as "*any computational process, derived from machine learning, statistical modeling, data analytics, or artificial intelligence, that issues simplified output, including a score, classification, or recommendation, that is used to substantially assist or replace discretionary decision making for making [employment] decisions that impact natural persons*". In this definition, the actual *replacement* of decision-making ability that impacts humans is attributed to AI systems.

In Canada, the proposed Artificial Intelligence and Data Act (AIDA), introduced as part of the Digital Charter Implementation Act (2022) recognised that the design, development and deployment of AI systems across provincial and international borders should be consistent with national and international standards to protect individuals from potential harm; it defines an AI system as "*a technological system that, autonomously or partly autonomously, processes data related to human activities through the use of a genetic algorithm, a neural network, machine learning or another technique in order to generate content or make decisions, recommendations or predictions*[88]".

In Japan, according to the Basic Act on the Advancement of Public and Private Sector Data Utilisation, the term "AI-related technology" means "*technology for the realisation of intelligent functions, such as learning, inference, and judgment, by artificial means, and utilisation of the relevant functions realized by artificial means*"[89].

---

[87] 10 U.S. Code §2358.
[88] Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts, 44th Parl., 1st Sess., 70-71 Elizabeth II, 2021-2022 (First Reading), at: https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading.
[89] Act No. 103 of December 14, 2016.

In Russia, the official definition of AI in the "Strategy for the development of AI in Russia until 2030" stands for "*a set of technological solutions that allow to simulate human cognitive functions (including self-learning and search for solutions without a pre-set algorithm) and get results that are comparable, at least, with the results of human intellectual activity when performing specific tasks*[90]".

Meanwhile, in the UK, the National Security and Investment Act 2021 (Notifiable Acquisition) (Specification of Qualifying Entities) defines AI as the "*technology enabling the programming or training of a device or software to (i)perceive environments through the use of data; (ii)interpret data using automated processing designed to approximate cognitive abilities; and (iii)make recommendations, predictions or decisions; with a view to achieving a specific objective*". However, the policy paper "Establishing a pro-innovation approach to regulating AI" published in July 2022 as part of the National AI Strategy introduced a different approach, claiming that "*a detailed universally applicable definition [of AI] is not needed*", with a view to regulate the *uses* of AI rather than AI itself. More specifically, current UK AI policy is concerned that lack of granularity (as in the case of EU Law, which attempts to provide a relatively fixed definition of new forms of technology) could hinder innovation; thus, it aspires to allow regulators to set out and evolve more detailed definitions of AI according to the *ad hoc* domains or sectors, while the core characteristics of AI are set on a nation-wide level in order to inform the scope of the framework.

Finally, and on the basis of the UK's National AI Strategy, **is an AI definition actually worth establishing**? Admittedly, the fact that an adequate definition is difficult to provide should not, just by itself, discourage any effort to try. However, as technology progresses, if language and Law don't respond to the changes at a rapid pace, they quickly become stale and outdated. One would then suggest that no definition is necessary, and the inevitable gaps be filled with pieces or sets of legislation examined on an individual basis and on the *ad hoc* needs of society. As in the case of the "computer programme" definition, the preamble of the draft EU Directive on the legal protection of computer programs indicates that it is not legally defined, "*as has been recommended by experts in the field that any definition of what constitutes a program would of necessity become obsolete as future technology changes the nature*

[90] Shchitova, A. (2020). Definition of Artificial Intelligence for Legal Regulation, 2nd International Scientific and Practical Conference on Digital Economy (ISCDE 2020), *Advances in Economics, Business and Management Research*, volume 156, p. 618.

of programs[91]". Progress quickly renders legal definitions for technology-based solutions obsolete[92]. One counterargument though is that too much laxity or over-the-top discretion at *what* is regulated and *when* that is regulated could potentially lead to underregulated AI or legal fragmentation.

Perhaps there is a middle ground: one could focus on the mutually agreed characteristics of AI and provide a sort of "umbrella" definition that would be further specified in the following articles of legislative piece. These include, *inter alia*, that:

(1) AI is an artificial, or man-made system,

(2) that is programmed to operate with varying degrees of autonomy,

(3) to learn how to adapt to new inputs after receiving appropriate training,

(4) and to produce outputs such as recommendations, predictions, classifications, associations, prioritisations or/and take action,

(5) in order to interact with the virtual and/or physical environment.

[91] PROPOSAL FOR A COUNCIL DIRECTIVE ON THE LEGAL PROTECTION OF COMPUTER PROGRAMS, COM/88/816 FINAL, OJ C 91, 12.4.1989, p. 9.

[92] Zalewski, T. (2021) Definition of artificial intelligence, in: Lai, L., Świerczyński, M. (eds.), *Legal and Technical aspects of Artificial Intelligence*, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw, p. 16.

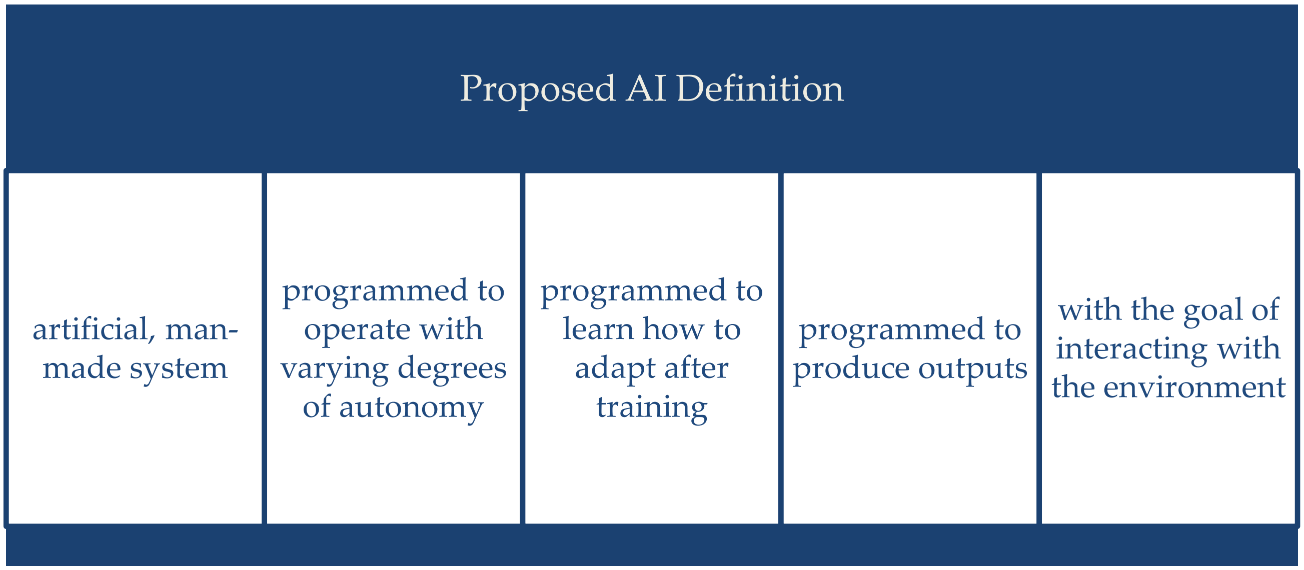


Figure 8

Created by the author

By using a generalised but targeted approach in order to define AI, the legislator neither confines himself within the strict boundaries of the Law, nor allows extreme flexibility to the technology industry; that is because any new technological developments will probably use the existing AI groundwork as their foundation and they will, therefore, in one way or another, share the aforementioned characteristics - thus, they will find themselves within the regulatory scope of the AI definition.

2.3. The legal status of AI: is personhood attributable?

One of the first questions that a legal scholar finds themselves asking is *what* is AI, in the legal sense; is it an object (*res* in Latin)? Is it a separate legal entity like a human? Or is it a quasi-being capable of consciousness or sentience "*in the image and likeness of man*"[93]? Or should a new legal status be formulated to fit this new technological marvel as it progresses rapidly through modern history?

The legal person is a being that has the ability to hold rights and bear duties, which is described as "legal capacity". Legal persons are, apart from natural persons (humans), juridical persons like corporations, religious and governmental entities. It has been argued that the most advanced autonomous systems have the ability to "communicate", to perceive certain external events and to even showcase some degree of "creativity". The quantitative question, i.e., how much autonomy or "creativity" they display, is the key to deciding on whether a distinctive "quasi-legal personality" could be attributed to advanced AI systems[94]. In other words, "self-awareness" as an attribute of AI that would predispose it to have rights (?) and duties.

As S. Chesterman so eloquently put it, "[juridical persons] *have personality because a legal system chooses to give it to them*" and "*personality is granted to achieve policy ends, such as encouraging entrepreneurship, or to contribute to the coherence and stability of the legal system*[95]". In short, that means that the legislator chose, after assessing the needs and goals of financial/social policy, to introduce the concept of juridical personhood. Similarly, in the case of AI, the Law could identify the factual triggers that would justify the need to devise a new form of legal status (like a "robot personhood"), or to include AI to an existing form of legal personality, as in the case of animals. Animals are generally recognised as objects in the legal sense, but with a limited moral status that allows them to hold certain rights ("animal rights", for example no hunting, experimenting or selective animal breeding). By reverse analogy, certain obligations could be attributed to AI.

Describing advanced AI as a "quasi-legal person" could be preferred, as the most appropriate term to convey a different understanding of the terms "person" and "subject of the Law", in such a way that these concepts are no longer considered identical, but that the broader conceptual scope of the latter should also include certain entities, other than persons, in those exceptional cases such as those referred

---

[93] That question would certainly become timely in the case of conscious AGI.
[94] *Supra* 79.
[95] Chesterman, S. (2020). Artificial Intelligence and the Limits of Legal Personality. *International & Comparative Law Quarterly*, 69(4), 819-844. doi: 10.1017/S0020589320000366, p. 823.

to (and to the extent that) the legislature makes them the bearers of rights and obligations or (and) recognises their capacity for legal action. Therefore, the legislative attribution of quasi-legal personality on an AI entity would not necessarily imply the acquisition of an unlimited legal capacity, but that capacity may be defined (that is to say, limited) either expressly by the legislature itself or interpretatively, by reference to the legislative purpose pursued with said legislative decision; for example, in the context of liability laws.

2.4. The implications of AI accountability and liability

The determination of the legal status of AI raises a wider range of questions on how to ensure accountability and the attribution and allocation of liability when a robot runs and "decides" autonomously. In 2017, the European Parliament passed a resolution with recommendations to the European Commission on civil law rules on robotics[96] which defined a robot's autonomy as "*the ability of purely technological nature to take decisions and implement them in the outside world, independently of external control or influence*".

It has been suggested attributing AI systems with some form of legal personality to help address liability questions, such in the cased of Autonomous Vehicles in the case of driverless cars that cause road accidents. The 2017 European Parliament's Resolution with recommendations to the Commission on Civil Law Rules on Robotics called for a "*creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently*". The reference was made in a rather cautious manner, as the Parliament called on the Commission to "*explore, analyse and consider the implications of all possible legal solutions*" when "*carrying out an impact assessment of its future legislative instrument*" in the context of civil liability. The proposal was harshly criticised in an open letter

[96] European Parliament Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) (2018/C 252/25), para 59(f).

issued by "Artificial Intelligence and Robotics Experts, industry leaders, law, medical and ethics experts"[97]. The concept of an "electronic legal person" seems quite interesting in the context of liability, as the more autonomous robots are, the less they can be considered to be simple tools in the hands of other actors such as the manufacturer, the operator, the owner, the user, etc., and could, potentially, be deemed as *solely responsible* for any wrongdoings. Also, in the cases of applying "personhood" to AI systems, the establishment of a supervisory authority, or at least a system of compulsory registration of electronic legal persons has been mentioned as necessary[98]. However, as it is often found in the field of Law, things are not that simple.

Civil liability refers to the legal responsibility that an *individual* or *entity* has for any harm or injury caused to another person or their property as a result of action or negligence. Civil liability typically leads to an award of damages which can only be paid if the wrongdoer is capable of owning property; that hardly seems realistic in the case of AI. In 2018, an accident involving an automated vehicle that caused fatal injuries to a pedestrian highlighted some of the most important risks associated with AI systems, like accountability[99]. In the context of robotics, the conventional legal theory supports that liability is shared between the robot manufacturer, owner, operator, and any third party who might be involved in the operation or maintenance of the robot. One of the key issues in determining civil liability for robots is the question of whether robots can (and should) be exclusively held accountable for their actions.

Manufacturers are now designing AI that can learn from experience, adapt to changes and even anticipate and respond to human needs. The issue of liability becomes more complex when considering the potential actions of autonomous robots which are designed to operate without human intervention

---

97 "Open Letter to the European Commission Artificial Intelligence and Robotics", http://www.robotics-openletter.eu/

98 Bar, G. (2021) Robot personhood – what is anthropocentric artificial intelligence good for? in: Lai, L., Świerczyński, M. (eds.), Legal and Technical aspects of Artificial Intelligence, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw, p. 46.

99 National Transportation Safety Board, Highway Accident Report–Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian, Tempe, Arizona, March 18, 2018, NTSB/HAR-19/03, 2019, p. 1., https://www.ntsb.gov/investigations/accidentreports/reports/har1903.pdf. As mentioned in the report, "*The vehicle operator's prolonged visual distraction, a typical effect of automation complacency, led to her failure to detect the pedestrian in time to avoid the collision*" (p. 57).

and may make decisions that result in harm or damage to individuals or entities. AI autonomy is critical to determining liability. In the case of assistive algorithms in the medical field, the clinician ultimately makes the clinical judgement or decision, and so should be held liable. However, the American Medical Association recommends that AI developers should be held liable for medical errors made by autonomous AI algorithms[100].

Some argue that AI should be treated like any other man-made tool and that manufacturers and operators should bear the full responsibility for any harm caused by it. However, others argue that AI that can learn and make decisions should be treated as an autonomous being with its own "moral agency". In these cases of gradually increasing AI autonomy, it may be difficult to assign liability and determine whether a human is responsible for the AI's actions. Hence, the manufacturer, the operator of the AI system, or even the software programmer should be held liable for any injuries or damages caused by its failure. Ensuring that software engineers design and the manufacturers produce reliable and safe AI, as well as that operators understand how to properly use and maintain it, is crucial.

Under the current legal framework, AI cannot be held liable *per se* for acts or omissions that cause damages to third parties; similarly, the existing rules on liability cover cases where the cause of the AI's act or omission can be traced back to a specific human agent such as the manufacturer, the operator, the owner or the user and where that agent could have foreseen and avoided the robot's harmful behaviour.

Current EU Law considers the producer of a product and respectively the manufacturer of the defect component liable for damage[101]. Further, the damage must be caused by the product, in particular, by the defect of the product The Product Safety Directive requires that all consumer products to be safe[102], while said safety rules are complemented with market surveillance and the powers conferred to national

[100] However, liability for level 3 models should depend on whether an error was caused directly by the AI, or by the clinician after they were involved as the backup operator. See American Medical Association. (2019). *Augmented intelligence in health-care: payment and regulation*, https://www.ama-assn.org/system/ files/2019-08/ai-2019-board-report.pdf.
[101] "Damage" is defined as: (a) damage caused by death or by personal injuries, and (b) damage to, or destruction of, any item of property other than the defective product itself.
[102] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products

authorities under the Market Surveillance Regulation[103] and the General Product Safety Directive. Still, the existing legal framework seems unfitting to successfully regulate AI liability; for example, personality rights are currently not covered by product liability rights, as in the case of copyright cases due to AI-generated content[104] or bias leading to incorrect assessments by the AI, as in the case of the use of AI tools during CVs' screening before job interviews. At the same time, some critics of the regulation of AI liability approach argue that it could stifle innovation and technological development.

To address these challenges, legal experts have suggested implementing a comprehensive regulatory framework for AI liability, which would require manufacturers to meet certain safety standards, to ensure their products are secure and to provide clear warnings and instructions for safe use, as well as guidelines for maintenance and repairs. In the EU, in order to adopt a more flexible approach that would allow for technical innovative experimentation while still ensuring public safety, a Proposal for a Directive on adapting non contractual civil liability rules to artificial intelligence (**AI Liability Directive**)[105] and a revision of the Product Liability Directive[106] were published in late 2022.

The new rules cover liability claims based on the fault or omission of any person (providers, developers, users), for the compensation of any type of damage covered by national law (life, health, property, privacy, etc.) and for any type of victim (individuals, companies, organisations, etc.). Furthermore, two main safeguards are provided for: a) the rebuttable "presumption of causality" in cases of probable non-compliance (meaning that the burden of proof is eased), and b) the option of accessing evidence of AI-related incidents and information about high-risk AI systems (Art. 3). Without delving into the specifics of the new proposed framework, it is evident that it addresses legal uncertainties surrounding AI systems and provides a good first step towards a more comprehensive and updated framework;

[103] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011.
[104] See Brittain, B. for Reuters (2023) *AI-created images lose U.S. copyrights in test for new technology*, at: https://www.reuters.com/legal/ai-created-images-lose-us-copyrights-test-new-technology-2023-02-22/.
[105] COM(2022) 496 final, 2022/0303 (COD) Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)
[106] COM(2022) 495 final 2022/0302 (COD) Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on liability for defective products.

however, it is crucial to address any inevitable shortcomings and create a working relationship between the new proposed framework and the AI Act.

To address liability concerns related to AI, it is important to establish clear guidelines and regulations for proper AI use, like the EU AI Act. This could involve creating standards for design and programming, as well as establishing liability frameworks for different types of AI and their potential actions. It is also important to ensure that individuals and entities who use AI receive proper training and education on how to use and maintain them, for example through establishing training programs for operators, as well as incorporating training on robotics and liability into AI educational programmes.

Furthermore, it is crucial to consider the potential ethical concerns surrounding the use of robots, particularly in industries such as healthcare. For instance, if a robot is responsible for administering medication to a patient, who is responsible if the machine malfunctions or gives the wrong dosage? In such cases, legal and ethical standards need to be established that address these concerns and hold parties accountable for their actions. The ethical ramifications of AI use will be analysed further below.

2.5. AI, data privacy and cybersecurity concerns

One of the practical challenges of AI stem from the use and protection of data. As AI relies on the datafication process to progressively support humans in decision making processes in all spheres of life, personal data is used to generate intel[107]. The concept of data privacy has become increasingly relevant in the digital age, as individuals and organizations alike are constantly creating and consuming digital information. Companies are using consumer data to shape their marketing strategies, while governments are collecting personal information for national security. AI is making it easier for these entities to analyze this data and obtain insights that may not have been possible before.

---

[107] Fontes, C., Hohma, E., Corrigan, C., Lütge, C. (2022), AI-powered public surveillance systems: why we (might) need them and how we want them, Technology in Society, Volume 71, 102137, https://doi.org/10.1016/j.techsoc.2022.102137, p. 1.

From the "*continuous, systematic collection, analysis and interpretation of health-related data*[108]" to the deployment of CCTV systems for public surveillance and the increasing use of facial recognition technology, potential data theft, misuse of collected data[109], data privacy infringements and cybercrime are some of the legal and ethical issues that could potentially arise. For example, Chinese researchers took remote control of a Tesla model from 12 miles away in 2016 and they were able to interfere with the car's brakes, door locks and dashboard computer screen, among other electronic features[110]; as technology for improved user experience like efficient routing and mapping often requires personal/sensitive data from the user (e.g., location, navigation patterns), a significant portion of people are concerned on possible cybersecurity issues[111].

AI systems can quickly analyse vast amounts of data in order to analyse social media updates, purchase history, and digital footprints to offer users tailored advertisements or even prevent the spread of misinformation on social media platforms[112]. The ethical challenges of AI systems affect data privacy as they raise questions around the use of data, particularly in situations where the data owner is unaware that their information is being used. As autonomous systems become more interconnected and communicate through the internet of things, they also become more vulnerable to breaches of privacy; and as AI is increasingly being used in homes and private spaces, privacy concerns are raised and the ability for AI to collect sensitive information is questioned. In 2018, the Amazon voice assistant "Alexa" recorded a private conversation and sent it to a random contact of the user[113]; recently, Snapchat's

---

[108] World Health Organization, Surveillance in Emergencies, World Health Organization, 2021. https://www.who.int/emergencies/surveillance.

[109] For example, a man was wrongfully arrested based on face recognition, see Hill, K. for the New York Times (2020), *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*, at: https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html.

[110] Solon, O. for the Guardian (2016) *Team of hackers take remote control of Tesla Model S from 12 miles away*, at: https://www.theguardian.com/technology/2016/sep/20/tesla-model-s-chinese-hack-remote-control-brakes.

[111] Othman, K. (2021). Public acceptance and perception of autonomous vehicles: a comprehensive review. *AI Ethics 1*, 355–387, https://doi.org/10.1007/s43681-021-00041-8.

[112] Bergengruen V., Perrigo, B. for The Times (2021) *Facebook Acted Too Late to Tackle Misinformation on 2020 Election, Report Finds*, at: https://time.com/5949210/facebook-misinformation-2020-election-report/.

[113] Wolfson, A. for the Guardian (2018) Amazon's Alexa recorded private conversation and sent it to random contact, at: https://www.theguardian.com/technology/2018/may/24/amazon-alexa-recorded-conversation

chatbot "My AI" revealed that it was aware of the user's location, even though it had previously reassured them that it didn't have access to it[114].

The Law has attempted to regulate these issues in a progressively more determined fashion: GDPR[115] has been the most pioneering piece of data protection legislation worldwide calling for lawfulness, fairness and transparency (Art. 5) and the EU AI Act provides for certain safety nets such as the conditional use of real-time remote biometric identification systems[116] (Art. 5 of Title II), as in the case of public surveillance.

The benefits of AI are clear, but there is still a long way to go to ensure that the technology is used ethically and safely. The relationship between data privacy and AI could be analysed in much greater detail[117]; however, for the purposes of the present thesis, the integration of the most fundamental legal and ethical principles into AI development should be put at the centre of the analysis, as it follows below.

2.6. Infusing legal principles and ethics into AI regulation: The Key Pillars

2.6.1. Transparency

In Houston TX, an AI system was used in 2012 to evaluate the school district's' teachers; however, after some bad ratings and layoffs, a few of them took their case to court and a federal judge ruled that use of an algorithmic evaluation programme violated their civil rights, as the teachers had no way of knowing

---

114 See the viral exchange at user's @rewolfe27 profile on Twitter, at: https://twitter.com/rewolfe27/status/1648785906361417728?s=20 (19/04/2023).

115 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

116 "Biometric data" means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data.

117 See more at: Datatilsynet – The Norwegian Data Protection Authority (2018) *Artificial intelligence and privacy Report*, at: https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

if the programme operated fairly or faulty and the software company refused to provide details on the used algorithms, deeming them as "*trade secrets*"[118].

Hence, the need for disclosure of all relevant information of an AI system so that others (e.g., users, third parties, regulators, etc.) can make informed decisions is the one of the most fundamental requests for AI use. In short, **AI transparency** refers to the need to describe, inspect and reproduce the mechanisms through which AI systems make decisions and learn to adapt to the environment and the governance of the data used.

Ensuring transparency, in the sense of access to the way the AI operates and its structure so that it can be examined and understood, is essential for the integration of ethics into AI. Transparency must be ensured from the beginning of the development of the AI system and information must be provided about the data used in its training, its structure and mode of operation, its purpose of operation and its use, so that any problematic code present can be accurately detected and consequently error correction and the incorporation of ethical and legal rules can be achieved.

The parameters to be considered to ensure AI transparency are depicted as follows:

AI Transparency

1.	**Identifiable AI:** AI systems should present themselves as AI (for example, as virtual assistants), so that the human user can make an informed decision on whether to interact with the AI or not.
2.	**Options alternative to AI:** the user should retain the option to interact with a human instead of an AI system, if they so decide.
3.	**Transparent AI System Processes**: the AI system's data sets, capabilities, processes and limitations should be communicated to end-users in a manner appropriate to the *ad hoc* use.

[118] Langford, C. for Courthouse News Service. (2017). *Houston Schools Must Face Teacher Evaluation Lawsuit*, at: https://www.courthousenews.com/houston-schools-must-face-teacher-evaluation-lawsuit/. The federal judge decided in favour of the teachers, see Document 91 on Case 4:14-cv-01189, as filed in Texas Southern District on 05/04/17 at: https://www.courthousenews.com/wp-content/uploads/2017/05/HoustonTeachers.pdf.

More specifically, humans need to be aware that they are interacting with an AI system and provided with all necessary information to access that fact.

	Topics	Be transparent by disclosing..
1.	Environmental costs	- how much energy was needed to train/test the system - what kind of energy the servers use (renewable or fossil) - what is done to compensate for the CO2 emissions or other environmental harm
2.	Employment	- the need to use people in low-cost countries for content moderation and training data creation, or clickworkers to develop and maintain AI systems (if this was the case) - what is done to guarantee fair wages and healthy employment conditions for all contributors to the system
3.	Business model	- the purpose of the system, what it can and can't be used for - the reasons for deploying the system - the design choices of the system - the degree to which the system influences and shapes the organisational decision-making process - who is responsible for the development and control of the system
4.	User's rights	- all information in a clear user interface and understandable, non-technical language - that people will be informed when they interact with an AI system - that people will be informed when decisions about them are being made by an AI system - if/that/how people can opt out, reject the application of certain technologies, or ask for human interaction instead - if/that/how people can check and/or change their personal data that is being used in the system - if/that/how people can request an explanation for the given outcome or decision. An explanation should be free-of-charge. - if/that/how people can challenge a given outcome or decision - if/that/how people can give feedback on the system
5.	Data	- what kind of data is being used in the system, which variables - why these variables are needed - how/when/where the training and test data was collected - what was changed/corrected/manipulated in the data, and why this had to be done - how the required data is gathered from the user - if/that people explicitly need to agree to use their data - in which country the data is stored and where the provider is headquartered

Picture 1 *See footnote 119 for source.*

On the left are some examples of information that should be provided by different topic, if the need for transparency were to be truly honed[119].

As regards the first topic, if we attempted to make a correlation between the need for transparency and the environmental aspect of AI, as it will be analysed further in the following section of the present thesis, we could mention the Aarhus Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters[120]. The Convention not only enshrines the dissemination of information, i.e., the obligation to openly

[119] Picture source: Woudstra, F. (2020). *Ethical Guidelines for Transparent Development and Implementation of AI - an Overview*, Filosofie in actie Blog, at: https://www.filosofieinactie.nl/blog/2020/4/9/ethical-guidelines-for-transparent-development-and-implementation-of-ai.

[120] See the 2005/370/EC Decision on conclusion of the Aarhus Convention by the Council of the EU. Two relevant Directives were adopted: a) Directive 2003/4/EC on public access to environmental information and b) Directive 2003/35/EC providing for public participation in respect of the drawing up of certain plans and programmes relating to the environment and amending with regard to public participation and access to justice Council Directives 85/337/EEC and 96/61/EC. Afterwards, Regulation (EC) 1367/2006 of the European Parliament and of the Council of 6 September 2006 on the application of the provisions of the Aarhus Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters to Community institutions and bodies entered into force on 28 September 2006.

provide environmental information to the public, but also imposes a concrete obligation for the competent authorities to share on their own initiative all environmental information that is relevant to each case "*when all options are open and effective public participation can take place*" [Art. 6(4)]. Similarly, the environmental impact of AI or the positive environmental implications of AI could be interpreted as "environmental information" and thus be subject to transparency obligations pursuant to the Aarhus Convention and all relevant European and national legislation.

In general, to achieve transparency in AI, the quality and quantity of the data used in models must be accurate and sufficient. AI systems and their decisions should be explained in a manner adapted to the stakeholder concerned, who must be informed of the system's capabilities and limitations. The data sets and the processes that yield the AI system's decision, including those of data gathering and data labelling as well as the algorithms used, should be documented to the best possible standard to allow for traceability and an increase in transparency. This also applies to the decisions made by the AI system, which enables identification of the reasons why an AI-decision was erroneous which, in turn, could help prevent future mistakes. Transparency in AI is of critical value; it allows the user and third parties to identify and deal with technical errors and to prevent human rights violations at their "formation" phase.

The **GDPR** contains some provisions that could be used in AI transparency frameworks, as is the possibility of processing personal data with respect to the rights of the subjects of processing (natural persons) and to the principles of the Regulation, already "by design" of the processing system. In accordance with the principle of data protection by design (para 1 of Art. 25 GDPR), at the time of designing the processing systems and determining the means of processing, the controller must incorporate and implement appropriate measures and use privacy-enhancing technologies, such as pseudonymization of personal data, as soon as possible (i.e. replacing personally identifiable information with artificial identifiers), encryption (encoding personal data so that only those who are authorised to read them), minimise data processing and incorporate the necessary safeguards into the processing, in such a way as to meet the requirements of the GDPR and protect the rights of data subjects ("Data Protection Principle"). In other words, regulation by design means that the processing system is designed from the outset in such a way as to respect the rights of individuals and legislative authorities, without any further action. Similarly, relevant to AI applications and large-scale data analysis is Article 25 para. 2 of the GDPR, which provides that, by default, only personal data which are necessary for each specific purpose of the processing should be processed.

Furthermore, the draft **EU AI Act** provides for strict transparency obligations as regards high-risk AI systems, the requirements of high-quality data, documentation and traceability, transparency, human oversight, accuracy and robustness; for non-high-risk AI systems, limited transparency obligations are imposed, for example in terms of the provision of information to flag the use of an AI system when interacting with humans: transparency obligations are applied for systems that

- **(i)** interact with humans,
- **(ii)** are used to detect emotions (emotion recognition systems) or determine association with (social) categories based on biometric data, or
- **(iii)** generate or manipulate content ('deep fakes[121]').

Pursuant to Art. 13, high-risk AI systems[122] shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users. The information should specify, *inter alia*, the characteristics, capabilities and limitations of performance of the high-risk AI system and the intended purpose and the level of the accuracy, robustness and cybersecurity of the system.

---

[121] Pursuant to para. 3 of Art. 52, an AI system that generates or manipulates image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful is defined as a "deep fake". Its user shall disclose that the content has been artificially generated or manipulated, unless the use is authorised by law to detect, prevent, investigate and prosecute criminal offences or it is necessary for the exercise of the right to freedom of expression and the right to freedom of the arts and sciences guaranteed in the Charter of Fundamental Rights of the EU, and subject to appropriate safeguards for the rights and freedoms of third parties.

[122] Pursuant to Art. 6: "1. *AI system shall be considered high-risk where both of the following conditions are fulfilled: (a) the AI system is intended to be used as a safety component of a product, or is itself a product, covered by Union harmonisation legislation, (b) the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation. 2. In addition to those high-risk AI systems, AI systems referred to in Annex III shall also be considered high-risk*" (systems used in the areas of biometric identification and categorisation of natural persons, management and operation of critical infrastructure, Education and vocational training, employment, workers management and access to self-employment, Law Enforcement etc.

However, in general, access to the source code of the algorithm is most of the time impossible because it is considered as a valuable asset of the companies and stakeholders; therefore, regulatory requirements for AI "openness" of varying rigour and consequences in case of non-compliance should be deployed.

2.6.2. *Explainability*

AI methods, especially ML methods, are opaque. In the typology provided by professor Burrell of Berkeley University, opacity in algorithms is seen in a three-dimensional form of:

a) intentional corporate or state secrecy[123], while lax regulation and corporate competition have been cited as the underlying causes for the phenomenon[124],

b) technical illiteracy, that is, the inability for the general public to "read code" in order to understand how the AI system works, and

c) complexity due to the scale of data systems: as in the case of ML, DP and LLMs, the mere size of computational resource to support them combined with the complicated code that is put through the lens of the "mechanism of the algorithm" is what yields complexity (and thus opacity[125]).

Said opacity renders the full explainability of AI systems a difficult, but maybe not unattainable, task. To address this issue, **Explainable AI (XAI)** has been proposed as the tool to "decrypt" or "decode" the way a complex AI system works. It aims to create a suite of techniques that produce more explainable models whilst maintaining high performance levels[126], in order to ensure that algorithmic decisions can

---

[123] Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Societ*y, 3(1). https://doi.org/10.1177/2053951715622512.

[124] Pasquale, F. A. (2015). *The Black Box Society: The Secret Algorithms that Control Money and Information*, Harvard University Press.

[125] *Supra* 123.

[126] A. Adadi and M. Berrada. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), *IEEE Access*, 2018, Vol. 6, p. 52138.

be explained to third parties in general in non-technical terms, so as to ensure end-user comprehensibility.

As it is known, *the reason why* a decision was made is a core element of the Law on an international scale. To justify is to prove to be just, right, or reasonable[127]. The "*duty to justify*" has been crystallised in the fundamental right to a fair trial (i.e., the right to a reasoned decision), as the latter is provided for in Article 6 of the European Convention of Human Rights[128]. In legal theory, the notion of "discovery", i.e., how a decision is reached, and "justification", i.e., how a decision is justified (especially in judicial proceedings[129]) are two separate sides of the explanation "coin". The reasons behind a verdict, the legal grounds used and the logical sequence between events and evidence gathered and accessed have to fulfil certain requirements in order to qualify as permissible justification.

Such requirements include the need for the reasoning process which crystallises in the provided justification to be:

- **a)** *clear*, i.e., when the thoughts of the decisive body are recorded in a clear manner, by using the appropriate wording, grammar and syntax. The decision must not be limited to the repetition of stereotyped expressions or be based on contradictory elements.
- **b)** *specific*: it has to be specific to the case, that is, when all of its elements refer to the *ad hoc* situation. If the necessary link to the specific case is not demonstrated, this condition is not met.
- **c)** And, perhaps the most difficult to "cross off the list", a proper justification is *sufficient*. It must be supported by elements, facts and reasoning which are accessed as of adequate quality and quantity. It is regarded as sufficient when it contains all its elements in such a way that no gaps or doubts remain as to the correctness of the decisive body's judgement. It is when the decision-maker properly assesses the facts, however material or immaterial they may seem. Of course,

[127] Luiz Silveira, L. (2014). Discovery and Justification of Judicial Decisions: Towards More Precise Distinctions in Legal Decision-Making, *Law and Method*, ReM 2014-09, doi: 10.5553/REM/.000007.
[128] Council of Europe. (1950). Convention for the Protection of Human Rights and Fundamental Freedoms. In Council of Europe Treaty Series 005. Council of Europe.
[129] See Dymitruk, M. (2021) *The role of XAI in automated decision-making in the judiciary,* in: Lai, L., Świerczyński, M. (eds.), Legal and Technical aspects of Artificial Intelligence, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw, p. 290.

> adequacy presupposes points a and b but also requires a deeper analysis of both the factual circumstances and (in case of a judicial decision) the legal background of each case. In AI, the "Law" behind a system's decision, or, more correctly, output, is its code.

The reasons for an AI act or outcome are sometimes unknown and incomprehensible. Hence, a cause-and-effect relationship between the AI's "decision" and *what specifically led to it* should be established. As evidenced by the analysis of Adadi and Berrada, the importance of ensuring high levels of the explainability are important in order "to discover, to justify, to improve and, perhaps most important, to control [AI]"; and maybe, to even trust it.

Also, indicators for a right to explanation can be found in GDPR: first, Art. 15 regulates the right of access by data subject, stipulating that the subject shall have the right to obtain from the controller confirmation as to whether or not their personal data are being processed, and, where that is the case, access to the personal data and, *inter alia*, to the existence of automated decision-making, including profiling, and to *meaningful information about the logic involved* – although without delving into details on what signifies as "meaningful" - as well as the significance and the envisaged consequences of such processing for the data subject. Second, Art. 22 regulates automated individual decision-making including profiling and obliges the data controller to "*implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision*" in certain cases (para. 3). Therefore, keeping in mind the AI's staggering increase in use, henceforth AI systems should be able to provide justifications in order to be in compliance with the "right to explanation" in the cases covered by GDPR.

Phillips et al. propose some integral principles of XAI[130]:

---

[130] Philips, J., Hahn, C., Fontana, P., Yates, A., Greene, K., Broniatowski, D. and Przybocki, M. (2021), *Four Principles of Explainable Artificial Intelligence*, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, doi: 10.6028/NIST.IR.8312.

AI Explainability Principles

1.	**Meaningful Explanation:** a system fulfills the meaningful principle if the intended recipient understands the system's explanation(s). Also, each audience's needs, level of expertise and ability to understand said explanation will vary on an *ad hoc* basis.
2.	**Explanation Accuracy:** the explanation provided should be targeted to the specific audience and be accordingly intelligible and detailed.
3.	**Knowledge limits**: the system should be able to identify its capabilities and limitations whenever a situation arises when the systems identifies a case that it was not designed or approved to [reliably] operate, in order to increase trust in the system and to prevent misleading and dangerous outputs.

An array of explainable methods that fall under the term of XAI have been proposed in literature, which include true transparency (interpretable models such as decision tree, rules and linear models) and post-hoc interpretations, additional techniques used to lighten up the darkness of complex black-box models such as Deep NNs (and that either by generating local explanations for particular inputs or by globally explaining the entire model[131]). For example, the use of LIME (Local Interpretable Model-Agnostic Explanations), the technique that approximates any black box ML model with a local, interpretable model to explain each individual prediction, has been described as a tool that visualizes an AI model's decision-making algorithms and therefore "explains" the AI system.

In the absence of access to the code, Sandvig et al. detail and compare several forms of algorithmic audit as a possible response towards code explainability[132]. However, alleviating AI explainability problems seems will not be accomplished by a single tool, a programme or piece of legislation, but some combination of regulations or audits of the code itself and of the algorithms comprising it could be deemed incredibly useful. In the end, transparent and Explainable AI technology allows for better decision-making, trust, and, therefore, **accountability**.

---

[131] *Supra* 126, at p. 52151. The authors provide a thorough review of the existing portfolio of XAI approaches on pp. 52147-52151.
[132] See Sandvig C., Hamilton K., Karahalios K, et al. (2014) *Auditing algorithms: Research methods for detecting discrimination on internet platforms*. In: Annual Meeting of the International Communication Association, Seattle, WA, pp. 1–23.

2.6.3. *Accountability*

According to a 2018 survey[133], respondents were asked if they felt comfortable sending their Autonomous Vehicles (AVs) to work, knowing they would be responsible themselves in the event of an accident – 72% of respondents answered negatively, 10% positively, while 18% were in neutral status. Then, respondents were asked on how do laws regarding AVs affect their desire to buy one; 10% of respondents believe that laws will not affect their desire, 50% were neutral, and 40% believe that laws would affect their desire AVs. These results show how liability can greatly influence public attitudes towards AI.

The catalyst for trust towards AI systems is accountability, in the sense of who will bear responsibility for the harmful effects that an AI system may cause. In other words, accountability refers to the need to explain and justify one's decisions and actions to its partners, users and others with whom the system interacts. Responsibility refers to the role of people themselves and to the capability of AI systems to answer for one's decision and identify errors or unexpected results.

The OECD AI Principles state that "*AI actors should be accountable for the proper functioning of AI systems (…) based on their roles, the context, and consistent with the state-of-the-art*". This means AI actors should design, install, and monitor processes that include documenting AI system decisions, conducting or allowing auditing, and providing adequate response to risks and redress mechanisms where justified. It is no coincidence that accountability, responsibility and regulation are closely interlinked; using once again the AVs example: who will be responsible in the case of physical damage to persons, for example a pedestrian, by a miscalculation of the vehicle's system? The driver? The pedestrian who should have paid more attention? Or the AI system's developer?

---

[133] Martínez-Díaz, M., Soriguera, F.: Autonomous vehicles: theoretical and practical challenges. *Transp. Res.* Procedia 33, 275–282 (2018). https://doi.org/10.1016/j.trpro.2018.10.103

According to OECD, accountability in the AI ecosystem should be shared by everyone directly and indirectly involved in or affected by the development or use of an AI system. All actors, as depicted below by the OECD, should manage risks based on their roles, the context, and following the state-of-the-art:

Suppliers of AI knowledge and resources	Actors in the AI Lifecycle	Users of the AI system	Stakeholders
• "From whom" • e.g. dataset creators, curators, open-source community incl. big tech, importers, distributors, etc.	• "By whom" • e.g. data scientists, developers, domain experts, modellers, deployers, operators, system integrators, auditors, certifiers, management, etc.	• "For whom" • e.g. procurers, doctors using AI, AV drivers, public sector using AI, etc.	• "Unto whom" • e.g. affected individuals of communities, patients, children, workers, civil society, data subjects, vulnerable groups, bystanders, etc.

Figure 9

OECD DIGITAL ECONOMY PAPERS No. 349 (2023) *ADVANCING ACCOUNTABILITY IN AI GOVERNING AND MANAGING RISKS THROUGHOUT THE LIFECYCLE FOR TRUSTWORTHY AI*, p. 20.

In the field of personal data protection, a privacy tool that is complementary to protection by design and derived from the principle of accountability is the Data Protection Impact Assessment study to be carried out before developing high-risk technologies. Article 35 of the GDPR provides that if processing, in particular when carried out using new technologies, taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk for data subjects, the controller must carry out an "*impact assessment*" before processing, i.e., to systematically describe the processing, the purposes and legitimate interests pursued, to assess the necessity and proportionality of the processing operations, the risks they pose to the rights of the subjects and the measures envisaged to minimize the risks. That could be used in the field of AI as well, in order to identify the potential responsibilities that would arise in case of miscalculations or accidents.

The fact that AI systems are often used in critical applications, such as healthcare and aviation, further complicates things. In these contexts, an AI system making a mistake could have serious or even deadly consequences. As a result, there is a growing demand for accountability measures that ensure AI systems are safe and reliable; this includes developing standards for testing and certifying AI systems, as well as establishing new and specialised legal frameworks for liability in the event of an AI-related accident.

For example, in the field of medical AI, the responsible party varies from the clinician in the case of assistive algorithms to the AI developer in the case of high or fully automated systems. Therefore, a risk-based approach to responsibility and autonomy in medical AI systems should inform regulation, clinical validation, and clinician engagement[134]; similarly, a multidisciplinary approach as regards AI accountability is needed. AI decision-making raises the question of moral accountability; there is a critical question of who should be held morally responsible for the actions of machines. While some people argue that machines are incapable of free will, the possibility of creating autonomous AI agents with self-learning capabilities blurs the divide between human and machine action.

According to the 2020 UNESCO Recommendation on the Ethics of Artificial Intelligence "*The ethical responsibility and liability for the decisions and actions based in any way on an AI system should always* ***ultimately be attributable to AI actors*** *corresponding to their role in the life cycle of the AI system. Appropriate oversight, impact assessment, audit and due diligence mechanisms, including whistle-blowers' protection, should be developed to ensure accountability for AI systems and their impact throughout their life cycle*" (paras. 42-43).

It is important to recognize that AI accountability is an ongoing process; as AI technology continues to evolve and integrate into new areas of our lives, we will need to continually reassess our (legal) approach to it.

[134] Bitterman, Danielle & Aerts, Hugo & Mak, Raymond. (2020). Approaching autonomy in medical artificial intelligence. *The Lancet Digital Health*. 2. e447-e449. 10.1016/S2589-7500(20)30187-4.

2.6.4. *Ethical, Social and Human Rights considerations*

As a human creation, AI also carries the problematic elements of human societies; human errors that are brought to the fore and further intensified when incorporated in an artificial agent like AI. The past years there have been multiple controversies over AI systems yielding biased and discriminatory results. For example, some AI systems have already caused alarm by propagating extremist content, exhibiting anti-Muslim bias or inadvertently revealing personal data, while a chatbot built using GPT-3 told a person to commit suicide[135].

One of the fears surrounding AI is that it will displace human workers. Automation can lead to the redundancy of certain professions, leading to unemployment and reduced job satisfaction. And not just that, as AI is now also used during the employment process: a new Law in New York city (Law 2021/144) aimed to deal with employment decision bias, requiring that a bias audit be conducted on an automated employment decision tool prior to the use of said tool, while also requiring that candidates or employees that reside in the city be notified about the use of such tools in the assessment or evaluation for hire or promotion, as well as be notified about the job qualifications and characteristics that will be used by the automated employment decision tool. To this end, as AI becomes more prevalent in policing and hiring decisions, it is essential to consider the social impact of AI development on humans.

One of the most potentially dangerous uses of AI, however, are facial recognition systems, the widespread use of which and the ever-increasing mass surveillance of individuals, especially in public places through cameras, are generating intense concerns both in terms of personal data and, more generally, the fundamental rights of citizens, such as the protection of personal data, the freedom of assembly and the right to privacy. Nowadays, the limits of these rights are being tested with highly controversial results.

Directly linked to risks to fundamental human rights is the ethical-philosophical question of whether AI systems can understand and incorporate ethical rules to their functioning and outputs; whether and

[135] Future of Life Institute. (May 2022) *General Purpose AI and the AI Act*, p. 4

how there can be "moral machines". The global community, and not just cognitive scientists but also philosophers, legal experts and scholars has been alarmed by the risks arising as the presence of AI in human life becomes stronger. For instance, when an AV faces the decision to save the occupant or a pedestrian, what should the vehicle's decision be? Research and collaboration between technology developers and ethicists will determine how to address these ethical concerns as AI technology advances. For example, the "Moral Machine" experiment gathered decisions by users using moral dilemmas where a driverless car must choose the lesser of two evils, such as killing two passengers or five pedestrians, or a homeless person as opposed to a business executive[136]. In countries where the rule of law is particularly strong—like Japan or Germany—people were more likely to kill jaywalkers than lawful pedestrians. But, even with these differences, universal patterns revealed themselves, as most players sacrificed individuals to save larger groups, most players spared women over men, and dogs were more likely to be spared than cats and more likely to be spared than criminals[137].

Ethics, however, is a value associated with humanity and conscience and that does not have a uniform meaning; on the contrary, it depends on each person's cultural background, upbringing, educational level, personal values and so forth. Therefore, how can morality be determined? And how can ethical guidelines be set on a wider context and then incorporated into an AI system? In the attempt to answer these questions, the field of the study of "robot ethics" or "AI ethics" has been formed in order to examine how ethical rules could be translated into algorithmic code[138].

---

[136] At www.moralmachine.net.

[137] Lester, C. for the New Yorker (2019) A Study on Driverless-Car Ethics Offers a Troubling Look Into Our Values, at: https://www.newyorker.com/science/elements/a-study-on-driverless-car-ethics-offers-a-troubling-look-into-our-values.

[138] On that, see the bottom-up and top-down moral approaches at: Wallach, W., Allen, C. & Smit, I. Machine morality: bottom-up and top-down approaches for modelling human moral faculties. *AI & Soc* 22, 565–582 (2008). https://doi.org/10.1007/s00146-007-0099-0

2.6.4.1. *From fundamental rights to ethical AI*

Microsoft's chatbot Tay learnt to chat by analysing and engaging in conversations with humans on Twitter. Within 24 hours, Tay spoke like an angry, confused, racist misogynist[139]. In the words of P. Boucher, "*in a way, Tay is a collective failure because it acquired its unpleasantness from how humans tend to interact online*[140]". In 2017, a German government commission released some guidelines for automated vehicles, among which one was that "*In the event of unavoidable accident situations, any distinction based on personal features (age, gender, physical or mental constitution) is strictly prohibited*[141]." The protection of human dignity and upholding the most fundamental human rights in the context of AI is of outmost importance if humans hope to truly regulate AI and thus, control, AI use. Thus, the field of "**AI ethics**" has emerged in response to the range of individual and societal harms that can be caused by the misuse, abuse, poor design or unintended consequences of AI systems, as a set of values, principles and techniques that use widely accepted standards of right and wrong in order to guide ethical behaviour in the development and use of AI technologies.

AI Ethics promote the protection of human rights in the AI industry. The Treaty on European Union and the Charter of Fundamental Rights of the EU provide some fundamental rights examples: the first "*is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities. These values are common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity and equality between women and men prevail*." (art. 2); the second provides for specific provisions on the right to life, the right to the integrity of the person[142], the right to liberty and security, the right to respect for private and

---

[139] Hunt, E. for the Guardian (2016) *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*, at: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter.

[140] Boucher, P. (2020) "*Artificial intelligence: How does it work, why does it matter, and what can we do about it*?", Panel for the Future of Science and Technology EPRS, European Parliamentary Research Service, p. 22.

[141] Federal Ministry of Transport and Digital Infrastructure, Ethics Commission, *Automated and Connected Driving*, at: https://bmdv.bund.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.pdf?__blob=publicationFile, p. 7.

[142] Respect for their physical and mental integrity.

family life, the freedom of thought, conscience, religion and expression, the right to non-discrimination[143] and many more. On an international level, the Universal Declaration of Human Rights (UDHR) as adopted by the United Nations General Assembly on 10 December 1948 sets out 30 basic human rights, which include the right to freedom and equality and the right to privacy.

Human dignity encompasses the idea that every human being possesses an "intrinsic worth", which should never be diminished, compromised or repressed by others[144]; in our case, by AI. Hence, AI systems must be created with consideration for human physical and mental well-being, personal and cultural identity, and fulfillment of basic needs. It is of outmost importance to respect, serve, and protect these aspects of humanity when developing AI technology.

The **principle of equality and non-discrimination** in AI calls for the equal respect of all human beings with equal respect for their moral worth and dignity. This goes beyond simply avoiding discrimination based on objective justifications. In the context of artificial intelligence, equality means that the system's operations should not produce *biased* results. A key question is whether AI can be truly neutral and fair or whether it is doomed to reflect the stereotypes of the society by which it is created and operates in. Additionally, it is important to show respect for potentially vulnerable individuals and groups through AI, such as women, people with disabilities, ethnic minorities, children, and others who may be at risk of exclusion. To achieve this, the data used to train AI systems should be inclusive and represent diverse population groups. It is important to use accurate and complete data to avoid incorrect conclusions, which means that the data itself should be diverse.

To this end, various ethical guidelines of AI use have been established; the most known of all is the **UNESCO Recommendation on the Ethics of Artificial Intelligence**, as adopted by all 193 UNESCO

---

143 Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited (art. 21).
144 Christopher McCrudden, Human Dignity and Judicial Interpretation of Human Rights, *European Journal of International Law*, Volume 19, Issue 4, September 2008, Pages 655–724, https://doi.org/10.1093/ejil/chn043.

Member States in November 2021 and published in 2022[145]: one of its values provides that "*No human being or human community should be harmed or subordinated, whether physically, economically, socially, politically, culturally or mentally during any phase of the life cycle of AI systems. Throughout the life cycle of AI systems, the quality of life of human beings should be enhanced, while the definition of "quality of life" should be left open to individuals or groups, as long as there is no violation or abuse of human rights and fundamental freedoms, or the dignity of humans in terms of this definition*".

2.6.4.2. An ethical approach to the AI Lifecycle

The Independent High-Level Expert Group (HLEG) on AI set out **four ethical principles**, rooted in fundamental rights, which must be respected in order to ensure that AI systems are developed, deployed and used in a trustworthy manner[146]:

a) Respect for human autonomy: AI should be designed to augment, complement and empower human cognitive, social and cultural skills and also follow human-centric design principles and leave meaningful opportunity for a "human-in-the-loop".

b) Prevention of harm: the protection of human dignity as well as mental and physical integrity of human persons is vital.

c) Fairness: ensuring equal and just distribution of both benefits and costs of AI use and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatization.

d) Explicability: the decision-making processes of the AI system should be explicable; other explicability measures (e.g., traceability, auditability and transparent communication on system capabilities) may also be required.

[145] Available at: https://unesdoc.unesco.org/ark:/48223/pf0000380455.locale=en
[146] High-Level Expert Group on AI (2019) *Ethics Guidelines for Trustworthy AI*, p. 11-13.

a) An Ethical Impact Assessment

An ethical evaluation of AI systems is proposed at all stages of their operation, i.e., the decision to create them, their design and afterwards, in order to examine each time who controls them, who benefits from their operation, who is adversely affected, how they are used, whether they exhibit discriminatory behaviour and, in general, whether they respect ethical and legal requirements.

UNESCO has sought to develop an **Ethical Impact Assessment** (EIA) tool methodology, in support of regulatory action, while OECD promotes human-rights due diligence such as human-rights impact assessments (HRIAs). Focusing on the EIA, it can be a structured process which helps AI project teams, in collaboration with the affected communities, to identify and assess the impacts an AI system may have, allowing to reflect on its potential impact and to identify needed harm prevention actions. An AI ethical impact assessment could be the solution towards establishing an ethical groundwork for AI systems to develop; it involves a process of identifying ethical implications of AI systems, assessing their significance, and proposing corrective measures to avoid negative societal impacts. The purpose of the AI ethical impact assessment is to ensure that AI systems are designed and implemented in a way that maximizes their benefits while minimizing their negative impacts. It also helps to foster trust and accountability, since trust between stakeholders and the technology is helped to be built when AI systems are designed and implemented in a transparent and ethical manner and AI systems actors can be held accountable in the event of unintended consequences. For example, the AI Fairness 360 toolkit (AIF360) by IBM is an open-source software toolkit that can help detect and remove bias in ML models, as it enables developers to use state-of-the-art algorithms to regularly check for unwanted biases from entering their ML pipeline and to mitigate any biases that are discovered[147].

However, there are several challenges to conducting an EIA. One of the challenges is that AI systems are complex, which makes it difficult to identify all their ramifications, as the design of AI systems is usually characterised by several layers of decision-making components, the effects of which may be difficult to determine. Furthermore, concrete ethical standards and principles that guide AI are not yet fully established. While there are general ethical principles such as fairness, accountability,

---

[147] An open project at IBM, see: https://www.ibm.com/opensource/open/projects/ai-fairness-360/.

transparency, and privacy, there is no consensus on how these principles should be applied to AI systems and to which degree.

To this end, collaboration and stakeholder engagement are critical. The input of diverse stakeholders such as regulators, industry associations, civil society organizations, and end-users is necessary to identify potential ethical implications of AI systems. It is also important to involve interdisciplinary experts to ensure an in-depth analysis of the ethical issues involved. Moreover, it is essential to consider the effects of AI on different -if not all- aspects of society on an ongoing basis, so no aspect is left out (e.g., minorities, labour rights, environmental protection, omission of people of colour bias, etc.). Maybe also deploying regulatory tools to make ethical frameworks more concrete could contribute to the foster trust, accountability, and credibility in AI systems.

b) Combating discrimination and biases in AI

The concept of discrimination and the concept of bias in AI are two concepts with different formal meanings, although they are often intertwined. Discrimination centres around discriminatory and unequal treatment as regards race, gender, etc. and is considered illegal by international, European and (most) national laws, whereas the term "bias" is used to convey stereotypical perceptions that lead to statistical deviations in AI; it is much more difficult to detect and usually stems from much deeper sources within society itself. Bias is when an error systematically favours a specific subset of data or a specific subpopulation. For example, if a variable's predicted value is consistently lower for one subgroup in the data, such as the salary of women with respect to equally qualified men for an equivalent job, the salary variable is biased[148]. Similarly, a consumer study of an image search on a popular search engine revealed that only 11% of results for the term "CEO" were female while at the

---

[148] IDB-OECD (2021), *Responsible use of AI for public policy: Data science* toolkit, https://publications.iadb.org/publications/english/document/Responsible-use-of-AI-for-public-policy-Data-science-toolkit.pdf.

time actually 27% of US CEOs were female[149] and studies have found that 80% of Black mortgage applicants are more likely to be rejected, along with 40% of Latino and 70% of Native American applicants[150].

OECD mentions certain cases of biases, which include:

- Historical bias: Pre-existing patterns in the training data, e.g., societal biases.
- Representation bias (and limited features): Incomplete information due to missing attributes, inadequate sample size, or total or partial absence of data from sub-populations[151].

Biases in computer systems have been identified as early as 1996 by Friedman and Nissenbaum, who mentioned "preexisting bias" having its roots in social institutions, practices, and attitudes: either by individuals who have significant input into the design of the system, such as the client commissioning the design or the system designer (e.g., a client embeds personal racial biases into the specifications for loan approval software) or by society at large, such as from organisations (e.g., industry), institutions (e.g., legal systems), or culture at large (e.g., gender biases present in the larger society that lead to the development of educational software that overall appeals more to boys than girls)[152], AI systems are largely affected by the prejudices of their human creators. It should be mentioned that biases are not deliberately inserted into the code; more often than not, biases are deeply rooted into the developer's consciousness, as affected by the society they have been born and raised in, and they can enter an AI system completely unintentionally.

In order to reduce and potentially eradicate biases, the need to emphasise inclusivity and diversity in the AI ecosystem has been mentioned; it is crucial that actors of all genders, races and backgrounds are

---

[149] Langston J. for UW News (2015) *Who's a CEO? Google image results can shift gender biases*, available at: https://www.washington.edu/news/2015/04/09/whos-a-ceo-google-image-results-can-shift-gender-biases/.

[150] Hale, K. (2021) *A.I. Bias Caused 80% Of Black Mortgage Applicants To Be Denied, available at: https://www.forbes.com/sites/korihale/2021/09/02/ai-bias-caused-80-of-black-mortgage-applicants-to-be-denied/*.

[151] OECD DIGITAL ECONOMY PAPERS No. 349 (2023) *ADVANCING ACCOUNTABILITY IN AI GOVERNING AND MANAGING RISKS THROUGHOUT THE LIFECYCLE FOR TRUSTWORTHY AI*, p. 24.

[152] Friedman, B., Nissenbaum, H. Bias in Computer Systems, *ACM Transactions on Information Systems,* Volume 14 Issue 3, p. 334.

involved in AI processes. Furthermore, an important strand of literature seeks to implement mathematical fairness metrics to assess a model's impartiality[153]; hence, there is a great need for "**debiasing**" of AI systems: that is, the ex-ante removal of all factors that can lead to biased results (e.g., gender), or, more generally, the application of select methods to address bias by achieving certain forms of statistical parity[154]. For example, a thorough analysis of the training data should be made, in order to identify and remove potential biases[155].

c) Approaches of ethical principles integration into AI

The integration of ethical principles into AI can be done by either using the "**top-down**" or "**bottom-up**" approaches. The top-down refers to the existence of certain rules for the operation of the AI, which are converted into an algorithm and incorporated into the AI. These rules may come from legislation, religion, general ethics, philosophy and more[156]. The bottom-up (or stochastic) approach, by contrast, enables a robot to "learn" from past experience and revise its algorithm over time by instilling certain core values and limitations to the system as it is "born"[157]; one benefit of this method is that it enables AI to adjust to the current moral and social circumstances, thus creating a flexible connection with society. However, a drawback is that the AI system may display unforeseeable actions that harm specific groups without the ability to investigate and determine the cause of the issue[158].

The best solution for managing AI involves using both top-down and bottom-up approaches, thus creating a third "**hybrid**" approach. This allows for oversight and control, as well as flexibility and

---

[153] Bellamy, R. et al. (2018) *AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias,* https://doi.org/10.48550/arXiv.1810.01943.

[154] Balayn, A., Gürses, S. (2021). Beyond Debiasing: Regulating AI and its inequalities, European Digital Rights, p. 23.

[155] On debiasing models, see for example Saravanan, A., *FineDeb: A Debiasing Framework for Language Models, at: https://doi.org/10.48550/arXiv.2302.02453.*

[156] See Wallach, Wendell, and Colin Allen, Moral Machines: Teaching Robots Right from Wrong (New York, 2009; online edn, Oxford Academic, 1 Jan. 2009), https://doi.org/10.1093/acprof:oso/9780195374049.001.0001.

[157] See UNESCO (2017). *Report of World Commission on the Ethics of Scientific Knowledge and Technology on Robotics Ethics*, http://unesdoc.unesco.org/images/0025/002539/253952E.pdf.

[158] Tasioulas, John, First Steps Towards an Ethics of Robots and Artificial Intelligence (June 30, 2019). *Journal of Practical Ethics*, Volume 7, Number 1, June 2019, King's College London Law School Research Paper.

adaptability. The so-called "hybrid machine ethics" provide a convergence of the two approaches, typically by having a lighter moral framework than top-down processes while still allowing the experiential learning of bottom-up processes[159].

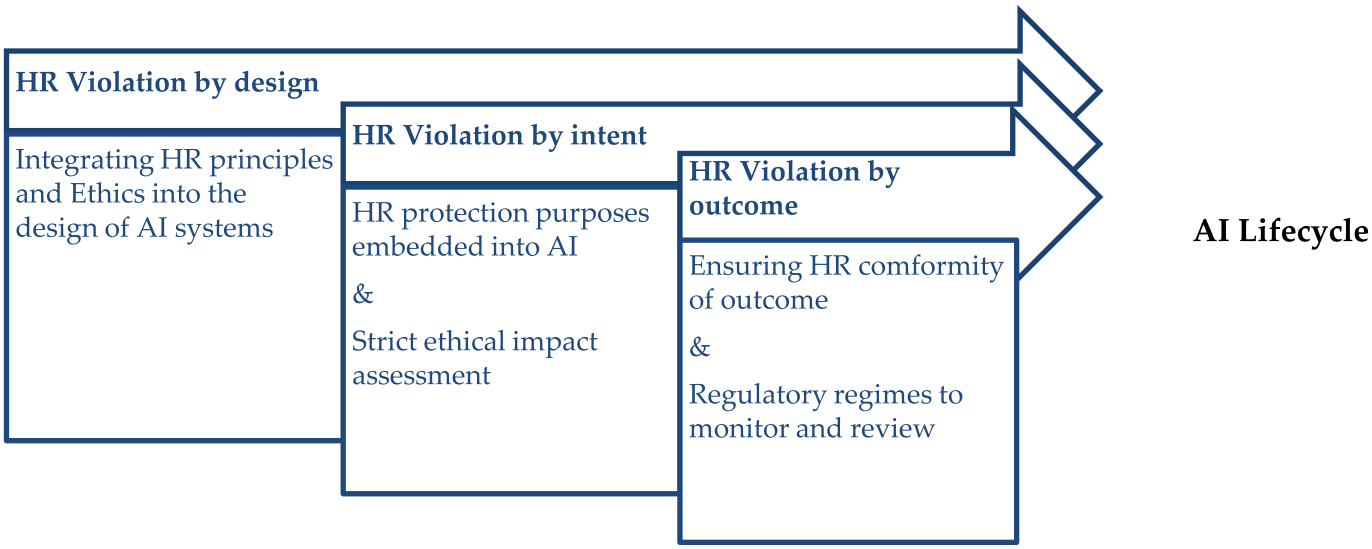


Figure 10

Kriebitz, A., (2023). *Protecting and Realizing Human Rights in the Context of Artificial Intelligence: A Problem Statement*, TUM Research Brief (additions by the author), at p. 6.

An important prerequisite for the integration of ethical norms in AI is that, on one hand, philosophers and ethical theorists do not examine the issue from a distance and purely theoretically, but seek practical applications; on the other hand, that programmers escape from their strict technological framework and deal with ethical and social concerns, realising that AI is a field with socio-technological implications[160] and legal consequences. We need anthropocentric AI to be friendly and safe, but we should avoid its

---

[159] See Shardelow, C. 2021. *Avoiding the Basilisk: An Evaluation of Top-Down, Bottom-Up, and Hybrid Ethical Approaches to Artificial Intelligence*, University of Nebraska-Lincoln.
[160] Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A. and Hall, P. (2022). *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, https://doi.org/10.6028/NIST.SP.1270.

anthropomorphisation. The great challenge for AI designers is the lasting implementation of human values in all autonomous systems[161].

2.7. The EU AI Act: an overview

The rapid advancement of AI technologies has necessitated the development of comprehensive regulatory frameworks to ensure responsible, ethical and fair AI deployment. The European Union (EU) has recently unveiled the groundbreaking EU AI Act, a legislative proposal designed to govern the ethical and legal considerations surrounding AI implementation. The EU AI Act represents a significant milestone in global AI governance, building upon preexisting regulations such as the General Data Protection Regulation (GDPR); as all-encompassing legal framework, the Act provides Europe with a pioneering role in shaping AI development, deployment, and regulation, in order to ensure compliance with the most fundamental European values and rights. As the president of the European Parliament put it, it is "*legislation that will no doubt be setting the global standard for years to come*[162]".

According to Art. 288 TFEU[163], "*A regulation shall have general application. It shall be binding in its entirety and directly applicable in all Member States*", as opposed to a Directive, which only indicates a general framework to be transposed into national Law by Member States. The choice of a Regulation as a legal instrument in the case of the AI Act, like the GDPR, is justified by the need for a uniform, comprehensive and horizontal application of the AI regulatory framework in the EU sphere and for avoiding legal

---

[161] Bar, G. (2021) *Robot personhood – what is anthropocentric artificial intelligence good for*? in: Lai, L., Świerczyński, M. (eds.), Legal and Technical aspects of Artificial Intelligence, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw, p. 51.

[162] O' Carroll, L. for the Guardian. (2023). *EU moves closer to passing one of world's first laws governing AI*, at: https://www.theguardian.com/technology/2023/jun/14/eu-moves-closer-to-passing-one-of-worlds-first-laws-governing-ai.

[163] *Treaty on the Functioning of the European Union*, OJ C 326, 26.10.2012, p. 47–390, ELI: http://data.europa.eu/eli/treaty/tfeu_2012/oj

fragmentation through national legislation. From the day of its entry into force, in will be binding for all 27 Member States and individuals can rely on the rights and obligations outlined in it without the need for national legislation to implement it, pursuant to the principle of direct effect. In this case, the legal process is streamlined by allowing individuals to directly invoke their rights before national courts, rather than having to wait for their government to transpose the Regulation into domestic law, as in the case of Directives.

Also, like the GDPR, according to L. Floridi, the AI Act is "extraterritorial", in the more technical sense that it has territorial extension[164], or aterritorial, which will further extend the so-called "Brussels effect"[165] whereby companies end up complying with EU regulations even in other countries because it is more practical to have a single approach globally, enabling the EU to extend de facto its laws internationally[166].

Apart from the AI definition, which has already been analysed in the present thesis, the AI Act, as approved by the European Parliament in June 2023[167], adopts a **risk-based approach**, distinguishing between unacceptable AI practices, high-risk AI systems, and AI applications carrying minimal risks. Through this tiered system, the Act facilitates a balanced approach to regulating AI, aiming to prevent harm while promoting innovation and economic growth.

Some selected important provisions of it, as they stand in July 2023, are the following:

---

[164] Scott, J. (2014). Extraterritoriality and territorial extension in EU law. *The American Journal of Comparative Law*, 62(1), 87–126.
[165] Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press.
[166] Floridi, L. (2021). The European Legislation on AI: a Brief Analysis of its Philosophical Approach. *Philosophy & Technology*. 34, 215–222, https://doi.org/10.1007/s13347-021-00460-9.
[167] European Parliament. (2023). *Draft European Parliament Legislative Resolution 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)), https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_EN.html#_section1.

	Provisions	Article
1.	**Prohibited AI practices:** **a)** AI systems that deploy subliminal or purposefully manipulative or deceptive techniques **b)** biometric categorisation systems using sensitive or protected attributes and characteristics **c)** social scoring, profiling and evaluation or classification systems[168] **d)** emotion recognition systems **e)** facial recognition databases	Art. 5
2.	**High-risk AI systems:** **a)** AI systems that are used in products falling under the EU's product safety legislation, as provided for in Annex II of the Act. **b)** Systems referred to in Annex III, which include Biometric and biometrics-based systems, management and operation of critical infrastructure, law enforcement, assistance in legal interpretation and application of the law, and more.	Art. 6 and Annexes II-III
3.	**Foundation models - Generative AI:** **a)** Transparency obligations, ensuring safeguards for design and development, make public training data summary of data protected under copyright law	Art. 28b, 52
4.	**Mitigation of biases through data governance measures**	Art. 10, 28b
5.	**Conformity assessment procedures to ensure accountability**	Art. 43
6.	**Regulatory Sandboxes** to facilitate the development and testing of innovative AI systems under strict regulatory oversight	Art. 53
7.	**Promotion of AI R&D** in support of socially and environmentally beneficial outcome	54a
8.	**Extraterritorial reach of the AI Act**	Art. 2(ca-cc)

[168] See Euronews. (2020). *Social scoring: Could that Facebook post stop you getting a loan or a mortgage?*, at: https://www.euronews.com/2020/06/24/social-scoring-could-that-facebook-post-stop-you-getting-a-loan-or-a-mortgage

The EU AI Act represents a groundbreaking legislative proposal that attempts to be the first to address the multifaceted challenges of AI governance. However, it is still a long way from entering into force.

SECTION 3: Artificial Intelligence and Green Transformation

3. AI and Sustainability: the digital and green transition at a crossroads

3.1. The Twin Transition

The climate crisis, or "climate change" is unequivocally linked to human activities that led to global warming through the emission of greenhouse gases[169]. In order to tackle, to the extent possible, to limit and to mitigate the emissions attributed to humans, the international community has gradually attempted to shift the attention of national policies towards environmental protection, energy efficiency, sustainable agricultural practises, renewable energy and circular production and consumption of resources ("circular economy[170]"), between others. With the pioneering prowess of the EU, the term "**climate neutrality**", i.e., the state of achieving net zero greenhouse gas (GHG) emissions by balancing those emissions so they are equal (or even less than) the emissions removed through the planet's natural absorption, and the term "**environmental sustainability**" have been increasingly incorporated into global, regional and national policy.

As is the term "**green transition"**, which refers to the fundamental shift in production and consumption patterns to allow us to live "within planetary boundaries". It means mitigating climate change impacts by introducing climate-friendly lifestyles and taking environmental costs into account, while including addressing the loss of biodiversity and its multiple ecosystem services that are crucial to healthy living and to resilient societies, in order to shift the global economy towards a more sustainable and

[169] Lee, H. et al. (2023) *Summary for Policymakers* in: Synthesis Report of the IPCC Sixth Assessment Report, p. 4.
[170] Circular Economy is a new model that does not exhaust the natural resources but aims to make it increasingly possible for the utilization and reuse of the goods already produced, in the context of sustainability. In a circular economy, the value of products, materials and resources is kept in the system, for as long as possible and the waste production is minimised. Circular economy is essentially a consumption and economic model that will not deplete the already diminished natural resources but will aim at an ever-greater utilisation and reuse of the goods and products already produced. Whereas in the linear economy products are discarded at the end of their life, thus creating tonnes of waste per year, the aim of the circular economy model is to enable goods and products to remain in the market cycle for as long as possible.

environmentally-friendly direction. The primary objective of the green transition is to reduce carbon emissions, the world's dependence on fossil fuels, and promote the use of renewable energy sources, in order to address major global issues such as climate change, pollution, and resource depletion, and is essential for improving the well-being of both humans and the planet; it will also greatly benefit the economy and public health, as clean energy reduces the cost of energy production, leading to a decline in energy bills for the average citizen; at the same time, if air pollution is decreased, public health is improved.

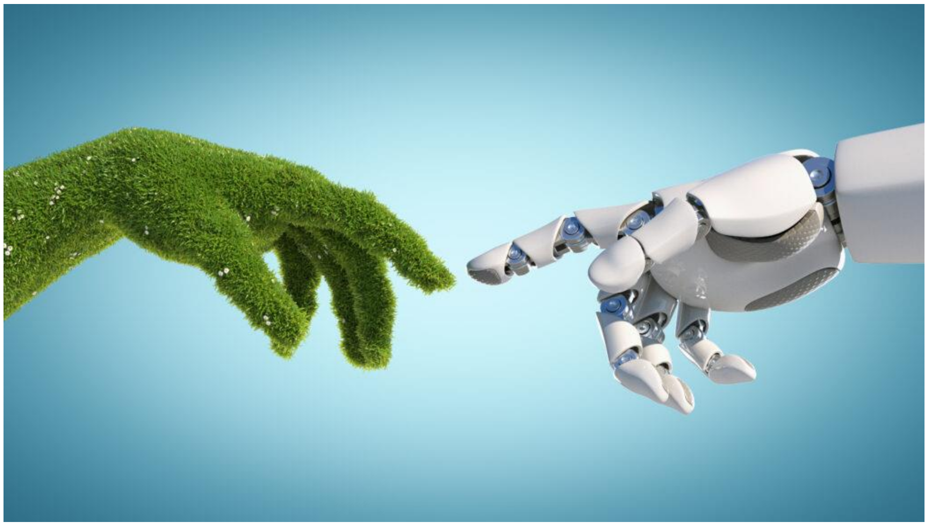
The transition to green energy and the adoption of circular economy and sustainable agriculture practices, *inter alia*, is dependent on government policies and regulations and requires the support of the private sector for innovation and development. And not just the private sector's back is needed, as financing models must be developed to support green energy initiatives and climate change education of the citizens be incorporated in societies, having in mind that the green transition is essential for the future of mankind. It will not happen on its own and requires a political, financial and societal push; with the adequate backing from governments, consumers, and private sector entities, the "greening" of technology and society can be the decisive factor of the future of our world.

On the other hand, the **digital transition** refers to the widespread adoption of digital technologies in various aspects of society, including communication, business, education, environment and entertainment, which has been driven by the rapid development of computer hardware and software, the internet, and mobile devices. In recent years, this transition has accelerated dramatically since the outbreak of the COVID-19 pandemic, which has forced many people to rely on digital tools for work, education, and social activities. The rapid digitalisation of society has enabled unprecedented levels of connectivity and communication, breaking down barriers of distance, time, and language, access to information and engagement in political and cultural activities in the "digital world".

As regards the interplay between the two transitions, which is known as "**the Twin Transition**", technology has been hailed as the most promising instrument, or "a force multiplier[171]" to achieve

[171] CIO (December 9th, 2022) *The AI Sustainability Paradox – And How to Solve It,* at: https://www.cio.com/article/415787/the-ai-sustainability-paradox-and-how-to-solve-it.html.

environmental sustainability in economy and society *quickly* and *efficiently*; in the case of AI, this has led to the usage of the term "**Green AI**[172]" as well.



Picture 2

The Twin Transition

The European Chemical Industry Council (CEFIC), *Innovation And A Supportive EU Policy Framework, The Duo That Unlocks Industry's Potential To Go Through The "Double Twin" Transition*, available at: https://cefic.org/media-corner/newsroom/innovation-and-a-supportive-eu-policy-framework-the-duo-that-unlocks-industrys-potential-to-go-through-the-double-twin-transition/

As per the principles of the Recommendation on the Ethics of Artificial Intelligence by UNESCO (paras. 17-18), the environment and ecosystems should "flourish" under ethical AI Use:

"*Environmental and ecosystem flourishing should be recognized, protected and promoted through the life cycle of AI systems. Furthermore, environment and ecosystems are the existential necessity for humanity and other living beings to be able to enjoy the benefits of advances in AI. All actors involved in the life cycle of AI systems must comply with applicable international law and domestic legislation, standards and practices, such as precaution,*

[172] See Schwartz, R., Dodge, J. et al. (2019) Green AI, https://doi.org/10.48550/arXiv.1907.10597.

designed for environmental and ecosystem protection and restoration, and ***sustainable development****. They should reduce the environmental impact of AI systems, including but not limited to its carbon footprint, to ensure the minimization of* ***climate change*** *and environmental risk factors, and prevent the unsustainable exploitation, use and transformation of natural resources contributing to the deterioration of the environment and the degradation of ecosystems".*

Similarly, para. 25 considers the implementation of risk assessment procedures and the adoption of measures in order to preclude the occurrence of environmental harm, pursuant to the "do no harm" principle and the precautionary principle, as important. Furthermore, para. 31 is dedicated to **sustainability** itself:

"*The development of sustainable societies relies on the achievement of a complex set of objectives on a continuum of human, social, cultural, economic and environmental dimensions. The advent of AI technologies can either benefit sustainability objectives or hinder their realization, depending on how they are applied across countries with varying levels of development. The continuous assessment of the human, social, cultural, economic and environmental impact of AI technologies should therefore be carried out with full cognizance of the implications of AI technologies for sustainability as a set of constantly evolving goals across a range of dimensions, such as currently identified in the Sustainable Development Goals (SDGs) of the United Nations*". Finally, three separate recommendations are given in the environment and ecosystems policy field[173].

The Twin Transition[174] has also been called "**The Green and the Blue**[175]" formula. The interplay between the digital and the green transition has been a pillar of EU policy following the EU Green Deal, with the publication of the "A Europe fit for the digital age" strategy[176] and the direct mentions of the ecological and digital transitions in various European Commission documents. The EU Green Deal communication itself notes that "*digital technologies are a critical enabler for attaining the sustainability goals of the Green Deal*

---

[173] See paras. 84-86.

[174] For more on the Twin Transition see Muench, S., Stoermer, E., Jensen, K., Asikainen, T., Salvi, M. and Scapolo, F., (2022) *Towards a green and digital future*, EUR 31075 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-52451-9, doi:10.2760/977331, JRC129319.

[175] Floridi, L., Nobre, A.-C. (2020) *The Green and the Blue: How AI may be a force for good*, OECD Forum, available at: https://www.oecd-forum.org/posts/the-green-and-the-blue-how-ai-may-be-a-force-for-good.

[176] European Commission, *Shaping Europe's Digital Future*, at: https://commission.europa.eu/system/files/2020-02/communication-shaping-europes-digital-future-feb2020_en_4.pdf.

in many different sectors" and "(…) *data, combined with digital infrastructure (e.g., supercomputers, cloud, ultra-fast networks) and artificial intelligence solutions, facilitate evidence-based decisions and expand the capacity to understand and tackle environmental challenges"*, while domains in which digital technologies are expected to play a role include energy grids (p. 6), consumer products (p. 8), air and water pollution monitoring (p. 9), mobility (p. 10), and food and agriculture (p. 12). Following on that, the "*New industrial strategy for Europe*" asserts that Europe needs an industrial sector that becomes greener, more digital and circular[177].

Hence, making sure that the digital transition, in particular through the staggering production and utilisation of AI systems, becomes an ally of the green transition, will be one of the most important challenges of our society; perhaps it already is. For example, urban areas are home to more than 50% of the world's population[178] and are where most of the economic activity is located and most of the building activity takes place; simultaneously, according to studies, they are responsible for approximately 75% of global CO2 emissions. The UN predicts that 80% of the world's population will live in urban areas by 2050[179], while urban areas are identified as particularly vulnerable to the impacts of climate change and the expected increase in their population will intensify the urban GHG emissions; in this case, the digitalisation of urban functions towards their "decarbonisation" and the creation of smart cities and communities seems like an ideal mission for AI.

AI-enabled technologies have huge potential to support positive climate action. From digital twin technology that model the Earth, to algorithms to make data centres more efficient, AI applications already support the green transition[180]. However, the policy documents tend to presume a harmonious relationship between the digital and ecological transitions, overlooking the trade-offs that may need to

---

[177] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, *A New Industrial Strategy for Europe*, COM/2020/102 final, p. 2-3.

[178] United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Population Prospects 2019: Highlights*, p. 6.

[179] United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Urbanization Prospects: The 2018 Revision*

[180] OECF Panel (Nov 15 2022) "*What is the environmental footprint of artificial intelligence*?", Informed by experts from the OECD.AI Expert Group on AI Compute and Climate and the Global Partnership on AI (GPAI) Responsible AI Working Group (RAI)/

be struck between them, and how this could or should be done[181]. Apart from their undeniably positive potential, AI systems also raise important concerns regarding the natural resources they need to be produced and their environmental operational costs, as well as the GHG the emit.

3.2. AI: a double-edged sword towards environmental sustainability?

3.2.1. The environmental footprint of AI

Various research over the years has highlighted the negative environmental impact of Information and Communication Technology (ICT) production and use; for example, in 2017 a forecast was made that by 2025 data centres could account for 10% of global electricity use[182]. According to estimates, the global CO2 emissions of the Information and Communications Technology (ICT) sector account for around 2% of global CO2 emissions, but this figure is hard to estimate precisely given the distributed nature of global computing infrastructure[183].

The computational power, or "**compute**," of AI systems refers to the number of computational resources needed to train and run a ML system. Typically, the bigger and more complex a system is, and the larger the dataset on which it is trained, the greater the amount of compute required. Therefore, the system's demands in energy use, water consumption and construction material need are also increased.

Focusing on AI, the amount of compute used to train the largest DL models has increased 300,000 times in the time span between the years 2012-2018[184]; it has been estimated that training the open-source ML

---

[181] Cowls, J., Tsamados, A., Taddeo, M. et al. The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Soc 38*, 283–307 (2023). https://doi.org/10.1007/s00146-021-01294-x
[182] Anders, A. (2017). Total Consumer Power Consumption Forecast, Nordic Digital Business Summit Conference.
[183] See Malmodin, J., & Lundén, D. (2018). The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015. *Sustainability*, 10(9), 3027. https://doi.org/10.3390/su10093027.
[184] Amodei, D., Hernandez, D. (2018) *AI and Compute*, at: https://openai.com/blog/ai-and-compute.

framework called "BERT" is roughly equivalent to a trans-American flight[185], carbon-wise (each passenger flying from London to New York and back generally generates about 1 tonne of CO_2[186]). Similarly, recent findings on "BLOOM", a 176-billion parameter LLM, have concluded that its training run emitted 25 times more carbon than a single air traveller on a flight from New York to San Francisco[187].

As evidenced below, since 2010, of all ML systems, LLMs are increasingly demanding the most computational resources.

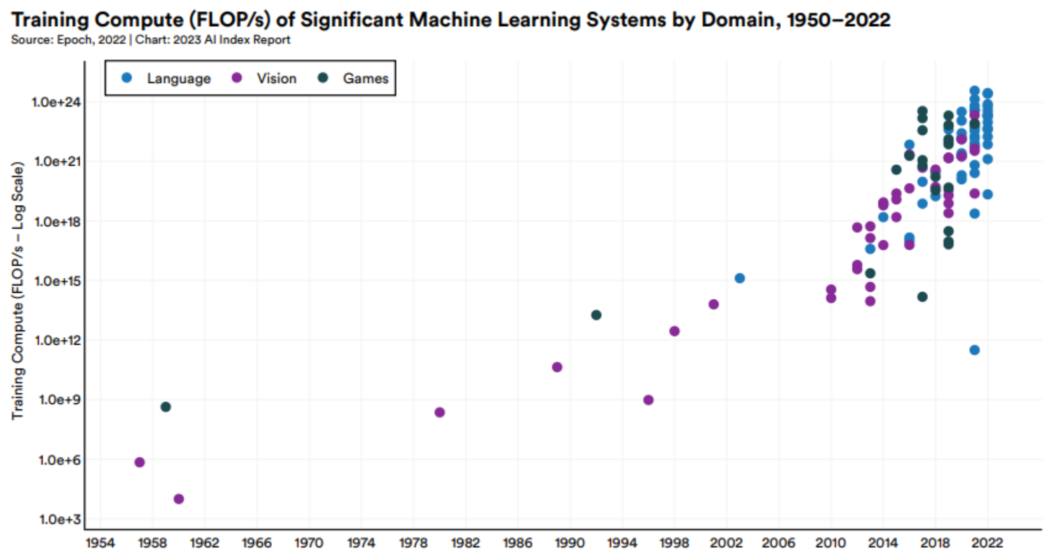


Figure 11

Training Compute of Significant ML Systems by domain, 1950-2022

Maslej N., et al. (2023). *The AI Index 2023 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, p. 57.

---

185 Strubell, E., Ganesh, A., McCallum, A. (2019). *Energy and policy considerations for deep learning in NLP*, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, p. 3648.

186 Kommenda, N. for The Guardian (2021). *How your flight emits as much CO2 as many people do in a year*, at: https://www.theguardian.com/environment/ng-interactive/2019/jul/19/carbon-calculator-how-taking-one-flight-emits-as-much-as-many-people-do-in-a-year.

187 Maslej N., et al. (2023). *The AI Index 2023 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, p. 120.

To put the emissions of said ML Systems into perspective, the following real-life examples are provided:

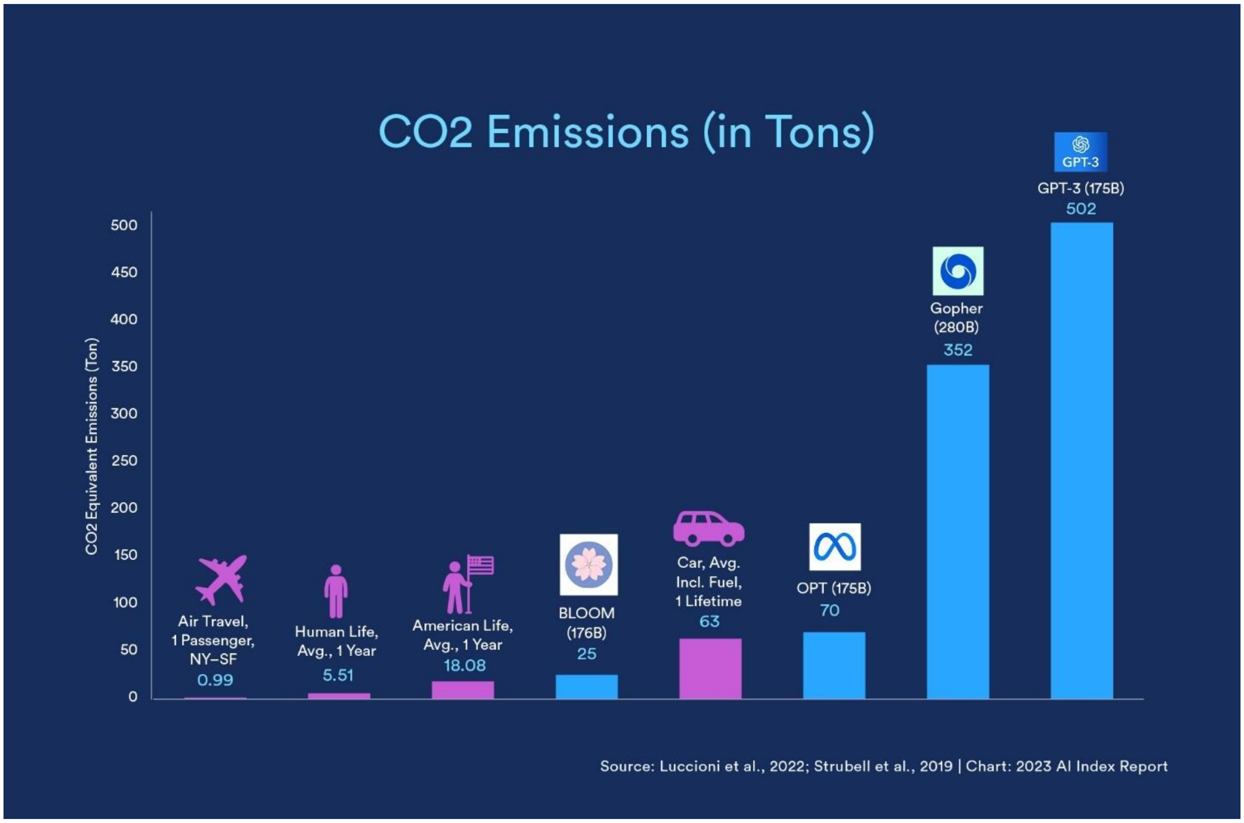


Figure 12

CO2 Equivalent Emissions in Tonnes by ML Models and real-life examples

Source: https://hai.stanford.edu/news/2023-state-ai-14-charts

For reference, the average emissions per person on an annual basis in Germany amount to around ¼ of the above emissions of an American, that is, 5.09 tonnes[188].

[188] Our World in Data (2021), CO2 per capita, at: https://ourworldindata.org/grapher/co2-per-capita-marimekko?tab=table.

However, the environmental impact of the ICT sector, and in particular, AI systems, should not only measured in GHG tonnes. AI can also negatively contribute to the use of natural resources; for example, further automating extraction in the mining sector through "mining robots" can take over the mining tasks may further increase the yield of rare materials and increase their depletion rate[189]. Moreover, AI can have system-level impacts on society that affect the climate: for example, even though AV technology can introduce efficiency gains for driving, the technology also lowers the barrier to driving and can induce new demand for individualised transportation. It is thus possible that self-driving cars will cause the total emissions associated with transportation to increase[190].

3.2.2. How to measure the environmental impact of AI?

A "carbon footprint" accounts for the GHG (greenhouse gas) emissions of a device or activity, expressed as **carbon dioxide equivalent** (CO2eq). CO2eq are calculated based on comparing the global-warming potential (GWP) of different GHG to that of carbon dioxide (CO2). For instance, methane has a 100-year GWP 25 times that of CO2– this means that it is equal to 25 CO2eq[191]. When applied to a product like a smartphone, a carbon footprint estimation considers emissions that occur during constituent activities, like the extraction of raw materials, manufacturing, transportation, lifetime usage and how the product is disposed of[192]; in other words, one would more correctly calculate the carbon (or environmental, in general) footprint of a product by examining its whole "**life cycle**" and not just its operating phase. In the case of AI things are more complicated, as determining the carbon footprint of a type of AI system (e.g., an LLM) or the entire sector (as in the case of Information Communication Technologies, ICT) can

[189] Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., & Zhang, W. (2018). The Rise of Artificial Intelligence under the Lens of Sustainability. *Technologies, 6(4)*, 100. https://doi.org/10.3390/technologies6040100

[190] Global Partnership on AI Report in collaboration with Climate Change AI and the Centre for AI & Climate. (2021). *Climate Change and AI: Recommendations for Government Action*, p. 48.

[191] Luccioni et al. (2022) *Estimating the Carbon Footprint of BLOOM*, https://doi.org/10.48550/arXiv.2211.02001, p. 3.

[192] Cowls, J., Tsamados, A., Taddeo, M. et al. (2023) The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Society* 38, 283–307. https://doi.org/10.1007/s00146-021-01294-x., p. 290.

be a daunting task that yields only partial results due to transparency issues, corporate/industrial confidentiality claims and various methodological challenges of GHG monitoring.

Several ways to measure the environmental impact of AI have been proposed; however, a global standard of measurement is absent; exactly what is measured and how it is measured depends largely on company strategy and literature used. Furthermore, neither the accounting methodologies for reporting carbon emissions of AI are standardised. For example, Lacoste et al. created an ML Emissions Calculator, which takes as input the details regarding the training of an ML model: the geographical zone of the server, the type of GPU (Graphical Processing Units) and the training time, and gives as output the approximate amount of CO2eq produced[193]. Similarly, researchers have found that three factors make AI research "red", that is, massive AI compute:

a) the cost of executing the model,

b) the size of the training dataset which controls the number of times the model is executed, and

c) the number of hyperparameter experiments, which controls how many times the model is trained.

It was calculated that the total cost of producing a result in ML increases linearly with each of these quantities[194].

In the case of data centres, **Power Usage Effectiveness (PUE)** is a metric used to evaluate their energy efficiency. PUE is the ratio of the total amount of energy used by a computer data center facility, including air conditioning, to the energy delivered to computing equipment. The higher the PUE, the less energy efficiency is achieved[195]. For example, Google Cloud Services has an average PUE of 1.1, meaning that only 11% of their total energy usage is not used for the servers themselves[196].

---

[193] Lacoste, A., Luccioni, A. Schmidt, V., Dandres, T. (20219) *Quantifying the Carbon Emissions of Machine Learning*, https://doi.org/10.48550/arXiv.1910.09700, p. 3. The calculator is available at: https://mlco2.github.io/impact/#compute.
[194] Dhar, P. The carbon impact of artificial intelligence. *Nat Mach Intell 2*, 423–425 (2020). https://doi.org/10.1038/s42256-020-0219-9, citing the work of Schwartz, R., Dodge, J. et al. (2019) Green AI, https://doi.org/10.48550/arXiv.1907.10597.
[195] Maslej N., et al. (2023). The AI Index 2023 Annual Report, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, p. 120.
[196] *Google Data Centers efficiency: How we do it*. https://www.google.com/about/datacenters/efficiency/internal/, 2023. According to the Uptime Institute's 2021 Data Center Survey, the global average of respondents' largest data centers is around 1.57.

While there is no universally-accepted approach for assessing the environmental impacts of ML models, the Life Cycle Assessment (LCA) methodology, which aims to cover all stages of the life cycle of a product or process, has also been proposed[197]. By using this methodology, one could calculate the environmental impact of raw material extraction, production and manufacturing, its transport costs (i.e., shipping emissions) towards the operation destination, the consumption of energy and water during the operation phase as well as the end-of-life impacts (like the disposal of materials).

In general, the carbon cost "phases" during the AI's lifecycle could be depicted as follows:

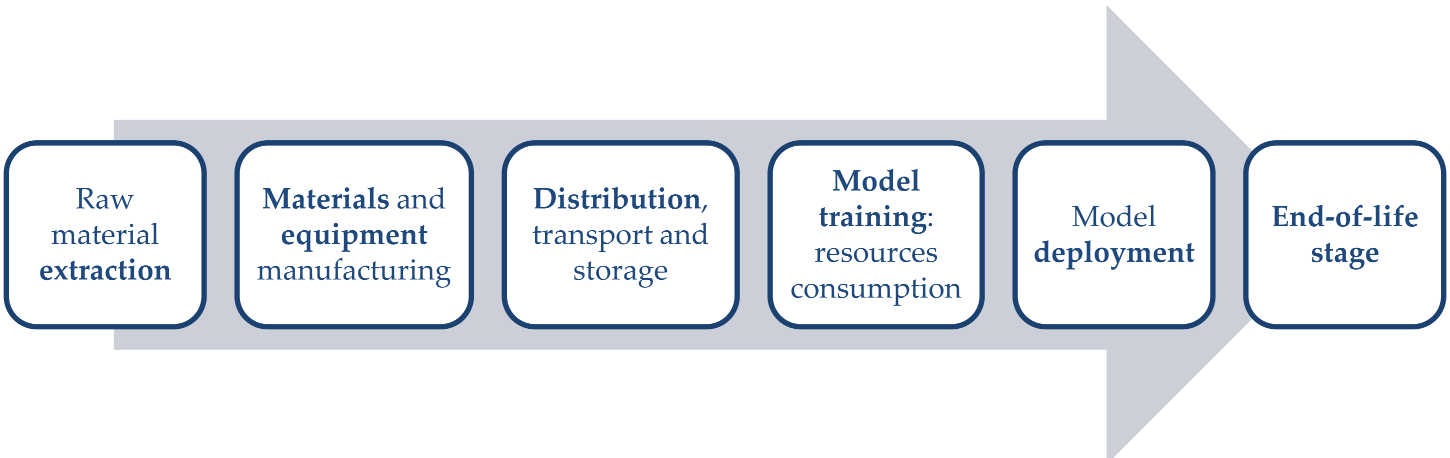


Figure 13

Created by the author

The "**raw materials**" phase relates not only to their extraction phase (for example, the environmental cost of extracting minerals like aluminium), but also to their assembly. The number of natural resources needed to build computer hardware most often than not lead to unsustainable processes that cause water and soil contamination, deforestation, groundwater pollution, biodiversity loss.

For example, on the Indonesian island of Bangka, as a Guardian investigation reports "*tin mining* [for semiconductor production] *is a lucrative but destructive trade that has scarred the island's landscape, bulldozed its farms and forests, killed off its fish stocks and coral reefs, and dented tourism to its pretty palm-lined beaches.*

[197] See Guinée, J.B.; Heijungs, R.; Huppes, G.; Zamagni, A.; Masoni, P.; Buonamici, R.; Ekvall, T.; Rydberg, T. (2011). Life Cycle Assessment: Past, Present, and Future. *Environ. Sci. Technol*. 45, 90–96.

The damage is best seen from the air, as pockets of lush forest huddle amid huge swaths of barren orange earth. Where not dominated by mines, this is pockmarked with graves, many holding the bodies of miners who have died over the centuries digging for tin[198]." Indonesia's national tin corporation, PT Timah, supplies companies such as Samsung directly, as well as solder makers Chernan and Shenmao, which in turn supply Sony, LG and Foxconn[199].

Furthermore, the cost of materials **production** and **manufacturing** of equipment is critical: air pollutant emissions, toxic waste disposal and water contamination from the manufacturing industry's operational activities are in no way negligible. The **distribution** of manufactured equipment from/to the other stages of production can also be environmentally damaging (oil spills, emissions, etc.), whether it is operated by the aviation or shipping industry, or over land by freight trains, commercial truck fleets etc. For example, according to data provided by the European Commission, in 2018 global shipping emissions represented 1.076 million tonnes of CO2, and were responsible for around 2.9% of global emissions caused by human activities. Hence, the decarbonisation of the transport sector should also be put at the forefront.

Afterwards, and as analysed above, the carbon footprint of AI **model training** through GHG emissions has already been calculated as particularly high; as it is during its **deployment** (operational phase), when AI compute also consumes energy and water resources, while also contributing to GHG emissions. As reported by the OECD, the United States data centre industry draws directly and indirectly from 90% of national watersheds and is one of the top ten water-consuming industries in the country[200].

However, a positive report from Denmark mentions that the excess heat from GlobalConnect's data centre in Copenhagen will be connected to the nearby municipality's district heating in 2023[201], while

---

198 Hodal, K. (2012). *Death Metal: Tin Mining in Indonesia for The Guardian*, available at: http://www.theguardian.com/environment/2012/nov/23/tin-mining-indonesia-bangka.

199 Simpson, C. for Bloomberg (2012). *The Deadly Tin Inside Your Smartphone*, available at: https://www.bloomberg.com/news/articles/2012-08-23/the-deadly-tin-inside-your-smartphone.

200 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, p. 29.

201 Nelson, M. for DigitalInfraNetwork (2023) *Next-gen submerged cooling installed in Danish data centre*, at: https://digitalinfranetwork.com/news/next-gen-submerged-cooling-installed-in-danish-data-centre/.

the Lefdal Mine Datacenter in Norway has already offered a case study in sustainable data centre architecture by using 100% renewable hydropower electricity[202]. These examples could potentially pave the way towards a broader harnessing of the "negatives" of AI and transforming them into "positives" regarding environmental impacts.

Finally, the total environmental cost of equipment dismantling, recycling (which also uses resources), collection, shipping to **disposal** sites, e.g., landfills, and the disposal itself (especially when no circular economy practices are followed) raises important environmental concerns. The amount of hazardous e-waste circulating in the world is estimated to be more than 7.3 kg per person, totalling 53.6 million metric tonnes, as reported in 2020[203], while ICT infrastructure accounts for about 12 million tons, or 25% of total global electronic waste[204].

In the EU, the Waste from Electrical and Electronic Equipment (WEEE) Directive[205] and the RoHS Directive[206] were introduced in early 2010's to tackle the issue of the growing amount of WEEE, while in March 2020 the Circular Economy Action Plan (CEAP) was published[207], which attached importance to electronic products by adopting the "Circularity of Electronic Products Initiative", aiming to address the challenges stemming from the rapid evolution of electronic products and the information and communication technologies (ICT) sector. Combined with the publication of the CEAP, the European Commission adopted a European Strategy for the industrial sector[208].

---

202 See the data centre's sustainability information on its website, at: https://www.lefdalmine.com/sustainability-3/#leading.

203 Forti, V. et al. (2020) *The Global E-waste Monitor 2020: Quantities, flows, and the circular economy potential*, at: https://collections.unu.edu/view/UNU:7737#viewAttachments.

204 Sustainable Infrastructure Alliance (2022), *The Roadmap to Sustainable Digital Infrastructure by 2030*, https://sdialliance.org/roadmap.

205 Directive 2012/19/EU of the European Parliament and of the Council of 4 July 2012.

206 Directive 2011/65/EU of the European Parliament and of the Council of 8 June 2011 on the restriction of the use of certain hazardous substances in electrical and electronic equipment.

207 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A New Circular Economy Action Plan for a Cleaner and more Competitive Europe, COM/2020/98 Final.

208 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - A New Industrial Strategy for Europe, 10.3.2020, COM(2020) 102 final. On more on the EU's circular economy policy see Kl. Pouikli, A. Delegkou, I. Tsakalogianni (2022). Circular Economy for a climate-neutral Europe: Mapping the latest policy and legislative developments *Carbon & Climate Law Review,* Volume 16, Issue 4 (2022) pp. 281 - 287 DOI: https://doi.org/10.21552/cclr/2022/4/8.

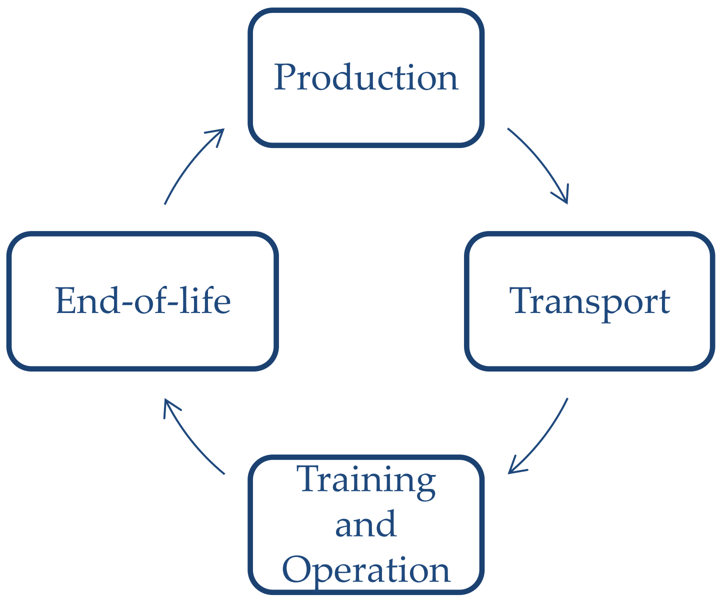


Figure 14

Created by the author

The above stages of the environmental impacts of AI are also reported by the OECD:

Direct environmental impacts AI compute resources lifecycle

Production	Transport	Operations	End-of-life
• Raw material extraction • Assembly • Manufacturing	• Distribution • Freight transportation • Handling & storage	• Energy consumption • Water consumption • Carbon footprint	• Collection & shipping • Dismantling & recycling • Waste disposal

Indirect environmental impacts AI applications

Positive impacts	Negative impacts
• Beneficial sectoral applications • Climate mitigation and adaptation • Environmental modelling and forecasting	• Harmful sectoral applications • Carbon leakage (net increase in emissions) • Consumption patterns and rebound effects

Picture 3

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

In the case of the BLOOM LLM, as regards the so-called "embodied emissions" associated to the computing equipment needed to train and deploy the model, its training represented approximately 7.57 tonnes for the servers and 3.64 tonnes for the GPUs, adding a total of 11.2 tonnes of CO2eq to its carbon footprint (which did not include the embodied emissions of the rest of the computing infrastructure like cooling equipment, due to lack of information). It is also important to add that embodied emissions of decentralized computing (e.g., desktops, laptops, smartphones), account for 40-80% of devices' lifecycle GHG emissions[209]. As regards electrical energy consumption, a total of 24.69 tonnes of CO2eq was emitted due to "dynamic energy consumption", whereas "idle energy consumption" (that is, when the system is powered on, but no processes are running) added 14.6 tonnes of CO2eq to the overall carbon footprint of the model's training[210].

[209] Kaack, L., Donti, P., Strubell, E., Kamiya, G., Creutzig, F. et al. (2021). *Aligning artificial intelligence with climate change mitigation*. ffhal-03368037f, p. 6.
[210] Luccioni et al. (2022) *Estimating the Carbon Footprint of BLOOM*, https://doi.org/10.48550/arXiv.2211.02001, p. 4.

Following on the above, the OECD has recommended also calculating the synergies, otherwise the "**interconnection**" between the separate stages of AI compute resources lifestyle; an example of such assessment is the Green Cloud Computing methodology report from the German Environmental Agency[211], which looks at:

(1) the abiotic depletion potential (i.e., use of minerals and fossil fuels). This makes the raw material indicator well suited to highlighting opportunities for saving resources in production and use and to addressing circular economy[212];

(2) global warming potential (i.e., impact on climate change);

(3) the cumulative energy demand (i.e., use of renewable and non-renewable energy); and

(4) water usage. As mentioned by the report, in the event that data centres are equipped with adiabatic cooling (called "Verdunstungskühlung" – "evaporative cooling") and are located at sites with acute water scarcity, this indicator can provide important information for the degree of their efficient operation[213].

3.2.3. The ways towards more sustainable AI: emphasis on legislation

In order to pave the way to more sustainable solutions as regards the production and operational environmental cost of AI, one solution suggested was to develop full general lifecycle assessments (as defined in ISO standards 14040:2006[214] and 14044:2006[215]) to assess their environmental impact across

---

[211] Gröger, J, Liu, R., Stobbe, L., Druschke, J., Richter, N. (2021), *Green Cloud Computing, Lebenszyklusbasierte Datenerhebung zu Umweltwirkungen des Cloud Computing*, Fraunhofer-Institut für Zuverlässigkeit und Mikrointegration, available at: https://www.umweltbundesamt.de/sites/default/files/medien/5750/publikationen/2021-06-17_texte_94-2021_green-cloud-computing.pdf.

[212] *Id*., at p. 140.

[213] *Id*.

[214] See ISO 14040:2006 Environmental management — Life cycle assessment — Principles and framework, at: https://www.iso.org/standard/37456.html

[215] See ISO 14044:2006 Environmental management — Life cycle assessment — Requirements and guidelines, at: https://www.iso.org/standard/38498.html.

their entire lifecycle. However, applying existing tools like product lifecycle assessments can be challenging due to the distinct compute needs of AI systems and software that can vary significantly in their composition and include many different individual and highly specialised products[216]. Furthermore, the vast majority of ML research and development still focuses on improving model accuracy, rather than balancing accuracy and energy efficiency[217]. So, a change on the mindset of the AI industry should also occur; as Wu et al. put it, "*To develop AI technologies responsibly, we must achieve competitive model accuracy at a fixed or even reduced computational and environmental cost*" and to develop and adopt a "sustainability mindset" for AI[218].

In 2021, Kaack at al. suggested certain levers to reduce AI-related GHG emission impacts in the public and private sectors[219]:

---

216 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, p. 23.
217 See Schwartz, R., Dodge, J., Smith, N., Etzioni, O. (2019). Green AI. CoRR, abs/1907.10597, http://arxiv.org/abs/1907.10597.
218 Wu, C-J., et al. (2022) *Sustainable AI: Environmental Implications, Challenges and Opportunities*, arXiv:2111.00364v2 [cs.LG], p. 7.
219 Kaack, L., Donti, P., Strubell, E., Kamiya, G., Creutzig, F., et al (2021). *Aligning artificial intelligence with climate change mitigation.* ffhal-03368037, p. 11.

<table>
<tr><th>Lever type</th><th>Compute-related (algorithm, infrastructure)</th><th>Application-related (immediate, systemic)</th></tr>
<tr><td colspan="3">Public sector</td></tr>
<tr><td>Economic instruments</td><td colspan="2">● Implement economy-wide or sector-specific carbon pricing to incentivize emissions reductions and mitigate rebound effects</td></tr>
<tr><td>Research, development & demonstration (RD&D)</td><td>● Support research in energy-efficient ML
● Support RD&D in energy-efficient, specialized, and low-resource hardware
● Support RD&D in data center operational efficiency</td><td>● Support interdisciplinary and applied ML research for climate-relevant applications of ML
◐ Provide mechanisms to advance the technological readiness of climate-beneficial AI applications (e.g. testbeds, demonstration projects, public procurement programs)</td></tr>
<tr><td rowspan="2">Regulation</td><td colspan="2">○ Employ a climate-cognizant technology assessment lens within AI strategies and when regulating ML-driven emerging technologies</td></tr>
<tr><td>● Implement clean electricity mandates (e.g., low-carbon portfolio standards)
● Implement efficiency standards for data center hardware and infrastructure</td><td>● Employ regulatory approaches to constrain sector-specific GHG emissions
◐ Reduce deployment barriers in relevant sectors and industries for AI applications that are beneficial to the climate</td></tr>
<tr><td rowspan="2">Best practices and standards</td><td colspan="2">◐ Develop interoperability standards for commercial ML approaches to prevent lock-in to particular solutions providers and facilitate a decentralized solutions provider space</td></tr>
<tr><td>● Develop and implement standardized metrics for evaluating model efficacy that include energy efficiency</td><td>● Implement data governance standards that spur impactful work and are mindful of privacy and ownership
● Require meaningful civic and stakeholder engagement in scoping, developing, and deploying ML-driven projects
○ Develop best practices and systematic approaches to weigh benefits and costs for ML applications</td></tr>
<tr><td>Monitoring and reporting</td><td colspan="2">◐ Develop measurement methodologies and guidance to estimate and report ML-related GHG emissions
◐ Mandate appropriate life-cycle transparency and reporting of GHG emissions for ML use cases, including both compute and application-related impacts</td></tr>
<tr><td rowspan="2">Capacity-building</td><td colspan="2">◐ Build in-house public-sector capacity in ML to facilitate governance and deployment
● Promote ML education and literacy among climate-relevant entities and in the public sector</td></tr>
<tr><td></td><td>◐ Incentivize ML workforce shifts towards climate-oriented entities (e.g., via placement programs)</td></tr>
<tr><td colspan="3">Private sector</td></tr>
<tr><td rowspan="2">Corporate climate action</td><td colspan="2">◐ Adopt organizational carbon pricing strategies that account for both compute- and application-related emissions (e.g., Scope 1, 2, and 3 emissions, including from cloud compute, as well as from products and services)</td></tr>
<tr><td>● Reduce wasteful model re-training and execution
● Make energy efficiency a central criterion in evaluating model efficacy
● Reduce GHG emissions across supply chains and product life cycle (including embodied emissions)
● Maximize energy efficiency in data centers and support related RD&D
● Shift compute load to geographies and times with lower carbon-intensity of the grid
● Purchase low-carbon electricity and invest in energy technologies to decarbonize the grid
○ Develop standardized ML platforms to facilitate rapid company-wide adoption of energy efficiency improvements</td><td>● Adjust business models to avoid ML applications that drive GHG emissions increases
● Encourage ML applications that drive GHG emissions reductions
◐ Measure and engage in voluntary reporting of the emissions impacts of ML products and services</td></tr>
</table>

Legend

● Policies that are ready to implement or already exist
◐ Policies that can be developed today
○ More analysis needed to develop policies
General climate policy levers

Picture 4

See footnote 219 for source.

As regards regulation, which is the perspective of the author of the present thesis, Kaack et al. cite clean electricity mandates and efficiency standards for hardware and infrastructure, as well as the employment of regulatory approaches for sector-specific GHG emissions; indeed, EU Directive 27/2012, as amended by Directive 2018/2002, set the target of improving energy efficiency by 32.5% across the whole of EU and to monitor this target with National Energy Efficiency Action Plans (NEEAPs[220]) every three years; in 2021, a recast proposal for the Directive was published, calling for 39% and 36% energy efficiency targets for primary and final energy consumption respectively[221]. The proposal was part of the "Delivering the European Green Deal" or "Fit for 55%" EU legislative package, as published in July 2021, which consisted of a set of inter-connected proposals[222]:

	Financial measures	**Targets**	**Standards**
1.	**Stronger Emissions Trading System** including in aviation: phase out the sector's free emissions allowances	**Updated Effort Sharing Regulation**: EU-wide reduction of **40**% in emissions by 2030.	Amendment of the Regulation on **CO2 emission standards for cars and vans**
2.	**Extending Emissions Trading** to maritime over the period of 2023--2025, road transport, and buildings	Updated Land Use Land Use Change and Forestry Regulation (**LULUCF**): new EU target of net greenhouse gas removals of 310 million tons of CO2 equivalent by 2030.	Revision of the Directive on deployment of **alternative fuels infrastructure**

[220] NEEAPs cover measures to improve energy efficiency and expected and/or achieved energy savings, including measures in the field of energy supply, transmission and distribution, as well as energy end-use, and shall be sent to the Commission, which may make recommendations.
[221] Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on energy efficiency (recast) COM/2021/558 final
[222] See COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS 'Fit for 55': delivering the EU's 2030 Climate Target on the way to climate neutrality, COM(2021) 550 final.

3.	**Updated Energy taxation Directive:** alignment of the minimum tax rates for heating and transport fuels with EU climate and environmental objectives	**Updated Renewable Energy Directive**: overall binding target to increase from 32% to **40**%.	ReFuel EU: More **sustainable aviation fuels**
4.	New Carbon Border Adjustment Mechanism	**Updated Energy Efficiency Directive:** 39% and 36% energy efficiency targets for primary and final energy consumption.	FuelEU: Cleaner maritime fuels

As evidenced above, the EU is a world leader in green, or climate policy; however, as per the suggestions of Kaack et al., **AI** "**climate assessment**" obligations have yet to be incorporated in legal systems *en masse*. To this end, the revised EU AI Act, as published in the second half of 2023, includes some encouraging provisions:

a) it proposes that AI compliance with environmental standards and a high level of environmental protection should be ensured (recital 60g),

b) that foundation models should be designed with capabilities enabling the measurement and logging of the consumption of energy and resources, and, where technically feasible, other environmental impact the deployment and use of the systems may have over their entire lifecycle (para d of Art. 28b),

c) that **fundamental rights impact assessment for high-risk AI systems should include the environmental impact of the system's "use"** (although this provision limits the system's assessment to perhaps just the deployment phase, rather than including the production and the training stages as well).

However, a **horizontal climate evaluation** of AI strategies and a **separate climate assessment** before the regulation of ML-driven technologies could be deemed particularly important in order to prevent or/and regulate the mass production and use of resource-intensive AI technologies that could be seen as dangerous to the environment. At the same time, the **Law should be used to enable and not hinder** the development of climate-positive AI technologies that contribute to the reduction of GHG emissions, the sustainable use of resources and the promotion of energy efficiency, as they will be presented next.

3.3. AI for climate neutrality: enabling the Twin Transition

3.3.1. AI and Climate Science

First and foremost, AI has contributed to the enhancement of remote sensing and weather forecasting; for example, in recent work scientists have used XAI to estimate precipitation over the contiguous United States from satellite imagery[223]. The UK's national weather service, the Met Office, incorporates AI in its weather forecasting models, as they use ML algorithms to analyse large datasets and improve the accuracy of short-term weather predictions[224].

The use of AI in short-term forecasting can also contribute to the accuracy and efficiency of the forecast; for example, UNESCO's Intergovernmental Hydrological Programme has developed an early warning system for flooding and drought in 11 African countries, which will collect, analyse and visualise flood-related images to identify and communicate practical solutions to affected communities in real time. This might entail issuing warnings to move livestock to higher ground, or to pack necessities and leave potential risk areas before the arrival of floodwaters[225].

---

[223] Hilburn, K.A., Ebert-Uphoff, I., Miller, S.D.: Development and interpretation of a neural-network-based synthetic radar reflectivity estimator using GOES-R satellite observations. J. Appl. Meteorol. Climatol. 60(1), 3–21 (2021)
[224] See Met Office (2022). *Embedding machine learning and artificial intelligence in weather and climate science and services: a framework for data science in the Met Office* (2022-2027). Exeter, UK.
[225] UNESCO (2022) *Being able to predict floods will save lives in West Africa,* https://www.unesco.org/en/articles/being-able-predict-floods-will-save-lives-west-africa.

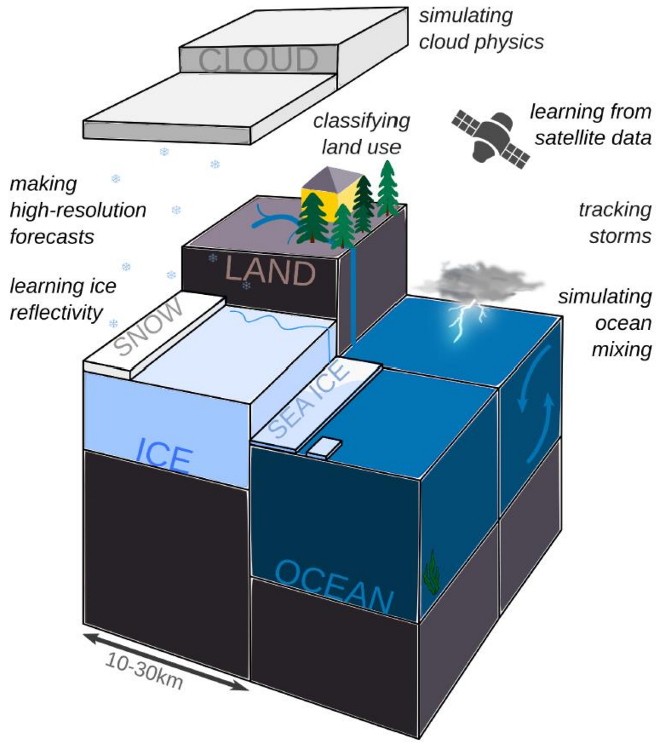
In a 2022 paper by Mamalakis et al.[226], several approaches to the relationship between AI, Meteorology and Climate Science are highlighted. Similar to weather forecasting, climate prediction at subseasonal, seasonal and decadal timescales is among the most important challenges in climate science, with implications for the economy and ecosystem management around the world[227]. The analysis of big climate data is crucial for understanding and predicting climate change; so, AI allows for the extraction of valuable insights from large datasets, such as historical climate records, satellite observations, and climate simulation outputs. By uncovering hidden patterns and correlations, AI algorithms enable scientists to discern climate drivers, understand the feedback mechanisms, and identify potential tipping points.

To this end, data-driven ML methods that leverage information from the entire globe (i.e., beyond predefined climate indices) have been suggested in the literature and they have shown improvements in predictive skill[228]. For example, a NN framework method for viewing climate patterns may be useful for quantifying differences in the timing and patterns of climate change across different climate models, and evaluating models against observational datasets, as well as for detection and attribution of anthropogenic climate change[229].

AI has revolutionised climate modeling and prediction by processing vast amounts of data and generating sophisticated models more accurately and at a higher speed than traditional approaches. ML algorithms have enabled the identification of complex climate patterns, leading to improved predictions of temperature changes, patterns and intensities of weather phenomena, and other variables. As

[226] Mamalakis, A., Ebert-Uphoff, I., Barnes, E. (2022). Explainable Artificial Intelligence in Meteorology and Climate Science: Model Fine-Tuning, Calibrating Trust and Learning New Science. In: *Holzinger, A., Goebel, R., Fong, R., Moon, T., Müller, KR., Samek, W. (eds) xxAI - Beyond Explainable AI.* xxAI 2020. Lecture Notes in Computer Science(), vol 13200. Springer, Cham. https://doi.org/10.1007/978-3-031-04083-2_16
[227] Ocean Studies Board, The National Academies of Sciences, Engineering, and Medicine et al. (2016). Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts. National Academies Press.
[228] DelSole, T., Banerjee, A. (2017) Statistical seasonal prediction based on regularized regression. *J. Clim.* 30(4), 1345–1361 and Stevens, A., et al. (2021) Graph-guided regularized regression of pacific ocean climate variables to increase predictive skill of southwestern us winter precipitation. *J. Clim.* 34(2), 737–754.
[229] Barnes, E.A., Hurrell, J.W., Ebert Uphoff, I., Anderson, C., & Anderson, D.(2019).Viewing forced climate patterns through an AI Lens. *Geophysical Research Letters*, 46, 13,389–13,398, https://doi.org/10.1029/ 2019GL084944.

depicted below in a schematic of a climate model, recent trends have created opportunities for ML to advance the state-of-the-art in climate prediction:



Picture 5

Schematic of a climate model, with selected opportunities to improve climate change predictions using ML

Rolnick D. et al. (2022) *Tackling Climate Change with Machine Learning*, https://doi.org/10.48550/arXiv.1906.05433, 42:36.

AI is well suited to help in forecasting and projecting climate-related hazards, whether in the form of improved long-term projections of regionalised events, such as sea-level rise, or in the form of early warning systems for extreme weather events[230]: by using AI to create large-scale climate models, long-

[230] BCG Gamma - AI for the Planet (2022) *How AI Can Be a Powerful Tool in the Fight Against Climate* Change, p. 10.

term climate projections include extreme events probability, e.g., floods, wildfires, etc. in order to "climate-proof" infrastructure and to prevent loss of human lives and property destruction.

Therefore, AI contributes to enhanced **climate risk assessment**: by processing vast amounts of socio-economic data and by combining them with climate models, ML algorithms are able to quantify climate risks at different spatial and temporal scales. This facilitates the identification of high-risk areas and assists in the development of targeted adaptation strategies, ensuring efficient resource allocation for climate resilience. AI can also provide fast approximations to certain physics simulations within climate and weather models that would otherwise be prohibitively time-intensive to run, which can be used to provide more localised and accurate risk predictions, as in the case of extreme weather events.

Through pattern recognition, neural networks can detect anomalies and provide an objective assessment of whether observed climatic events are attributable to climate change. Ultimately, AI in forecasting and Climate Science can bolster climate change detection and attribution, helping scientists to distinguish between natural climate variations and human-induced changes, and then, to help formulate the exact needed policies for climate action.

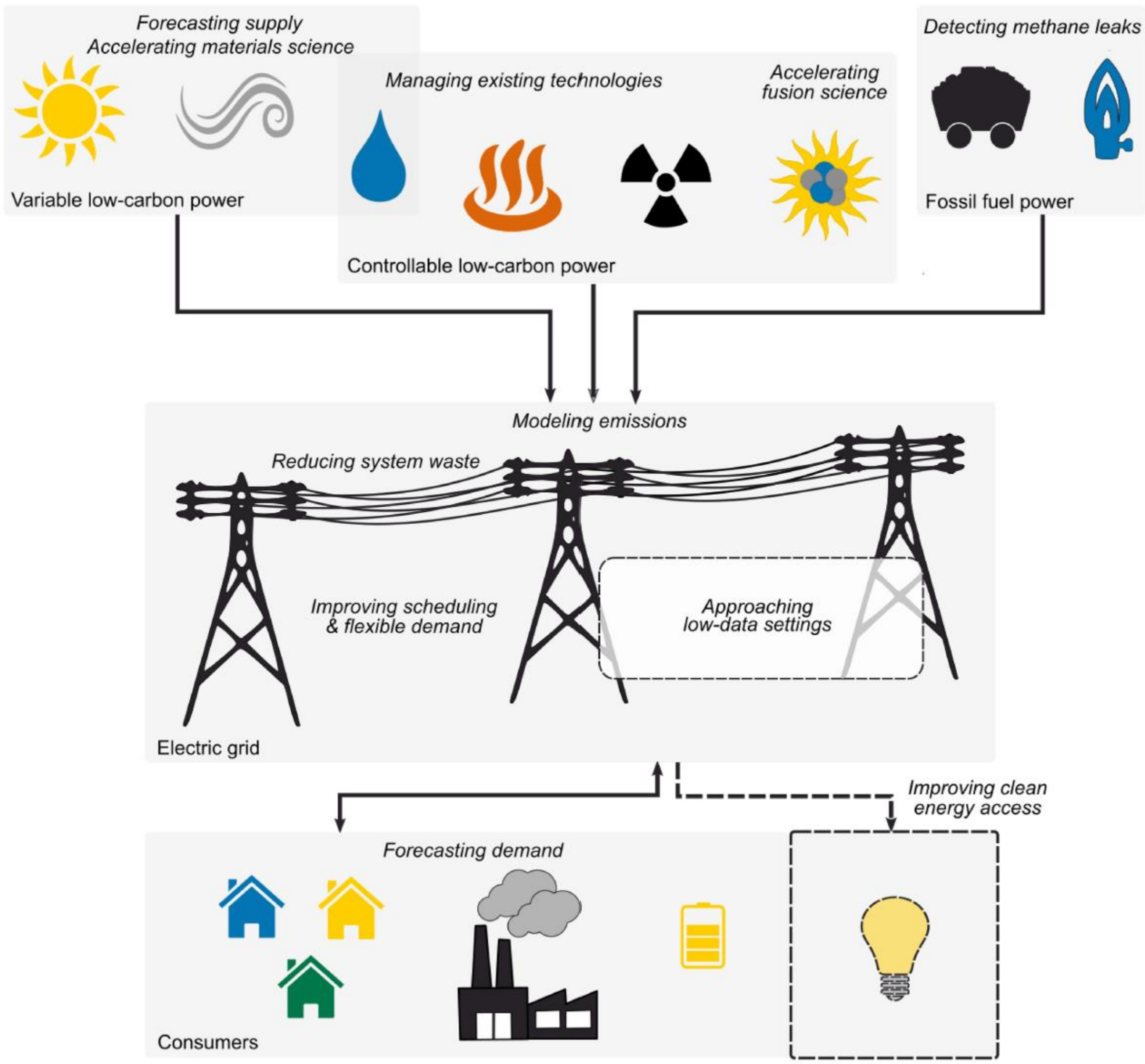
3.3.2. AI in Energy and Smart Cities

3.3.2.1. Energy Systems

Electricity is one of the industries that AI is poised to transform[231]. With the help of AI, it is possible to effectively combat emissions in the energy sector by employing accurate forecasting methods to

[231] Schematic of a climate model, with selected opportunities to improve climate change predictions using ML Rolnick D. et al. (2022) Tackling Climate Change with Machine Learning, https://doi.org/10.48550/arXiv.1906.05433, 42:6.

determine power supply and demand on the grid, enhancing renewable energy scheduling, and reducing life-cycle fossil fuel emissions through proactive maintenance.



Picture 6

Rolnick D. et al. (2022) *Tackling Climate Change with Machine Learning*,

https://doi.org/10.48550/arXiv.1906.05433, 42:7.

To balance power grids efficiently, and thereby enable the integration of large amounts of renewables, it is essential to forecast both electricity supply and demand, a function that AI can provide[232]. AI-powered **smart grids** offer intelligent energy management by optimising the supply and demand of electricity and their advanced algorithms analyse data from sensors to optimise energy distribution; this allows for better load balancing, reduced transmission losses, and improved integration of decentralised energy sources. Hence, when system optimisation and control are enhanced, so are its operation and efficiency; and apart from electricity grids, smart grids can also be used in heating and cooling systems.

AI can also significantly contribute to **energy efficiency** initiatives by enabling precise demand-response mechanisms. With the help of pattern recognition algorithms, energy consumption patterns can be identified and analysed, leading to the development of energy-saving strategies. This predictive capability allows AI systems to proactively adjust energy usage based on anticipated needs; thus, by actively managing energy usage in response to fluctuations in supply and demand, AI can minimise wastage and achieve lower electricity costs. For example, during peak demand periods, AI can automatically optimise heating, ventilation, and air conditioning (HVAC) systems to reduce energy consumption without compromising comfort levels. Similarly, in industrial settings, AI can optimise production schedules to align with periods of lower electricity costs or higher availability of renewable energy sources.

Furthermore, AI can enhance the effectiveness of demand-response mechanisms by predicting and optimising energy consumption patterns in real-time. In a home setting, by using AI-enabled tools (e.g., motion sensors and "smart thermostats"[233]), heating systems can learn the patterns of residents' presence and absence in a home; said data can be collected through various sources such as motion sensors, smart

---

[232] Global Partnership on AI Report in collaboration with Climate Change AI and the Centre for AI & Climate. (2021). *Climate Change and AI: Recommendations for Government Action*, p. 19.

[233] See Ghahramani A, Galicia P, Lehrer D, Varghese Z, Wang Z and Pandit Y. (2020). Artificial Intelligence for Efficient Thermal Comfort Systems: Requirements, Current Applications and Future Directions. *Front. Built Environ.* 6:49. doi: 10.3389/fbuil.2020.00049.

thermostats, or even wearable devices. With this information, AI algorithms can analyse and understand the typical schedule of residents, including their daily routines, work hours, and sleeping patterns and then automatically adjust the temperature settings to conserve energy. Further, by analysing historical data and considering various factors such as weather conditions, time of day, and user behavior, AI algorithms can accurately forecast future energy demands.

Moreover, AI-powered smart grids can enable two-way communication between energy providers and consumers. This bi-directional flow of information allows consumers to actively participate in managing their energy usage. Through smart meters and IoT devices connected to the grid, consumers can receive real-time feedback on their energy consumption habits and make informed decisions about when to use certain appliances or devices, and which ones to use.

Another aspect is that smart infrastructure and asset maintenance combine computer vision and automated Unmanned Aerial Vehicles (UAV or drones) for e.g., distributed energy generation assets (urban wind turbines which may be installed in hazardous locations such as on towers and bridges and integrated photovoltaics (PV) in dwelling facades). In places like energy infrastructure, AI application contributes in damage detection, damage prediction, damage classification, damage localisation, condition assessment and lifetime prediction[234].

Ultimately, ML applications make the grid more resilient for decentral and uncertain generation, enabling the uptake of more renewable energy, while citizens are supported in adopting flexible and efficient energy consumption patterns which benefits the households and the grid[235]. AI has the potential to become a revolutionary power by enabling smarter, more efficient, and sustainable energy systems; from optimising renewable energy generation to improving energy storage and efficiency, AI can bring tremendous benefits to the sector.

---

[234] European Parliament, Policy Department for Economic, Scientific and Quality of Life Policies (2021) *Artificial Intelligence in smart cities and urban mobility.*

[235] European Parliament, Policy Department for Economic, Scientific and Quality of Life Policies (2021) *Artificial Intelligence in smart cities and urban mobility*, p. 3.

However, as AI innovations have the potential to revolutionise the oil and gas industries as well by optimising processes and improving efficiency, it is crucial to consider the unintended consequences that may arise from these advancements; one of which is that AI could potentially lower operational costs for oil and gas companies, at the expense of the environment. By streamlining operations, enhancing extraction techniques and automating certain tasks, AI could potentially contribute to *more* emissions than before its deployment.

To address this challenge effectively, close collaborations between AI researchers and decision-makers in the electricity system are essential. These collaborations can help ensure that AI innovations align with broader sustainability goals and do not inadvertently contribute to increased emissions; by involving regulators and policymakers into innovation and industry, collective work can be made in order to implement safeguards that prevent any unintended negative consequences.

3.3.2.2. Smart Cities

a) The Revolution of Urban Living

The integration of AI with the concept of smart cities has the potential to revolutionise urban living, as it can help create more sustainable and efficient cities. By analysing data from various sources such as traffic, weather, and energy use, AI algorithms can provide insights for optimising city operations and reducing emissions and energy consumption. The philosophy of AI reflects on the imitation of nature in an attempt to find solutions to new problems and new interactions between the biological and the digital[236], which is the epitome of the creation of the "smart city" concept. As a recent report by the World Bank puts it, metropolises act as cradles of brands that produce sustainable, innovative products and can become incubators of green innovation[237].

---

[236] Polisena, V. (2021). *Philosophy of Artificial Intelligence: Ethics for Smart Cities,* AI Ethics.
[237] See the World Bank Report on making Cities Green in the context of climate change on Mukim, M., Roberts, M. (eds). (2023). *Thriving: Making Cities Green, Resilient, and Inclusive in a Changing Climate*. Washington, DC: World Bank. doi:10.1596/978-1-4648-1935-3, p. 279.

According to the definition used by the European Commission, *"a smart city is a place where traditional networks and services are made more efficient with the use of digital solutions for the benefit of its inhabitants and business. A smart city goes beyond the use of digital technologies for better resource use and less emissions. It means smarter urban transport networks, upgraded water supply and waste disposal facilities and more efficient ways to light and heat buildings. It also means a more interactive and responsive city administration, safer public spaces and meeting the needs of an ageing population*[238]".

Smart cities present seven key characteristics, as opposed to the "traditional" city concept[239]:

- **a)** Broad use of ICT: complex computer networks and interconnected Internet services;
- **b)** Pro-business ethos: cooperation and consultation between local government and communities with the industry[240];
- **c)** Openness: the city is seen as an open platform to foster innovation and progress;
- **d)** Real-time monitoring: the use of ICT to monitor the city's functions, for example in transportation networks;
- **e)** Increasing use of technologies and innovations: the city is seen as an innovation hub for the introduction of new technologies;
- **f)** Empowerment: The social aspect of smart cities is centred around the encouragement of citizen education and empowerment for active participation in public affairs[241].

---

[238] European Commission. *Smart cities: Cities using technological solutions to improve the management and efficiency of the urban environment*. Available at: https://commission.europa.eu/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en.

[239] Montes, J. (2020). *A Historical View of Smart Cities: Definitions, Features and Tipping Points*, doi.org/10.2139/ssrn.3637617.

[240] See a thorough analysis on the subject on Marchesani, F., Masciarelli, F., Bikfalvi, A. (2023). Smart city as a hub for talent and innovative companies: Exploring the (dis) advantages of digital technology implementation in cities, *Technological Forecasting and Social Change*, Volume 193, 122636, https://doi.org/10.1016/j.techfore.2023.122636.

[241] See Hollands, R. G. (2008). *Will the real smart city please stand up? Intelligent, progressive or entrepreneurial*? City, 12(3), 303-320. doi:10.1080/13604810802479126

g) Sustainability: Smart Cities foster sustainable tools to provide environmental-friendly solutions to everyday urban life, e.g., smart waste management systems and monitoring of urban emissions.

And as depicted below, Smart Cities are the cradle of the intelligent transformation of five applications, which render their respective sectors “smart”:

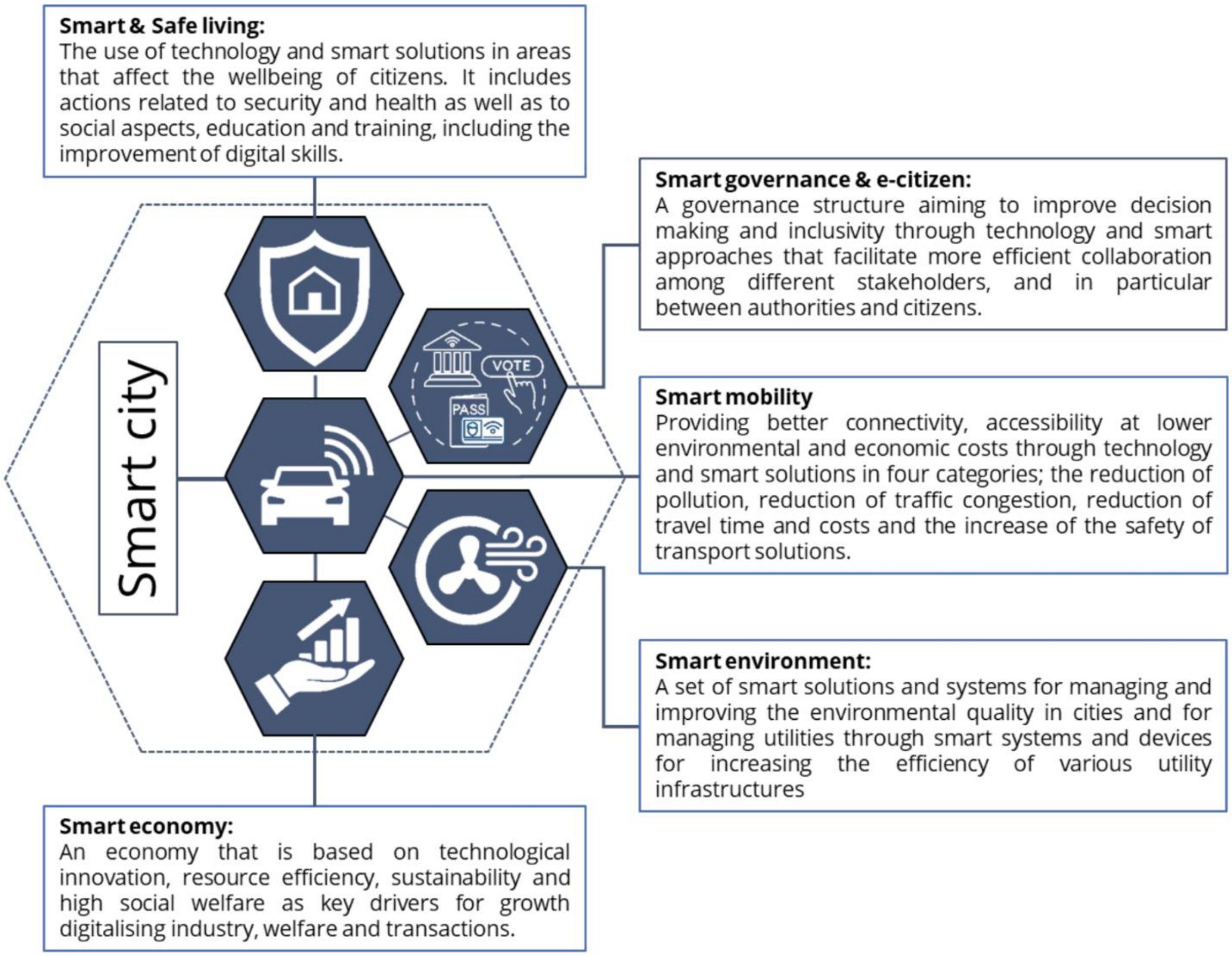
a) smart and safe living,

b) smart governance and e-citizen,

c) smart mobility,

d) smart environment, and

e) smart economy.



Picture 7

European Parliamentary Research Service, Panel for the Future of Science and Technology (2023) *Social approach to the transition to smart cities*, p. 10.

It could be said that Smart Cities encompass the idea of the Twin Transition through urban transformation.

One of the key benefits of AI in Smart Cities is its ability to **enhance urban infrastructure**, as AI-powered systems can analyse large amounts of data generated by sensors and devices throughout the city, allowing for real-time monitoring of critical infrastructure such as bridges, roads, and buildings. This tremendously helps risk management in case of potential accidents, and also enables prompt maintenance and repairs[242].

Furthermore, smart solutions in cities are ideally incorporated into **urban mobility**; AI-powered traffic management systems can reduce congestion and optimise traffic flows by analysing data from various sources, including traffic cameras and GPS devices. This can lead to shorter travel times, decreased pollution levels, and improved overall urban mobility. More on AI in transport will be analysed in the "Transport" section next.

Governance is another area where AI can have a significant impact on Smart Cities[243]. The fast and efficient service of the citizens and visitors of the city is an important part of city life; the digitalisation of applications and services can enhance citizen access to the public domain, improve interactivity and reduce costs and bureaucracy. By using ICT in city governance, it is possible to radically transform existing processes and upgrade the services provided into value-added ones, in order to make urban administration more effective and efficient and to promote the active participation of citizens in public affairs. As AI algorithms can process vast amounts of data from various sources, by utilising data predictive analysis better decision-making processes can be achieved; for example, city officials can gain actionable insights into urban phenomena such as crime rates and housing needs. In 2019, researchers used real data sets on the cities of Chicago and New York, USA, to run an algorithm that led to a spatio-temporal model for urban crime forecasting models[244]. This valuable knowledge that is acquired from digital means and processed by digital models can thus help enhance public safety and facilitate police work. Another example is the use of smart knowledge management and virtual

---

[242] Isern, J. Barranco, F., Deniz, D., Lesonen, J., Hannuksela, J., R. Carrillo, R. (2020). *Reconfigurable Cyber-Physical System for Critical Infrastructure Protection in Smart Cities via Smart Video-Surveillance*, https://doi.org/10.48550/arXiv.2011.14416.

[243] For more insights on public services and AI see Misuraca, G., van Noordt, C. (2020). *Overview of the use and impact of AI in public services in the EU*, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19540-5, doi:10.2760/039619, JRC120399.

[244] Catlett, C., Cesario, E., Talia, D., Vinci, A. (2019). Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments, *Pervasive and Mobile Computing*, Volume 53, https://doi.org/10.1016/j.pmcj.2019.01.003.

assistants to help with citizen-local governance interactions, while the creation of digital platforms of citizen participation ("Citizen Portals") and intra-city governance can lead to the establishment of an E-Governance model[245] and the creation of an E-Citizen; this could, potentially, improve city life and build citizens' trust towards their municipal authorities[246]. Moreover, proper allocation of resources, like tailored subsidy provision[247], can be achieved through "smarter" policing.

The so-called "**smart environment**" can also be achieved though digital solutions in cities. For example, applications of AI are for improved catchment area management on e.g., **water** quality and the prevention of flooding and contamination, water efficiency on the demand side, draught forecasting and planning, and adequate sanitation[248]. The management of **waste** and the prevention of leakages and waste contamination of soil and water and air also entail the utilisation of AI systems: by using "smart bins" the waste containers can easily be managed and monitored; afterwards, digital sensors can significantly cut waste processing times through separation and allocation of waste. Digital systems can also contribute to maintaining good **air quality** by using sensors to measure and monitor urban emissions, as well as **noise** levels.

Another significant advantage of AI in smart cities lies in **optimising resource management**. By leveraging AI algorithms cities can efficiently manage crucial resources like agricultural products, energy and water to cover the city's needs. AI can help analyze consumption patterns, identify areas for conservation and enable predictive maintenance, leading to more effective and sustainable use of resources. This has the potential to reduce costs, enhance environmental sustainability, and improve the overall efficiency of urban systems and thus, citizens' quality of life.

---

[245] Kumar, V. T. M. (ed.) (2014). *Governance for Smart Cities, Advances in 21st Century Human Settlements*, Springer Singapore, https://doi.org/10.1007/978-981-287-287-6.
[246] Kuzior, A., Pakhnenko, O., Tiutiunyk, I., & Lyeonov, S. (2023). E-Governance in Smart Cities: Global Trends and Key Enablers. *Smart Cities*, *6(4)*, 1663–1689. https://doi.org/10.3390/smartcities6040078
[247] European Parliament, Policy Department for Economic, Scientific and Quality of Life Policies (2021) *Artificial Intelligence in smart cities and urban mobility*, p. 2.
[248] *Id*., p. 4.

b) Smart Cities Initiatives in the EU

In the EU sphere, while the bulk of existing legislation covers digital aspects[249], certain multiple relevant programmes and initiatives have also been launched and implemented in recent years towards the gradual evolution of the traditional city model to the "digital and sustainable" city. Some include:

- **The EU Mission "Climate-Neutral and Smart Cities**": The so-called "Cities Mission" is part of the Horizon Europe 2021-2027 Programme, putting the city at the heart of energy transformation and urban reform and on the path towards climate neutrality and ecological transition, by having the selected cities draft a "Climate City Contract" (CCC). The CCC are political commitment agreements between municipalities and the European Commission and include: a) an Action Plan of existing and planned climate and digital actions and b) an Investment Plan, as a systematic map of the costs and funds needed by the city to achieve climate neutrality. As regards Greek Cities, in April 2022 the European Commission selected the city of Ioannina, Athens, Thessaloniki, Kalamata, Kozani and Trikala to participate in the mission, whose CCC is to be submitted by September 2023.
- **URBACT** is a European exchange and learning programme promoting urban development and the exchange of knowledge among cities for "sustainable change[250]".
- **The Intelligent Cities Challenge** (ICC) is a European Commission initiative that supports 136 cities in using cutting-edge technologies to lead the intelligent, green and socially responsible recovery through "Local Green Deals[251]".
- **The Smart Cities Marketplace**: it was created by merging two former platforms, the "Marketplace of the European Innovation Partnership on Smart Cities and Communities (EIP-

[249] See European Parliamentary Research Service, Panel for the Future of Science and Technology (2023). *Social approach to the transition to smart cities*, p. 41 and Annex VI for relevant EU legislation.
[250] See more at: https://urbact.eu/.
[251] See more at: https://www.intelligentcitieschallenge.eu/.

SCC Marketplace)" and the "Smart Cities Information System (SCIS)". It is a major market-changing undertaking that aims to bring cities, industries, SMEs, investors, banks, researchers and many other smart city actors together towards a "just and clean urban transition"[252].

c) Legal and ethical challenges in Smart Cities

Despite the numerous advantages that were presented above, AI in smart cities also presents legal challenges and ethical concerns.

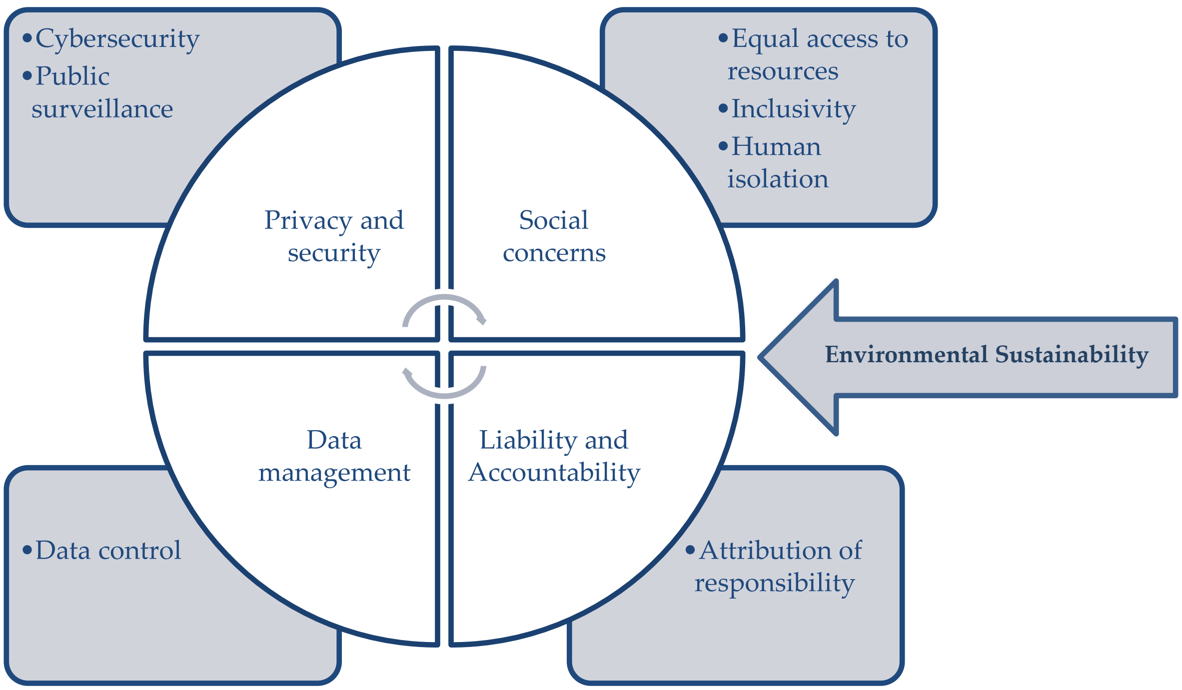


Figure 15

Created by the author.

[252] See more at: https://smart-cities-marketplace.ec.europa.eu/.

Privacy and Security

In Smart Cities, the sensor networks and the surveillance systems deployed at large in public spaces raise questions regarding the proper maintenance of **civil liberties**, e.g., freedom to assemble (in case of public protests) and freedom of speech. Hence, careful oversight of the systems is needed, so as to ensure that a balance is kept between state supervision and the protection of personal and collective freedoms.

Furthermore, as smart cities become increasing dependent on interconnected "smart" networks, they also become move vulnerable to **cyber threats and cybercrime**, e.g., damage on critical infrastructure through cyber-attacks, ransomware cases etc[253]. To this end, maintaining a strong digital protection system that is able to prevent and deal with potential attacks should be one of the top priorities of the governance agenda in smart cities.

Data management

One of the foremost challenges smart cities face when utilising citizens' **data** to optimise urban systems involves the collection, storage, and sharing of it. Managing citizens' consent through the management of their personal data and implementing robust data protection regulations through the city's processes, while fostering transparency in said procedures, are crucial; to this end, ensuring that citizens' personal information is protected and unauthorized access to sensitive data is prevented, a clear data protection policy throughout the city's development is absolute necessary.

Liability and Accountability

When smart city technologies result in negative consequences, such as accidents involving autonomous vehicles and cybercrimes, determining accountability and liability can be challenging. Developing legal frameworks that allocate responsibility to the appropriate parties while fostering innovation is essential for building trust among the city, citizens and businesses. More on the legal frameworks that could be established in order to not hinder innovation will be analysed in the "Regulatory Sandboxes" section.

---

[253] On more on ransomware, see Beaman, C. et al. (2021) Ransomware: Recent advances, analysis, challenges and future research directions, *Computers & Security*, Volume 111, 102490, https://doi.org/10.1016/j.cose.2021.102490.

Social concerns

Apart from the aforementioned techno-legal issues, there are ethical concerns that should be taken into account as well. Smart cities, and without prejudice to their tremendous positive impact on society, have the potential to exacerbate existing social inequalities if not carefully designed and implemented. As progress demands financial support, it is usually realised in the more privileged areas of a city; so, without **inclusive** policies, the more marginalised parts of cities could be left behind, widening the social divide.

Furthermore, and as for **accessibility** considerations, cities must ensure equal access to digital advancement and relevant opportunities for all citizens, irrespective of socioeconomic backgrounds. Issues like digital illiteracy should be dealt with; minorities such as immigrants, the elderly and people with disabilities should be given equal chances to participate in digital skills workshops, as well as in decision-making processes and open dialogue with the city officials, so as diversity of opinions is ensured.

As analysed in the previous sections of the present thesis, AI can be unfair, discriminatory and biased against, for example, ethnic minorities. A proper smart city should also develop **ethical guidelines and regulations** for the creation of a transparent, explainable, fair and ethical urban AI ecosystem, as well as implement diversity policies.

Last but not least, a problem that arose during the COVID-19 pandemic was **human isolation** due to remote work and indoor activities. When a city multiplies its digital features, education, work, entertainment and activities that were once physical can turn to digital or, in any case, indoor ones. This progressively bigger loss of human interaction would be a lot more evident in smart cities; so, it deserves careful consideration in the city's transitional policies.

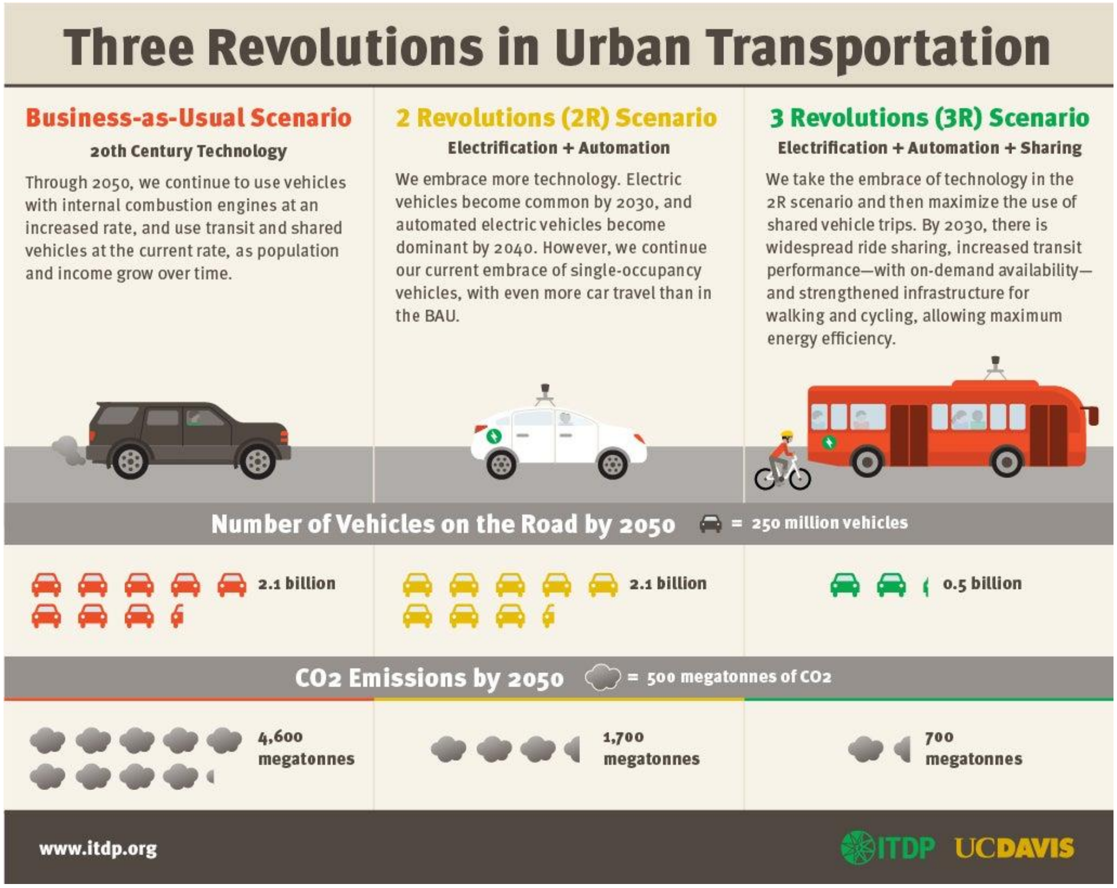
Environmental Sustainability

Without question, **the very core of the idea of "smart cities" includes the environmental sustainability aspect.** So, it is inevitable that when the ICT sector thrives, its ecological footprint, for example in the form of disposed ICT equipment and high energy consumption, also becomes intensified. A dilemma is then identified; while technology can enhance resource efficiency and reduce carbon footprints, it can also contribute to electronic waste accumulation and energy consumption. In order to not reach an impasse, a balance is needed between technological advancement and ecological sustainability; for example, through the adoption of green technologies, the constant evaluation of the environmental impact of the ICT sector and the thoughtful planning of urban interventions by licencing authorities.

In conclusion, AI and the digitalisation of society through the smart cities concept has the potential to significantly enhance human life. From optimising urban infrastructure and resource management to promoting sustainability and improving governance, AI-driven systems offer immense possibilities for urban development; however, the challenges are numerous. By addressing these challenges head-on and implementing thoughtful and ethical governance frameworks, cities can harness the transformative potential of technology while ensuring the well-being of their citizens.

3.3.3. AI in Transport

AI has emerged as a transformative technology that holds immense potential in revolutionising various sectors of the economy, including transportation. By augmenting traditional transportation systems with sophisticated AI capabilities, researchers and professionals are making strides in enhancing efficiency, safety, and sustainability. As depicted below, Urban Transportation, with the help of AI, can evolve into an electrification-automation-sharing transport sector and significantly cut emissions up to 700 megatonnes and the number of vehicles worldwide three times over, by 2050.



Picture 8

Kerlin, K. (2017) *You Say You Want a Transportation Revolution? How About Three of Them?,* at: https://www.ucdavis.edu/news/you-say-you-want-transportation-revolution-how-about-three-them-0

3.3.3.1. *Automated Vehicles (AVs)*

In order to transform the transport sector, automation plays a significant role; car manufacturers around the world are using AI in almost every aspect of the car manufacturing process and in the way they operate on the road. More specifically, AI is being applied to the automotive industry, including design, supply chain, production and post-production, but it is also being applied to “driver assistance” and “driver risk assessment” systems, and thus transforming the transport sector. Aftermarket services such as vehicle predictive maintenance and insurance are also being transformed through the use of AI.

In the EU, the words “automated” and “autonomous” can be used together. For example, Regulation (EU) 2019/2144[254] defines “automated vehicle” and “fully automated vehicle” based on their autonomous capability:

1. “Automated vehicle”: means a motor vehicle designed and constructed to move autonomously for certain periods of time without continuous driver supervision but in respect of which driver intervention is still expected or required;
2. “Fully automated vehicle” means a motor vehicle that has been designed and constructed to move autonomously without any driver supervision;

Many major car manufacturers are working to create their own autonomous cars and driving features. The AI system must be able to identify any type of external influence (infrastructure, weather, roadworks, other vehicles, pedestrians or any other object) that could be considered as an obstacle to a car and assess its potential effect on the vehicle’s trajectory in order to make appropriate corrections to the driver control system in real time.

An AI-powered vehicle uses an Advanced Driver Assistance Systems (ADAS), some types of which are:

- Automatic braking,
- driver drowsiness detection,
- lane departure warning,

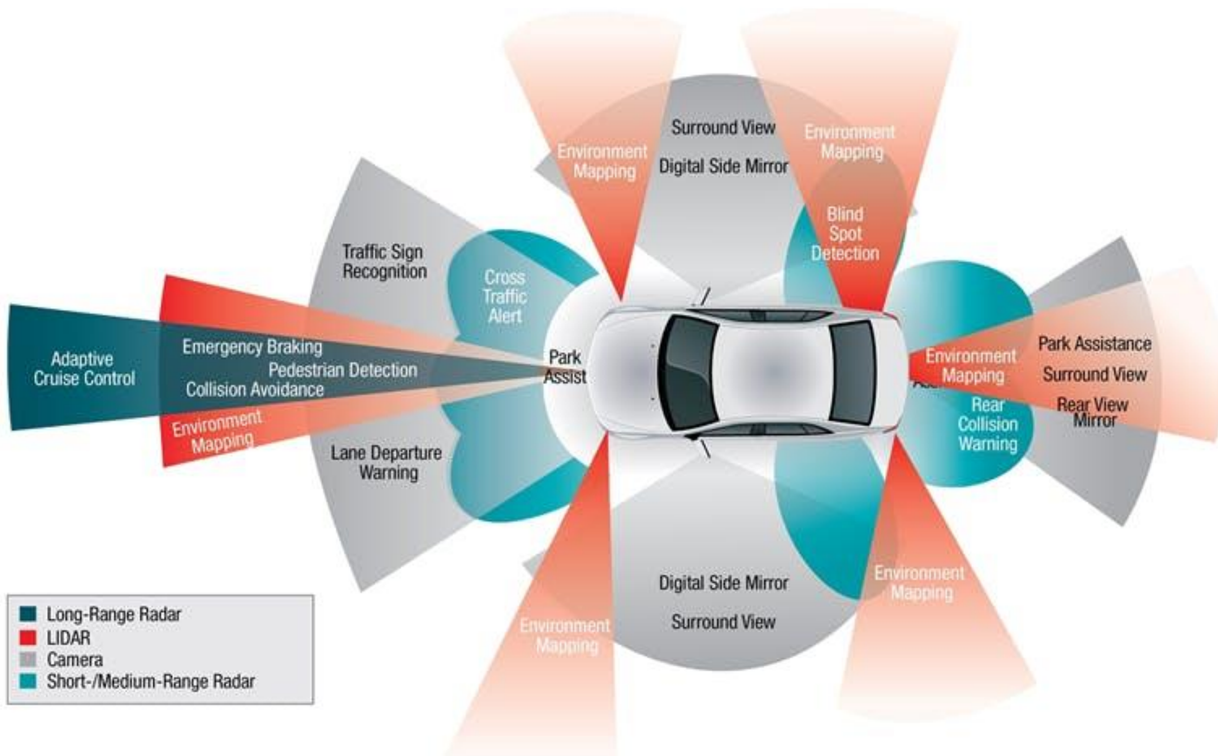
[254] REGULATION (EU) 2019/2144 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users, amending Regulation (EU) 2018/858 of the European Parliament and of the Council and repealing Regulations (EC) No 78/2009, (EC) No 79/2009 and (EC) No 661/2009 of the European Parliament and of the Council and Commission Regulations (EC) No 631/2009, (EU) No 406/2010, (EU) No 672/2010, (EU) No 1003/2010, (EU) No 1005/2010, (EU) No 1008/2010, (EU) No 1009/2010, (EU) No 19/2011, (EU) No 109/2011, (EU) No 458/2011, (EU) No 65/2012, (EU) No 130/2012, (EU) No 347/2012, (EU) No 351/2012, (EU) No 1230/2012 and (EU) 2015/166

- blind spot information for passing vehicles,

- warning of road speed limits,

- keeping the car in the lane, and

- maintaining a constant safe distance from the vehicle ahead.



Picture 9

An example of the systems of an AV

Wendt Z, Cook, J. (2018) *Saved by the Sensor: Vehicle Awareness in the Self-Driving Age*, at: https://www.machinedesign.com/mechanical-motion-systems/article/21836344/saved-by-the-sensor-vehicle-awareness-in-the-selfdriving-age

As AVs utilise advanced perception, decision-making, and learning algorithms to navigate roads without human intervention, they can accurately perceive their surroundings, predict traffic patterns, and react in real-time, offering a more efficient and safer mode of transportation. Thus, they can contribute to the reduction of traffic congestion and induce efficiency through the so-called "ecodriving" and platooning (that is, a "platoon" of connected AVs - or CAVs - is defined as a group of CAVs that exchange information, so that they can drive in a coordinated way, allowing very small spacings and, still, travelling safely at relatively high speeds[255]). Through AI-powered algorithms, AVs can communicate with each other and coordinate their movements efficiently; this coordination can lead to smoother traffic patterns, reduced travel times, and improved overall efficiency in transportation networks.

a) Environmental impact

As per their **environmental impact**, if we assume that the AVs will be electric, then they will positively contribute to reducing GHG emissions. However, if AVs are powered like conventional vehicles, or, in any case, they are not 100% electric, then some problems around their environmental impact arise. It is also possible that AVs will actually lead to an increase in overall energy consumption by facilitating individualised transportation and the increase of total distance covered due to the more comfortable nature of "having the car drive itself".

For example, AVs allow people to send their cars on "zero occupancy" trips: instead of paying for parking - for example - one could decide to send the car back home while at work, only to call it in at the end of the day. While this is certainly convenient, it is not necessarily sustainable from an environmental point of view[256]. Another point that researchers often identify is that since sitting in a car as a passenger is much less stressful than driving, people may be willing to take longer trips. A 2019 survey conducted in California[257], for example, reported that 21%-35% of AV owners take more long-distance trips - especially on weekends - in part because of automated operating systems.

---

[255] See Martínez-Díaz, M., Al-Haddad, C., Soriguera, F., Antoniou, C., (2021) Platooning of connected automated vehicles on freeways: a bird's eye view, *Transportation Research Procedia*, Volume 58, 2021, p. 479-486, ISSN 2352-1465, https://doi.org/10.1016/j.trpro.2021.11.064.

[256] Gailhofer, P. et al., The role of Artificial Intelligence in the European Green Deal, Study for the special committee on Artificial Intelligence in a Digital Age (AIDA), Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg, 2021.

[257] Hardman, S., Chakraborty, D., & Kohn, E. (2021). A Quantitative Investigation into the Impact of Partially Automated Vehicles on Vehicle Miles Travelled in California. UC Office of the President: University of California Institute of Transportation Studies. http://dx.doi.org/10.7922/G2XK8CVB

A survey of people who drive partially AVs - where autopilots can assist with driving tasks even although they don't take over the driving fully - found that over the course of a year, owners of such cars drove an average of nearly 5.000 miles (about 8,000 kilometers) more than those who had conventional vehicles, admitting that they were more willing to engage in traffic and make more long-distance trips because of the increased comfort and reduced stress that these cars allow. Similarly, a study conducted the same year in Sacramento, California, found that households that had access to chauffeur service - which the researchers compared to access to a driverless vehicle - traveled nearly 60% more than usual. They also stopped or reduced their public transit, ride-hailing, cycling and walking trips - which fell by 70%, 55%, 38% and 10% respectively - in favour of more car use[258] .

In addition, AVs may offer new Mobility as a Service (MaaS) services[259], which are likely to increase vehicle mileage because they will lead to additional empty trips for pick-up and drop-off compared to personal driving. AVs may also generate transport demand from currently underserved groups (e.g., older people and groups with limited driving). Thus, additional vehicle kilometres will be associated with higher emissions. Finally, the contribution of AVs to urban sprawl, i.e., the expansion of cities increasingly towards the suburbs, with various environmental impacts (like land clearance and deforestation) has been highlighted, as users will no longer be bothered by the longer distance to the city centre as they will not have to drive.

In this context, ensuring the **electrification** of the AV industry is of high importance. Electric vehicles are essential for decarbonising transport, provided that countries switch to renewable energy for electricity generation. AI can also optimise vehicle charging during periods of high electricity supply and enable the use of car batteries as an energy storage option for the grid, while also improving battery

---

[258] Harb M, Malik J, Circella G, Walker J. Glimpse of the Future: Simulating Life with Personally Owned Autonomous Vehicles and Their Implications on Travel Behaviors. *Transportation Research Record*. 2022;2676(3):492-506. doi:10.1177/03611981211052543

[259] Mobility as a Service (MaaS) is an innovative mobility service that aims to redesign the future of urban mobility by integrating multi-modal transportation and app-based technologies to enable seamless urban mobility, see Mitropoulos, L., Kortsari, A., Mizaras, V., & Ayfantopoulou, G. (2023). Mobility as a Service (MaaS) Planning and Implementation: Challenges and Lessons Learned. *Future Transportation*, 3(2), 498–518. https://doi.org/10.3390/futuretransp3020029.

and charging management in transport electrification[260]. ML has also been found more powerful than other methods to improve real-time power management in hybrid vehicles[261].

However, each fully automated car would generate about 4000 GB per day[262]; 1.7 million autonomous cars (0.2% of the global car fleet) would generate the same amount of data as the current global internet traffic[263]. The network technology and data processing infrastructure behind these volumes of data would require large amounts of electricity and produce life-cycle emissions. This is also an important issue that will need to be addressed.

b) AVs: Legal and ethical issues

Of course, the use of AVs incurs some legal and ethical concerns as well.

Accountability: As mentioned before in this thesis, the issue of attributing responsibility in case of malfunction, accidents and even life-threating situations directly or indirectly caused by AI is highly controversial and complicated; hence, it should be one of the top priorities of the industry. According to a 2021 survey, it was concluded that (future) legislations should limit the liability of passengers – and thus transfer it to the manufacturer or the AI's software engineer – or else people will not accept AVs[264]. Another study in the US[265] aimed to understand the impact of three factors on public acceptance of AVs, one of which was liability:

---

[260] See Rolnick, D. et al. (2019) *Tackling Climate Change with Machine Learning*, at: https://arxiv.org/abs/1906.05433.
[261] Ali, A.M., Söffker D., *Towards optimal power management of hybrid electric vehicles in real-time: A review on methods, challenges, and state-of-the-art solutions*. Energies, 11(3), 2018.
[262] Miller, R., *Autonomous cars could drive a deluge of data center demand*, 2017. Available at: https://datacenterfrontier.com/autonomous-cars-could-drive-a-deluge-of-data-center-demand/.
[263] Liu, R. et al., *Impacts of the digital transformation on the environment and sustainability*, Öko-Institut, issue paper under Task 3 from the "Service Contract on future EU environment policy" 2019. Available at: https://ec.europa.eu/environment/enveco/resource_efficiency/pdf/studies/issue_paper_digital_transformation_20191220_final.pdf.
[264] Othman, K. (2021). Public acceptance and perception of autonomous vehicles: a comprehensive review. *AI Ethics 1*, 355–387, https://doi.org/10.1007/s43681-021-00041-8.
[265] Martínez-Díaz, M., Soriguera, F. (2018). Autonomous vehicles: theoretical and practical challenges. Transp. Res. Procedia 33, 275–282. https://doi.org/10.1016/j.trpro.2018.10.103

- Respondents were asked if they felt comfortable sending their cars to a job, knowing that they would be liable themselves if an accident occurred: 72% of respondents said no, 10% said yes, while 18% were neutral.

- Respondents were asked "By law, if a car's autonomous system fails, the car is required to alert the driver and either give the driver control or stop. I feel safe knowing that this is required by law". The majority of respondents answered positively knowing that the car will alert the driver and stop if the autonomous system fails.

- Respondents were asked "how do the laws regarding autonomous cars affect your desire to buy one?" 10% of respondents believed that the laws would not affect their desire to buy an AV, 50% were neutral and 40% believed that the laws would affect their desire to buy an AV.

The aforementioned results show how liability can greatly affect public attitudes towards AVs and that a large percentage of respondents may elect not to drive an AV for legal/liability-related reasons. Returning to the AI theoretics, explainability has been mentioned as an inevitable first step towards a holistic accountability framework; by installing a data recorder or "black box" similar to the ones used in planes is often mentioned as a means to document inherent processes that led to a system failure[266].

Unless clear legal frameworks regarding liability are adopted, the issue of liability will continue to be a grey area for the industry, which will definitely hinder the progress of the mass influx of AVs in modern societies.

Trust and control: By using an AV, users actively relinquish control over many decisions - not only in driving *per se*, but also regarding their health and the safety of other traffic participants, like pedestrians and other drivers. Two surveys showed high levels of safety concerns, with 90.9%[267] and 54% of respondents concerned about vehicle safety respectively[268].

---

[266] Boch, A., Hohma, E., Trauth, R. (2022). Towards an Accountability Framework for AI: Ethical and Legal Considerations, Technical University of Munich Munich Center for Technology in Society, Institute for Ethics in Artificial Intelligence, at: https://ieai.sot.tum.de/wp-content/uploads/2022/03/ResearchBrief_March_Boch_Hohma_Trauth_FINAL_V2.pdf, p. 7.
[267] Rezaei, A., Caulfeld, B. (2020) Examining public acceptance of autonomous mobility. *Trav. Behav. Soc*. 21, 235–246, https://doi.org/10.1016/j.tbs.2020.07.002.
[268] Piao, J., McDonald, M., Hounsell, N., Graindorge, M., Graindorge, T., Malhene, N. (2016) Public views towards implementation of automated vehicles in urban areas. *Transp. Res. Procedia 14*, 2168–2177.

Privacy, cybersecurity and data protection: as technology for improved user experience (e.g., efficient routing and mapping) often requires personal/sensitive data from the user (e.g., location, navigation patterns), data protection and privacy issues arise. AVs will be able to collect and use a lot of data; at a minimum, they will need data on surroundings such as neighborhoods and other vehicles, users, their locations, habits, and possibly their status (e.g., whether a drunk driver is trying to control the vehicle, whether they need medical attention, etc.). A regulatory framework would guide decisions on what data are really necessary for the optimal functioning of AVs, where they can be processed, who should have access to them (the manufacturer, the Ministry of Transport, the police and other third parties) and under what conditions. Of course, proper safety nets should also be installed in order to prevent cyberattacks and hacking of AVs' control.

Ethical concerns: In the case of an accident, and as shown in the Moral Machine dilemma[269], various prejudices may arise (like choosing to hurt a homeless person crossing the street as opposed to a businessman); the AI's developers should ensure that proper and universally accepted ethical standards are put to place that exclude biases from the AV's operation. Furthermore, the adoption of AI-driven transportation systems may also lead to job displacement, requiring policy interventions and retraining programs for affected individuals.

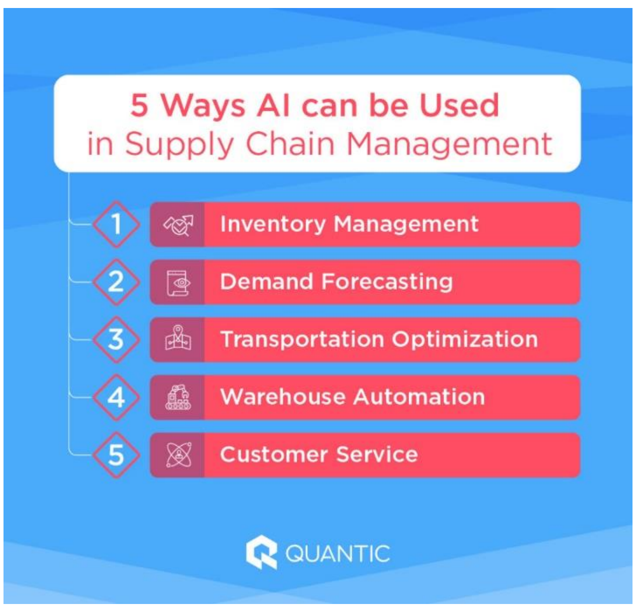
3.3.3.2. Transportation systems

Although AI is used in the automotive industry, it is also being rapidly applied in various fields of transportation, aiming to revolutionise the sector as we know it. Passenger and freight transportation are each responsible for about half of transport GHG emissions[270], so sustainable policies aim to

[269] *Supra* 137 and 138.
[270] R. Schaeffer, R. Sims, J. Corfee-Morlot, F. Creutzig, X. Cruz-Nunez, D. Dimitriu, and M. D'Agosto (2014). *Transport, in IPCC, Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Climate Change 2014: Mitigation of Climate Change, Chapter 8*. Geneva. O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth,

effectively decarbonise the sector by enhancing efficiency and reducing activity. For example, AI methods can provide predictions about patterns in mobility sectors such as aviation or train and naval shipping.

More specifically, **real-time traffic management** that is enabled by AI systems supports decision-making processes on various traffic control strategies and enhances the performance of the overall network[271]. Smart Traffic Management Systems are defined as technology solutions that municipalities can integrate into their traffic cabinets and intersections today for fast, cost-effective improvements in safety and traffic flow on their city streets[272]. These systems can dynamically adjust traffic lights (especially when using "smart traffic lights"), speed limits and on-ramp signalling in order to optimise traffic flow, reduce travel times and allow for the monitoring of the entire city's transportation network. By leveraging real-time data from sensors, smartphones, and connected vehicles, AI algorithms generate optimised traffic signal timings, rerouting suggestions, adaptive speed limits and even emergency routing for ambulances and police cars.



AI-powered transportation systems also have the potential to optimise **logistics** and **supply chain**

Picture 10

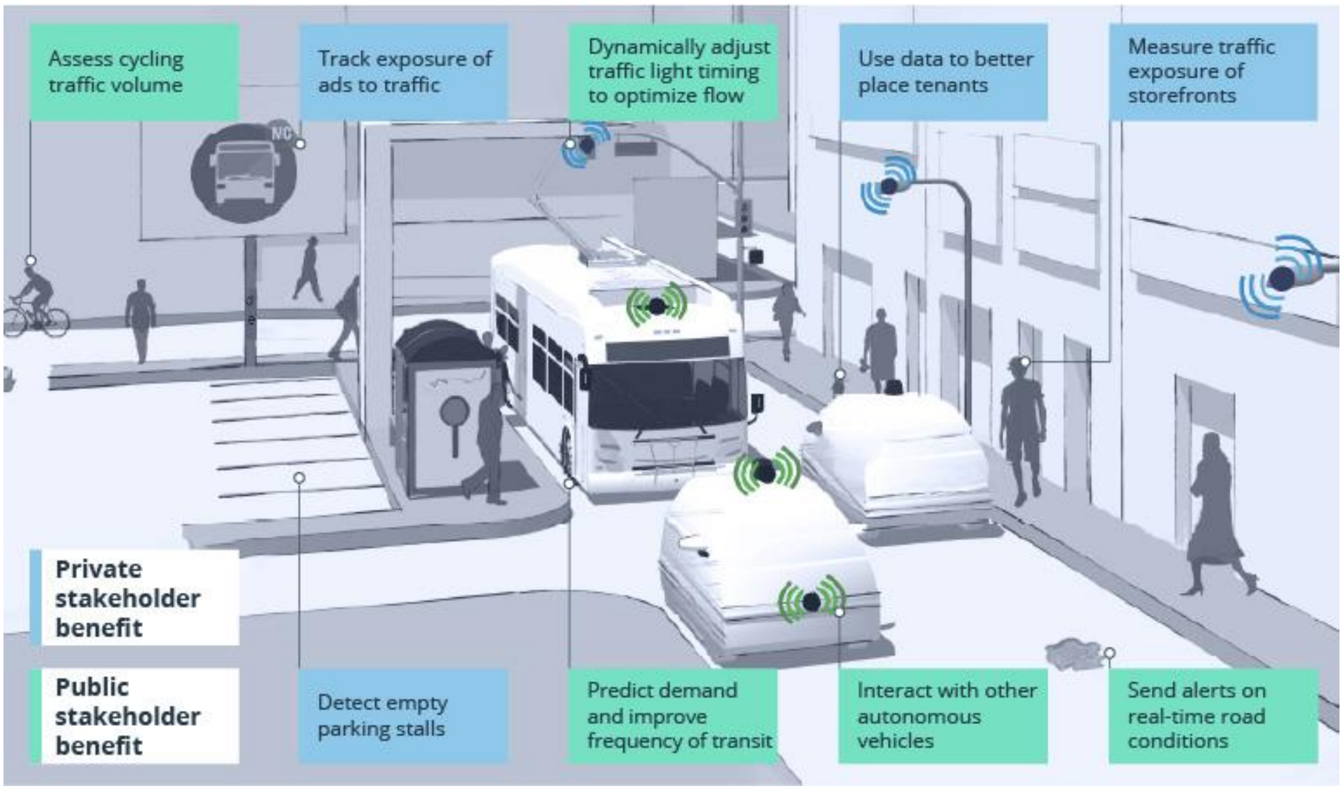
The Important Role of AI in Supply Chain Management & Logistics, at: https://quantic.edu/blog/2023/04/10/the-important-role-of-ai-in-supply-chain-management-logistics/

A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, and J. C. Minx (Eds.). Cambridge University Press, Cambridge

[271] Wismans, L., de Romph, E., Friso, K., Zantema, K. (2014). Real Time Traffic Models, Decision Support for Traffic Management, *Procedia Environmental Sciences*, Volume 22, 220-235, https://doi.org/10.1016/j.proenv.2014.11.022.

[272] Mazur, S. (2020). *Smart Traffic Management: Optimizing Your City's Infrastructure Spend,* at: https://www.digi.com/blog/post/smart-traffic-management-optimizing-spend.

management[273]. By leveraging AI algorithms, companies can better predict demand patterns, optimise routes for delivery vehicles, and reduce fuel consumption. By looking at the supply chain, AI-based automation can assist in the timely retrieval of an item from a warehouse and ensure a smooth journey to the customer, while also ensuring effective warehouse management and efficient fleet planning[274].



Picture 11

AI In Urban Mobility: an example

(2021) *AI in Urban Mobility: Give Citizens What They Want or Die Trying*, at: https://intellias.com/ai-in-urban-mobility-give-citizens-what-they-want-or-die-trying/

[273] See Min, H. (2010) Artificial intelligence in supply chain management: theory and applications, *International Journal of Logistics Research and Applications*, 13:1, 13-39, DOI: 10.1080/13675560902736537

[274] Jacobs, T. (2023) *Unlocking the Value of Artificial Intelligence (AI) in Supply Chains and* Logistics at: https://throughput.world/blog/ai-in-supply-chain-and-logistics/.

Last but not least, **shared mobility** can be largely enhanced by AI, in order to extensively optimise pick-ups and routes. Services like car-, bike- and ride-sharing can lead to higher utilisation of each vehicle, which means a more efficient use of materials[275]. The use of newer, more efficient vehicles and ideally electric ones, could increase with vehicle-sharing concepts and thus lead to reduction of GHG emissions[276]. By encouraging individuals to share vehicles or use alternative modes of transportation like bike-sharing, shared mobility reduces the number of cars on the road, thereby decreasing environmental impacts and alleviating congestion. Furthermore, the widespread adoption of shared mobility has the potential to transform the urban landscape by allocating more space for pedestrians, cyclists, and green areas, transforming streets into vibrant, livable spaces.

AI holds the potential to reshape the way we move, making transportation more efficient, environmentally friendly, and secure. Despite existing challenges, proactive measures and collaboration between governments, academia, and private enterprises can foster responsible AI implementation, unlocking its immense benefits and propelling society towards a smarter and more interconnected future.

3.3.4. *AI in Agriculture and Food Production*

Anthropogenic climate change is caused by multiple climate pollutants, with CO2, CH4, and N2O being emitted by agriculture and food production[277]; at the same time, changes in climate can jeopardise the viability of crops, cause shifts in precipitation and increase frequency of catastrophic weather events that

[275] See G. Hertwich, E. et al. (2019). Material efficiency strategies to reducing greenhouse gas emissions associated with buildings, vehicles, and electronics—a review. *Environmental Research Letters* 14, 4 (2019), 043004.
[276] Rolnick, D. et al. (2019) Tackling Climate Change with Machine Learning, https://arxiv.org/abs/1906.05433, 42:15.
[277] See more on Lynch, J., Cain, M., Frame, D. and Pierrehumbert, R. (2021) Agriculture's Contribution to Climate Change and Role in Mitigation is distinct from predominantly Fossil CO2-Emitting Sectors. *Front. Sustain. Food Syst.* 4:518039. doi: 10.3389/fsufs.2020.518039

affect crops and livestock health. To this end, the growing unpredictability of weather patterns makes agricultural planning and resource management increasingly challenging. According to one estimate, about a third of GHG emissions reductions could come from better land management and agriculture[278]. So, agriculture, as one of society's fundamental sectors, has begun to harness the potential of AI for improved productivity and sustainable practices.

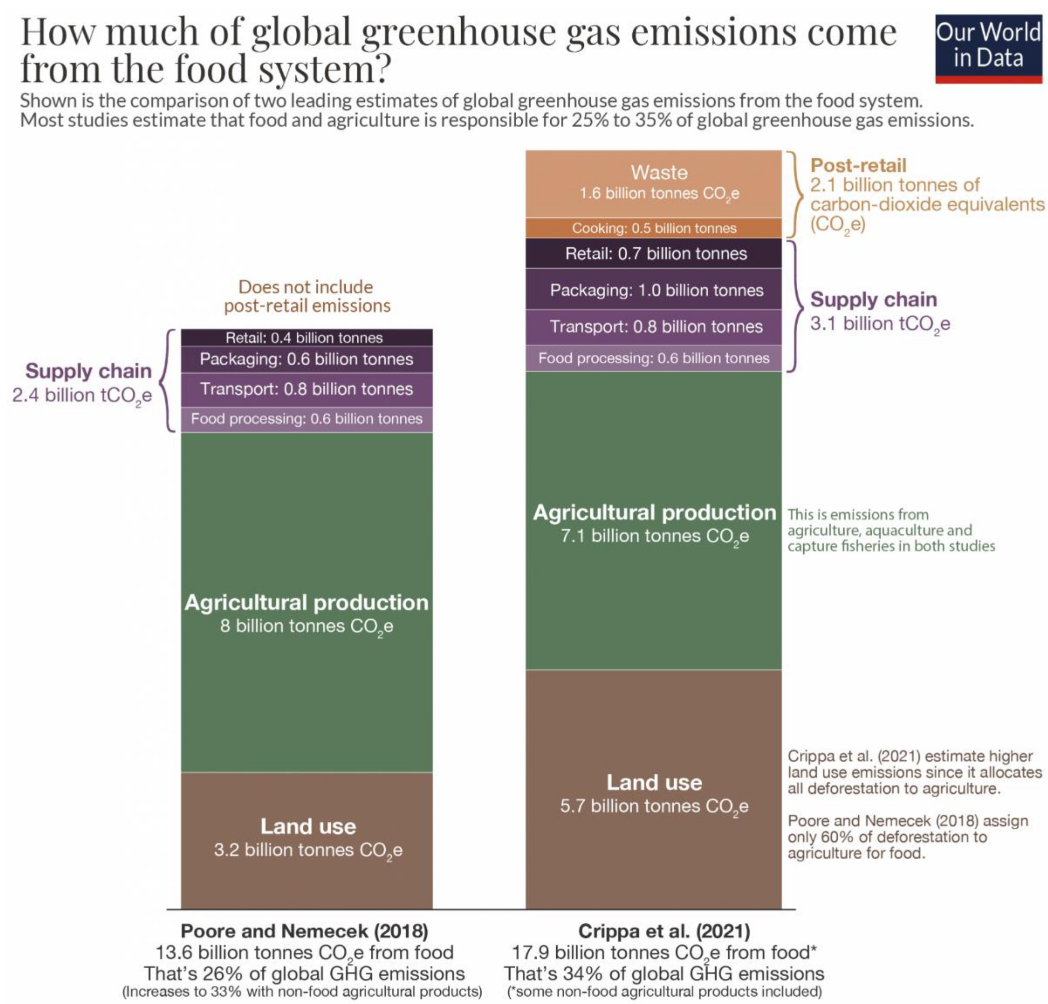


Figure 16

Emissions of the food system sector

by OurWorldInData, at: https://ourworldindata.org/greenhouse-gas-emissions-food

[278] Paul Hawken. (2015). Drawdown: The *Most Comprehensive Plan Ever Proposed to Reverse Global Warming*. Penguin Books.

Perhaps the most important use of AI in sustainable agriculture practices is through "**Precision Agriculture**", which is used as an umbrella term for AI applications in crop management, pest and disease management and optimisation of farming practices. More specifically:

Agricultural mapping, geomapping and field scouting: By using GIS (Geographic Information Systems) multi-layered interactive maps of agricultural land are generated for the visualisation of data and for spatial analysis[279]. Then, a forecast of forthcoming field conditions through evaluating the previous and current state of each field can be made, while also parts of land that require special assessment can be identified. In the end, by analysing field location, crop type, planting time, and weather conditions daily instructions for each separate land patch can be sent to crews, thus providing a a specialised targeted approach.

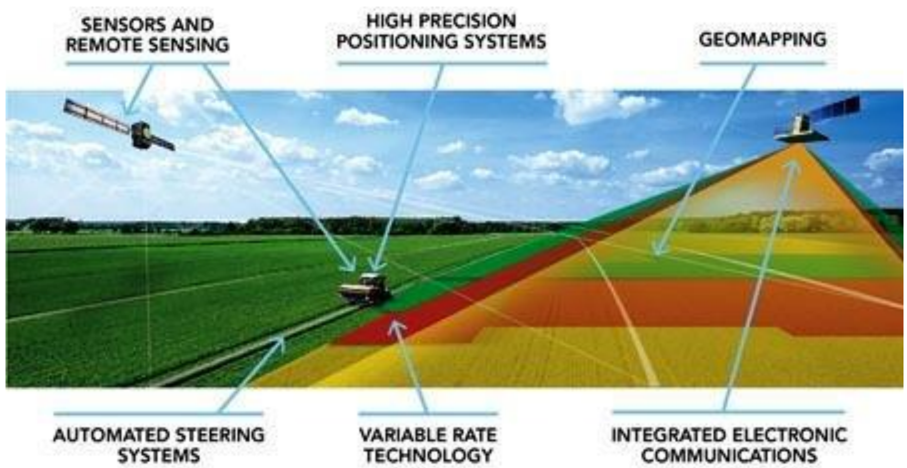
Crop production optimisation: ML algorithms can also analyse vast amounts of agricultural data to predict yield rates, identify crop diseases, analyse soil samples and determine optimal precision irrigation schedules while reducing water and chemical usage. By enabling precise resource allocation and decision-making, AI has the potential to maximise crop yields while minimising resource waste. A company called Prospera combines AI, computer vision and ML to monitor crop health, optimise irrigation, and improve farm management[280] and the World Bank has also developed more than 10 Climate- Smart Agriculture Investment Plans (CSAIPs) for Bangladesh, Zimbabwe, Zambia, Lesotho, Mali, Burkina Faso, Ghana, Cote D'Ivoire, Morocco and The Republic of Congo[281].

Pest and Disease Management: Effective pest and disease management is crucial for maintaining healthy crops. AI can also assist in identifying, monitoring, and combating pests and crop diseases, as drones equipped with AI-powered sensors can survey vast agricultural areas, detecting signs of pests

---

279 See Ghosh, P., & P. Kumpatla, S. (2022). *GIS Applications in Agriculture.* IntechOpen. doi: 10.5772/intechopen.104786.
280 See more at: https://prospera.ag/about-us/.
281 The World Bank (2021). *Climate Smart Agriculture,* at: https://www.worldbank.org/en/topic/climate-smart-agriculture.

or diseases in real-time. This prompt detection allows farmers to take targeted and timely action, minimizing crop losses and reducing the reliance on harmful ex post chemical interventions. AI applications help detect exact areas that have been infected with weeds and pests, reducing usage of herbicides and managing harvest loss in a timely way. Robots that use computer vision and AI are allowing to eliminate 80% of the volume of chemicals that are normally sprayed on the crops and bring down expenditure of herbicide by 90% in some cases. On a practical example, the International Crops Research Institute for the Semi Arid Tropics, built an ML downloaded by 12 million users globally that allows farmers to diagnose pest damage, plant disease and nutrient deficiencies by taking a photo of their affected crop[282]. FAO estimates that up to 40% of global crop yields are reduced each year because of pests[283]; thus, these types of AI applications become key assets.



Picture 12

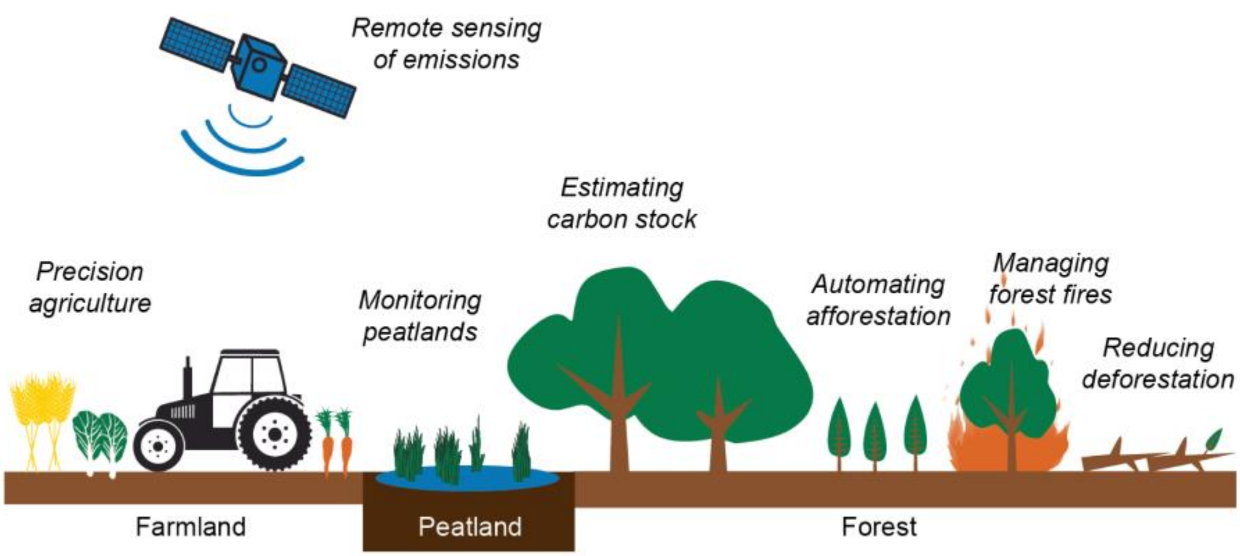
Jat, Ram & Dinesh et al. (2020). *Precision Agriculture and Conservation Agriculture for Resource Conservation, at:* https://www.researchgate.net/publication/344388003_Precision_Agriculture_and_Conservation_Agriculture_for_Resource_Conservation

[282] Gómez Mont C. (2021). *The Intersection between AI and the Climate and Biodiversity Crisis*, AI Ethics: Global Perspectives.
[283] IPPC Secretariat (2021). *Scientific review of the impact of climate change on plant pests – A global challenge to prevent and mitigate plant pest risks in agriculture, forestry and ecosystems*. Rome. FAO on behalf of the IPPC Secretariat, https://doi.org/10.4060/cb4769en.

Livestock monitoring: Of course, similar monitoring and assessment applications can be applied to livestock management as well: through AI, live video feeds of livestock can by analysed and animal behavior, health, and welfare be monitored. AI-powered systems can recognize subtle changes in livestock behavior that may indicate illness, heat stress, or inadequate nutrition, enabling early intervention and ensuring animal well-being.

3.3.5. AI, Land Uses and Biodiversity

Even though agriculture was reviewed separately, the positive contribution of AI towards sustainability in the Land Uses sector and Biodiversity should also be examined.



Picture 13
Opportunities to use AI in Land uses
Rolnick, D. et al. (2019) *Tackling Climate Change with Machine Learning*, at: https://arxiv.org/abs/1906.05433, 42:31.

3.3.5.1. *Forests and peatlands*

One significant application of AI is **predictive modelling** in forests; by analysing various factors such as climate data, soil composition, and historical forest dynamics, AI algorithms can forecast changes in forest conditions over time, which enables a more effective forest managers to mitigate the impact of natural disturbances like wildfires and droughts, as well as human activities such as logging and deforestation. For instance, AI can provide early warning systems to predict the spread of wildfires, allowing for timely evacuations and effective firefighting strategies. Rodnick et al. report the importance of sustainable management of peatlands, as they hold twice the total carbon in all the world's forests, making peat the largest source of sequestered carbon on Earth[284]; AI can help towards monitoring and protection of peatlands from droughts and fires that would lead to the emission of huge amounts of carbon into the atmosphere.

Forest management can also greatly benefit from AI, as satellites equipped with AI algorithms could analyse remote sensing data to detect and monitor changes in forest cover, identify illegal logging activities and poaching as well as estimate forest carbon stocks. Additionally, AI-powered drones can be used to monitor vast areas of forest and collect real-time data on forest health, species distribution, and habitat quality. Then, by evaluating all gathered data, AI can assist in the drafting of forest management plans, as well as calculating the need to increase or halt timber production and predicting optimal logging schedules and areas[285].

Furthermore, the important practice of **afforestation** (planting of trees) can be enhanced with AI, as ML can automate large-scale afforestation by locating appropriate planting sites, monitoring plant health,

---

[284] Faizal Parish, A. A. Sirin, D. Charman, Hans Joosten, T. Yu Minaeva, and Marcel Silvius. (2008). *Assessment on peatlands, biodiversity and climate change: Main report*. Global Environment Centre, Kuala Lumpur andWetlands International, Wageningen
[285] For more see Rana, P., Miller, D. (2019) Machine learning to analyze the social-ecological impacts of natural resource policy: insights from community forest management in the Indian Himalaya, *Environ. Res. Lett.* 14 024008, DOI: 10.1088/1748-9326/aafa8f.

assessing weeds, and analysing trends[286]. For example, a company called Mast uses end-to-end tech for reforestations by using satellites to create reforestation plans and then plants tailored seeds while monitoring their progress[287].

3.3.5.2. Ecology

The rapid decline in biodiversity poses a significant hazard for ecosystem collapse, which in turn undermines the indispensable life-sustaining functions provided by biodiversity to humanity. In this fight, AI can play a significant role in supporting efforts aimed at mitigating biodiversity loss. Having already recognised the need for innovative approaches, the integration of AI into biodiversity management has shown promise.

More specifically, AI has emerged as a game-changer in automating species **identification** processes, enabling accurate and efficient species **monitoring**. ML algorithms can analyse vast amounts of data provided by mobile and stationary stations (as depicted below), such as images or vocalisations, to identify and then classify species with remarkable precision. This approach reduces human error, increases monitoring coverage, and enables the identification of vulnerable species, contributing to their protection. These ML approaches have the potential to unlock ecological understanding and to promote conservation research on a scale difficult to imagine in the recent past[288].

---

[286] Rolnick, D. et al. (2019) *Tackling Climate Change with Machine Learning*, at: https://arxiv.org/abs/1906.05433, 42:33.
[287] See https://www.mastreforest.com/reforestation.
[288] Tuia, D., Kellenberger, B., Beery, S. et al. Perspectives in machine learning for wildlife conservation. *Nat Commun* 13, 792 (2022). https://doi.org/10.1038/s41467-022-27980-y.

Picture 14

Data used in biodiversity monitoring

Tuia, D., Kellenberger, B., Beery, S. et al. *Perspectives in machine learning for wildlife conservation*. Nat Commun 13, 792 (2022). https://doi.org/10.1038/s41467-022-27980-y

Besides optical data, a growing application of AI for biodiversity monitoring involves acoustic[289] and thermal[290] data from airborne sensors (drones, unmanned aerial vehicles-UAVs[291]). It is also remarkable that citizen science can also enable dataset collection at a scale impossible in individual studies, as in the case of eBird[292], which, by leveraging public enthusiasm for birdwatching, has logged more than 140 million observations which have been used for population and migration studies[293].

---

[289] Towsey, M., et al. (2012). *A toolbox for animal call recognition*, Bioacoustics, 21:2, 107-125, DOI: 10.1080/09524622.2011.648753

[290] Psiroukis, V. et al. (2021). Monitoring of free-range rabbits using aerial thermal imaging, *Smart Agricultural Technology*, Volume 1, 100002, https://doi.org/10.1016/j.atech.2021.100002.

[291] GPAI (2022). *Biodiversity & Artificial Intelligence, Opportunities and Recommendations*, Report, Global Partnership on AI.

[292] At: https://ebird.org/about.

[293] Rolnick, D. et al. (2019) *Tackling Climate Change with Machine Learning*, at: https://arxiv.org/abs/1906.05433, 42:41.

Furthermore, AI can greatly contribute to **predictive modelling**, as it can examine historical data and identify patterns in species behaviour or habitats' evolution. By analysing diverse variables, such as species distribution, threat levels, and stakeholder interests, AI algorithms can also optimise resource allocation towards species **conservation**. Last but not least, AI can power **Invasive Species Warning Systems** through prediction of areas prone to invasion and to timely identify invasion cases, thereby helping with early response protocols, as well as identifying poachers in protected ecosystems[294].

As we strive to protect and restore Earth's dwindling biodiversity, the integration of AI into biodiversity management holds tremendous promise. By harnessing AI's potential in biodiversity and ecosystems' protection, a sustainable path for preserving our planet's irreplaceable natural heritage can be charted.

3.3.6. AI and Circular Economy

3.3.6.1. The concept of circularity

Technological innovation, including AI, plays a major role in bringing the vision of "**circularity**" to life; in short, circularity or circular economy is an economic system that aims to minimise waste and promote resource efficiency by having resources "in a loop" throughout economy and society. It goes beyond the traditional "take-make-dispose" approach through recycling, reusing, and remanufacturing. In a circular economy, products are designed to be durable, repairable, and recyclable, and waste is viewed as a potential resource; this huge shift in patterns requires, of course, systemic changes in supply chains, business models and consumer behavior and is no easy feat.

Until now, economic growth has been based on a model of intensive use of available and imported resources without taking into account - at least to the extent that it should - the long-term effects on

[294] For example, Microsoft's Azor Machine Learning Program combats poaching. See more at: Liakhovich, O., Zolochevska, A. (2019). *Preventing Rhino Poaching with Microsoft Azure* https://devblogs.microsoft.com/cse/2019/05/07/preventing-rhino-poaching-though-microsoft-azure/

resources. The change proposed through Circular Economy is based on maximising the efficiency of existing resources by reducing losses in their use. This implies a total reassessment of resources while at the same time revising the established concept of waste. At the same time, it calls for abandoning the current approach of extracting limited raw materials from the ground, using them to produce a product and then disposing of them and burying them underground (the traditional model of "linear economy")[295]. This change requires a shift from the "supply, production, consumption and disposal" model to a model based on the fourfold "reuse, repair, refurbish and recycle".

The term "**circular economy**" comes from an economic model by Pierce & Turner (1990), which was the first coherent critique of the traditional system of linear economy[296]. The relationship between the economy and the environment is dominant in this model, which incorporates three functions of the environment: resource provision, waste absorption and source of benefit[297]. Already from this model the relationship between economy and environment is inseparable covering three economic functions of the environment: resource provision, waste absorption, source of benefit.

In the EU, the European Commission adopted the 1st Action Plan for Circular Economy in 2015, defining the concept of the circular economy as "*the transition to a more circular economy, where the value of products, materials and resources is retained in the economy as much as possible, and waste production is minimised, is an essential contribution to the EU's efforts to develop a sustainable, efficient and competitive low-carbon economy with efficient use of resources. Such a transition will be an opportunity to transform the European economy and give Europe new, sustainable competitive advantages*"[298]. As part of the EU Green Deal, the Circular Economy

---

[295] European Environment Agency. (2016). *Circular Economy in Europe – Developing the knowledge base*, EEA Report, p. 8-9.
[296] D. W. Pearce/R. K. Turner (1990). *Economica of Natural Resources and the Environment*, The John Hopkings University Press.
[297] Rizos, V., Tuokko, K. Behrens, A. *The Circular Economy A review of definitions, processes and impacts*, at: https://www.researchgate.net/publication/315837092_The_Circular_Economy_A_review_of_definitions_processes_and_impacts, p. 2.
[298] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS *Closing the loop - An EU action plan for the Circular Economy*, COM/2015/0614 final, p. 2.

Action Plan was published in March 2020 and presented measures to (a) make sustainable products the norm, (b) to empower consumers and buyers, and (c) to ensure less waste generation[299].

The core of Circular Economy concept consists of the following main axes:

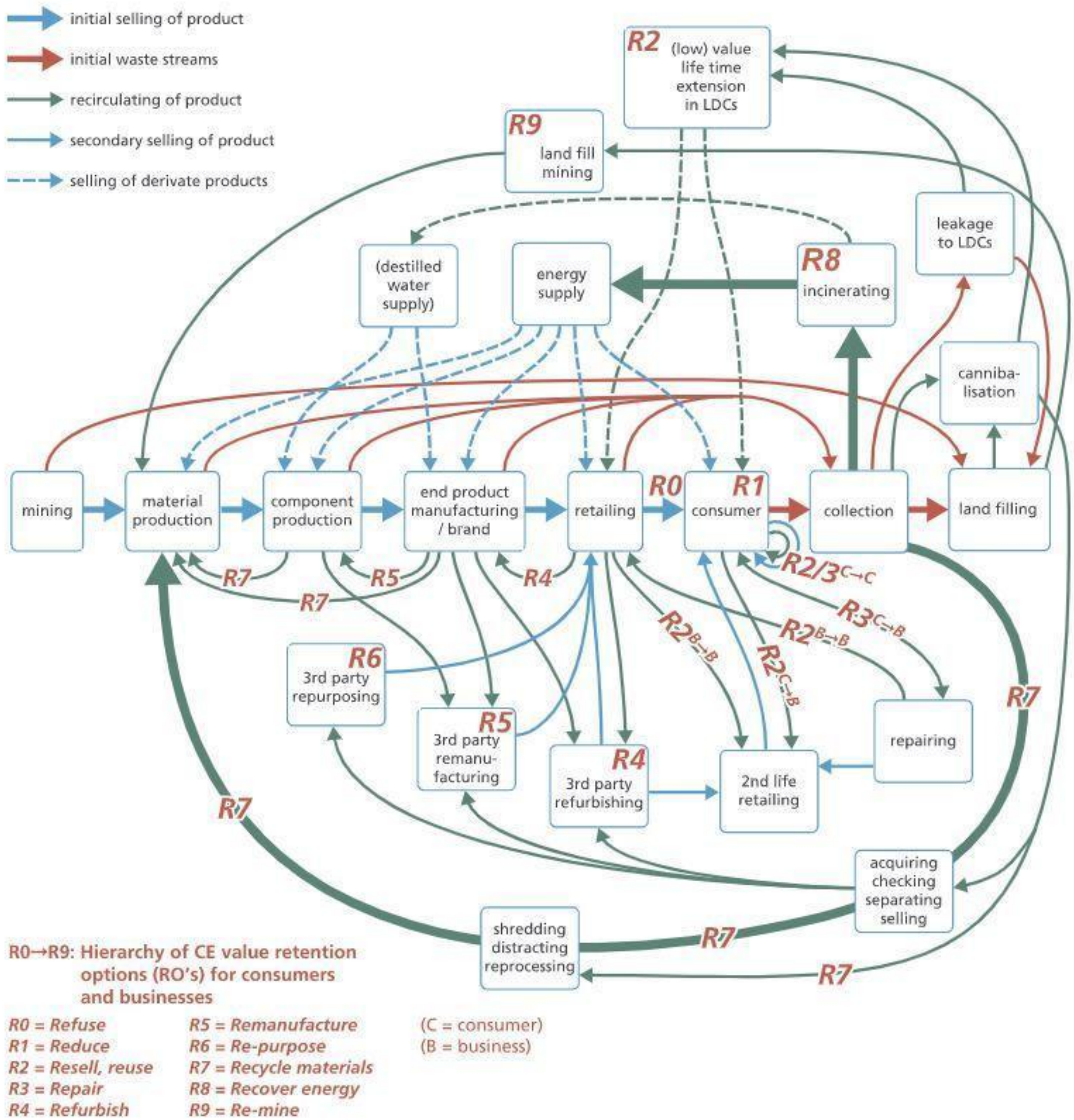
- recycling,
- resource efficiency,
- use of renewable energy,
- extending the life of products,
- remanufacturing,
- upgrading/renewal and reuse of products and their components, and
- changes in consumption patterns.

In the context of international scientific analysis, the current typology of Circular Economy is oriented around the application of **10 principles**, which, because they begin with the suffix re-, i.e., they imply repetition, hence the cycle of the process, are known as **10Rs:**

1. R0 = **Refuse**
2. R1 = **Reduce**
3. R2 = **Resell**, **Reuse**
4. R3 = **Repair**
5. R4 = **Refurbish**
6. R5 = **Remanufacture**
7. R6 = **Re-purpose**
8. R7 = **Recycle materials**
9. R8 = **Recover energy**

[299] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *A New Circular Economy Action Plan for a Cleaner and more Competitive Europe,* COM/2020/98 Final.

10. R9 = **Re-mine**



Picture 15

Reike, D., Vermeulen, W. J., & Witjes, S. (2018), doi: 10.1016/j.resconrec.2017.08.027

The first four principles (R0-3) are placed close to the consumer and can be linked to commercial or non-commercial actors involved in extending the life of the product. Scholars applying a clear hierarchy identify them as the most preferred R principles in the Circular Economy.

The second group of principles (R4-6) includes refurbishment, remanufacturing and re-purposing of materials; for these principles commercial business activity is the main driver, with often specialised actors with high levels of expertise as stakeholders.

The third group (R7-9) refers to traditional waste management activities, including recycling, various forms of energy recovery and, more recently, re-mining[300]. Many scholars who apply clear hierarchies to their Rs agree that these options are the least desirable. A key challenge here is how to achieve the application of higher value recycled materials, especially in countries where mass recycling is already well organised (especially in Northwest and Central Europe)[301].

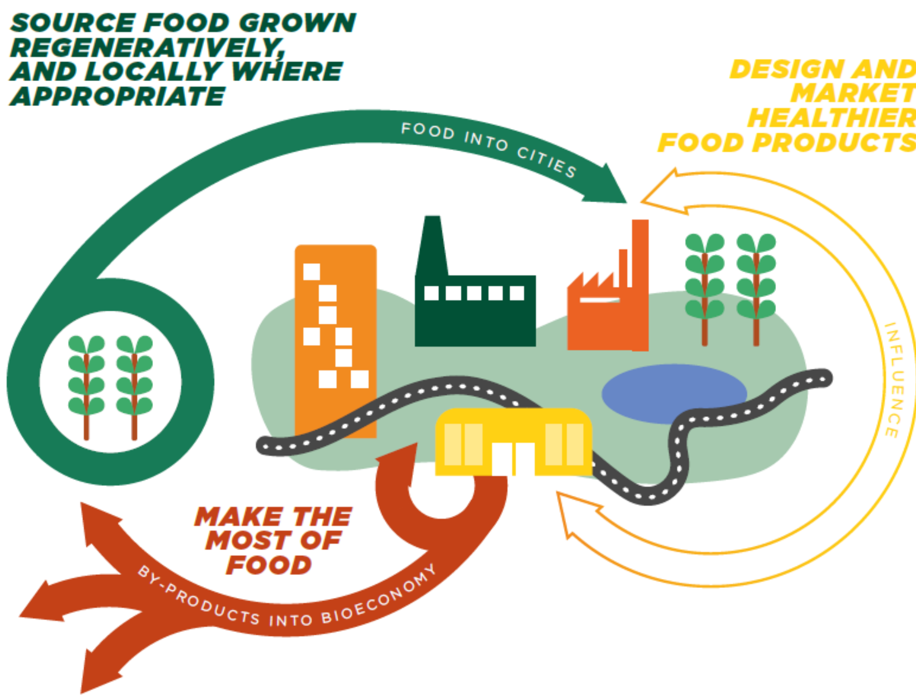
One of the key characteristics of Circular Economy is that, especially by utilising AI's potential, it can be integrated into the majority, if not all sectors of economy, as depicted below:

[300] Recovery of materials after the landfill phase, recovery of a limited set of available products or components from the materials (which is also called "*cannibalisation*"), high-tech landfill mining or urban mining (*i.e.,* using advanced waste collection and separation methods to extract iron, zinc, copper and plastics).

[301] See more at Vermeulen, W., Reike, D, & Witjes, S. (2019). *Circular Economy 3.0 - Solving confusion around new conceptions of circularity by synthesising and re-organising the 3R's concept into a 10R hierarchy,* at: https://www.researchgate.net/publication/335602859_Circular_Economy_30_-_Solving_confusion_around_new_conceptions_of_circularity_by_synthesising_and_re-organising_the_3R's_concept_into_a_10R_hierarchy .

Picture 16
Circular Strategies in modern economy
Neufield, D. (2022). Visualized: The Circular Economy 101, at: https://www.visualcapitalist.com/sp/visualized-the-circular-economy-101/

For example, changing our food system to a circular one can be one of the most important actions to tackle climate change, as its production and consumption has fuelled urbanisation and has supported a rapidly growing population, which has come at a huge cost to society and the environment in the form of food waste and no equal access to healthy food. For example, in Switzerland, the 2.8 million tonnes of food waste produced throughout the value chain is estimated to be equivalent to a land area equivalent to half of the country's agricultural land[302].



Picture 17

Circular Economy for food in Cities

Ellen MacArthur Foundation. (2019). Artifi*cial intelligence and the circular economy - AI as a tool to accelerate the transition*, at: http://www.ellenmacarthurfoundation.org/publications, p. 18

As an Ellen McArthur Foundation report depicted above, in a circular economy model[303] for cities in the field of food production, agriculture is regenerative and food is produced locally where appropriate; then, food is designed and marketed to become healthier, whereas food by-products are seen as value streams into bioeconomy. And with the rise of "ag-tech", as analysed in the above "AI in Agriculture

[302] EEA Report (2020). *Bio-waste in Europe — turning challenges into opportunities*, available at: https://www.eea.europa.eu/publications/bio-waste-in-europe

[303] For more on circular economy in the legislative field see Pouikli, K., Delegkou, A., Tsakalogianni, I. (2022). Circular Economy for a climate-neutral Europe: Mapping the latest policy and legislative developments, *Carbon & Climate Law Review*, Volume 16, Issue 4 (2022) pp. 281 - 287 DOI: https://doi.org/10.21552/cclr/2022/4/8 (EN)

and Food Production" section, a new set of food and agricultural solutions are emerging; apart from the harvest stage, AI could contribute to reduction of food waste during **processing**, **distribution** and **storage** as well as during **preparation** and **consumption**. For example, the Belgian company TOMRA uses automation to reduce food waste in food processing stages and help valorise produce which may not be suitable for direct sale[304]. AI could also accelerate food inspection time as well as cut waste during the preparation process, for example in commercial kitchens; the AI tool "Winnow Vision" can automatically track food waste, cut costs and save time in restaurants[305].

3.3.6.2. AI in Circular Industry

a) Consumer electronics sector

In a circular economic future, e-waste is a concept of the past[306]; to this end, AI-driven tools can assist in **a**. These variables often interact with each other in complex ways, making it challenging for human designers to find the optimal balance. However, AI can analyse vast amounts of data and identify patterns that humans might overlook, leading to more refined, well-rounded and **sustainable** designs. In this way, AI can be utilised to create electronics that are easier to repair, upgrade and recycle in order to honor the Circular Economy principles and reduce the environmental impact of the industry.

Another significant advantage of AI technology in product development is the ability to analyse feedback from existing products in use. By collecting data on user experiences, preferences, and various channels like customer reviews or usage statistics, AI algorithms can extract valuable insights. These insights can then be incorporated into new designs to address shortcomings.

More AI benefits in the so-called "Circular Manufacturing" (CM) are depicted below:

---

[304] See more at: https://www.tomra.com/en/food.
[305] See more at: https://www.winnowsolutions.com/company.
[306] Ellen MacArthur Foundation. (2019). *Artificial intelligence and the circular economy - AI as a tool to accelerate the transition*, at: http://www.ellenmacarthurfoundation.org/publications, p. 26.

Scale	*Entity impacted*	*CM strategy*	*AI benefits*
Micro	Product	Circular Design	Tracking the material starting from the acquisition phase. Prototyping the new product without wasting resources during the test phase Keeping high product modularity exploiting AI to prototype products Gathering data from smart products to improve next generation products on the basis of end-users behaviours
		Servitization	Tracking and monitor the product usage to improve the service provided
		Disassembly	Defining the best and most efficient disassembly path relaying on AI in collaborative robots
		Reuse	Tracking the product to monitor the conditions and evaluate whether the product reuse is possible
		Recycle	Tracking the product to monitor the conditions and evaluate whether and how the product components and materials can be recycled
	Process	Waste Management	Tracking the type of material present in the waste to evaluate its recyclability or disposal
		Resource Efficiency	Tracking energy, water, and other resources usage during the production process
		Cleaner Production	Tracking energy, water, and other resources usage during the production process to evaluate possible improvements
Meso	Network of firms	Closed-loop Supply Chain	Creating collaboration by mapping the most convenient circular path Forecasting return products quality, quantity, and time Tracking products in real-time to estimate the residual value Tracking vehicles to manage the loading of waste and recyclable resources
		Circular Design	Designing the product considering the actors involved in the value chain
		Remanufacture	Tracking the turned back product to monitor the conditions and evaluate whether the product can be remanufactured Exploiting product data, once returned, so to define the best remanufacturing path
Macro	Nations, Regions, Cities	Waste Management	Keeping track of circular performance to forecast it in nations, regions and cities Monitoring municipal waste type and quantity

Figure 17

AI benefits in CM

Federica Acerbi, F., Forterre, D.A., Taisch, M. (2021) Role of Artificial Intelligence in Circular Manufacturing: A Systematic Literature Review, *IFAC-PapersOnLine*, Volume 54, Issue 1, https://doi.org/10.1016/j.ifacol.2021.08.040, p. 370-371.

Focusing on waste, more than 50 million metric tons of e-waste is generated globally every year, averaging 7 kg of e-waste per capita[307]. AI can significantly improve **waste management systems**, as it can first keep track of the materials in the generated waste in order to evaluate is recyclability or suggest

[307] According to Statista, as published in May 2023, at: https://www.statista.com/topics/3409/electronic-waste-worldwide/.

its disposal; AI systems equipped with computer vision can efficiently identify recyclable materials, sort them accordingly, and separate contaminants. Then, automated processes can reduce possible contamination by using ML algorithms that continuously learn and adapt to optimise disposal processes.

In the core of business, according to the Ellen McArthur Foundation, products like consumer electronics require specific **circular business models**, which AI can help design and include features such as incentives to ensure a reliable return flow of products and components and efficient reverse logistics processes for collecting and transportation[308].

Circular business models represent fundamentally different ways of producing and consuming goods and services, which have the potential to promote the transition towards a more resource-efficient and circular economy and, in doing so, significantly reduce the environmental pressure resulting from economic activity. The environmental outcomes of circular business models also depend on their market penetration. However, the market share of these business models is currently limited. Linear business models are based on the following logic: use of natural resources, production of products for consumers, production of waste. In contrast, circular business models contribute to the circular economy by adhering to the three fundamental principles of the circular economy: a) Elimination of waste and pollution, b) Preservation of products and materials in use, and c) Regeneration of natural ecosystems.

---

[308] Ellen MacArthur Foundation. (2019). *Artificial intelligence and the circular economy - AI as a tool to accelerate the transition*, at: http://www.ellenmacarthurfoundation.org/publications, p. 28.

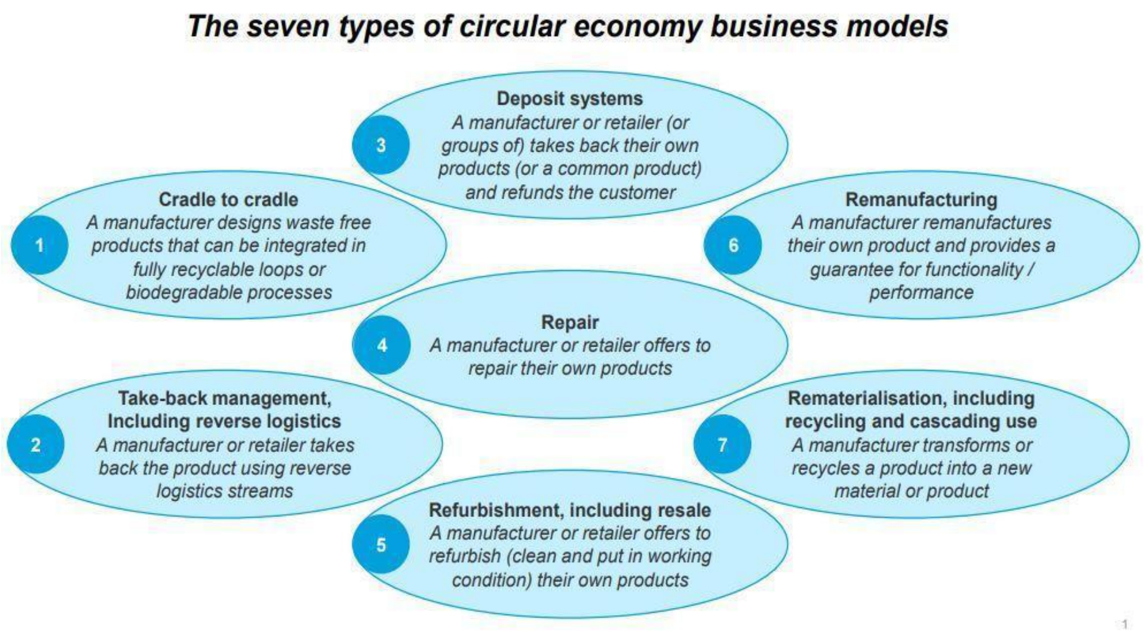


Figure 18

See Deloitte, *Circular Economy - From theory to practice, at:*

https://www2.deloitte.com/content/dam/Deloitte/fi/Documents/risk/Circular%20economy%20FINAL%20web.pdf

Of course, targeted policy interventions are needed to create the conditions for the wider adoption of circular policies and consequently circular business models in business activity. Ultimately, the transition to a much more circular and resource-efficient economy will require a wider penetration of circular business models into the whole structure of society and economy. In this context, policy and regulation can play an important role in addressing the market barriers, policy gaps and various biases that currently hinder the competitiveness of these business models.

b) Textiles

Textiles have long been an integral part of our daily lives and our society, providing employment for hundreds of millions of people and creating enormous economic value. However, the industry has a significant and negative impact on the environment, as, according to the European Environmental Agency, textile consumption in Europe in 2020 had on average the 4th highest impact on the environment and climate change and it was the consumption area with the 3rd highest impact on water and land use and the 5th highest in terms of raw material use and GHG emissions[309]. The EU has once again been a pioneer in recognising the need for more sustainable practices in the industry, as the first Strategy for Sustainable Textiles was published in 2022[310], following on the EU Green Deal's call for circularity in all sectors of economy.

As analysed in the previous section, AI can greatly contribute to the optimisation of production stages; and it is no different in the textile industry: AI can design business models that promote the need for recycled (synthetic or natural fibres) or renewable materials (such as cotton or wood-based fibres from sustainable agricultural practices), or both (such as recycled vegetable fibres) for production. Emerson is relying AI running on data collected from sensors on the factory floor to help provide a single pane view of operations for increased efficiency and decreased energy use[311].

Furthermore, it can enhance the collection and recycling process of textiles, in order to have fabrics broken down into fibres, which can be reused to make new yarns; this recycling of materials will naturally reduce GHG emissions in production, however attention should be paid to the case of energy-intensive recycling technologies which use high carbon energy sources.

---

[309] European Environmental Agency. (2022). *Textiles and the environment: the role of design in Europe's circular economy*, at: https://www.eea.europa.eu/publications/textiles-and-the-environment-the.

[310] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS EU, *Strategy for Sustainable and Circular Textiles*, COM/2022/141 final.

[311] See (2022). *The AI Sustainability Paradox – And How to Solve It*, at: https://unifiedguru.com/the-ai-sustainability-paradox-and-how-to-solve-it/ and Dell Technologies, *Driving industrial innovationto transform your business*, at: https://www.dell.com/en-us/dt/oem/industrial-automation.htm#scroll=off&tab0=0&video-overlay=6314733967112

D

Picture 18

Textiles production future scenarios

PACE Platform for Accelerating Circular Economy. (2021). *Circular Economy Action Agenda, Textiles'*, at: https://pacecircular.org/sites/default/files/2021-02/circular-economy-action-agenda-textiles.pdf

3.3.6.3. *Legal and ethical issues regarding AI and Circular Economy*

Although the positives around the use of AI in the field of Circular Economy are multiple, there are certain legal and ethical aspects which require attention.

First of all, one of the key concerns regarding AI in general and in the circular economy as well is **data privacy** and **protection**. AI systems rely heavily on vast amounts of data, including personal information, to derive insights and make informed decisions; for example, circular tracking and measurement devices that track a person's location and the product's condition in order to access the need for upgrades/repairs lead to the processing of sensitive geospatial data of the consumer[312]. Roberts et al. mention another example: smart energy meters can identify and store data on patterns in energy usage that point to when individuals wake up, go to sleep, go to work, are away and have guests over, amongst many others[313]. Hence, issues such as data ownership and users' consent as regards circular devices should be properly addressed in a more specific legal space than general frameworks like the General Data Protection Regulation (GDPR).

Furthermore, an issue we haven't touched upon is the labour market, as the staggering use of AI may lead to **job displacement** and the replacement of job roles with AI systems. In the field of Circular Economy, we have mentioned the contribution of AI to the optimisation of industry processes and the growing automation of processing, distribution, storage, etc.; however, the more processes like food sorting, product inspection, data collection, waste management and general factory work are automated, the greater is the possibility of total elimination of said positions.

To this end, it is crucial to develop mechanisms to support the reskilling and upskilling of workers affected by technological advancements and to ensure the so-called "**Just Transition**[314]". The Just Transition is understood as the transformation of economies and their sectors in the context of combining climate action and social inclusion, provided that there are no adverse economic and social impacts throughout said transition. According to a 2018 International Labour Office (ILO) report, while 18 million new green jobs are projected to be created by 2030, 14 economic sectors show employment losses of more than 10,000 jobs worldwide and 2 (petroleum refinery and extraction of crude petroleum)

---

[312] See Roberts, H., Zhang, J., Bariach, B. et al. (2022). Artificial intelligence in support of the circular economy: ethical considerations and a path forward. *AI & Soc*, https://doi.org/10.1007/s00146-022-01596-8 and Nobre, GC., Tavares, E. (2017). Scientific literature analysis on big data and internet of things applications on circular economy: a bibliometric study. *Scientometrics* 111(1):463–492. https://doi.org/10.1007/s11192-017-2281-6
[313] See Murray, D, Liao, J, Stankovic, L, Stankovic V. (2016). Understanding usage patterns of electric kettle and energy saving potential. *Appl Energy* 171:231–242. https://doi.org/10.1016/j.apenergy.2016.03.038
[314] See more on Healy, N., Barry, J. (2017). Politicizing energy justice and energy system transitions: Fossil fuel divestment and a "just transition", *Energy Policy*, Volume 108, 2017, Pages 451-459, ISSN 0301- 4215, DOI: 10.1016/j.enpol.2017.06.014.

show losses of 1 million or more jobs[315]. Just Transition calls for social policy and vocational integration through poverty reduction, protection of minorities, transformation of the labour sector, the creation of new employment positions and professional training; thus, it is integral to the concept of Sustainability. Internationally, the concept of Just Transition has been integrated into ILO guidelines[316], the preamble of the Paris Agreement[317], the Silesia Declaration on Solidarity and Just Transition during COP24[318] and the Just Transition Declaration during COP26[319], while the EU has adopted the important Regulation 2021/1056 on the Just Transition Fund[320]. By investing in **education** and **training** programs rather than layoffs or demotions, societies can try to mitigate the possible negative impact of AI on employment and create new opportunities in the Circular Economy driven by AI technologies.

Another important consideration, as usual in the case of AI, is ensuring that no **unfairness** or **biases** are integrated into the algorithm; as the AI system is trained on historical data, some it may inherently contain biases against a certain gender or ethnicities and favour certain regions or demographics. If biases are identified, the equitable distribution of Circular Economy practices and products may be at risk. Miller and Hosanagar touch upon the issue of perpetuating biases when using personal characteristics for personalised marketing[321]; the same phenomenon could appear when pricing and advertising circular economy products.

---

[315] International Labour Office. (2018). *World Employment and Social Outlook 2018: Greening with jobs* – Geneva: ILO, p. 43

[316] ILO. (2015). *Guidelines for a just transition towards environmentally sustainable economies and societies for all*, at: https://www.ilo.org/wcmsp5/groups/public/@ed_emp/@emp_ent/documents/publication/wcms_432859.pdf.

[317] Paris Agreement to the United Nations Framework Convention on Climate Change, Dec. 12, 2015, T.I.A.S. No. 16-1104: "*Taking into account the imperatives of a just transition of the workforce and the creation of decent work and quality jobs in accordance with nationally defined development priorities*".

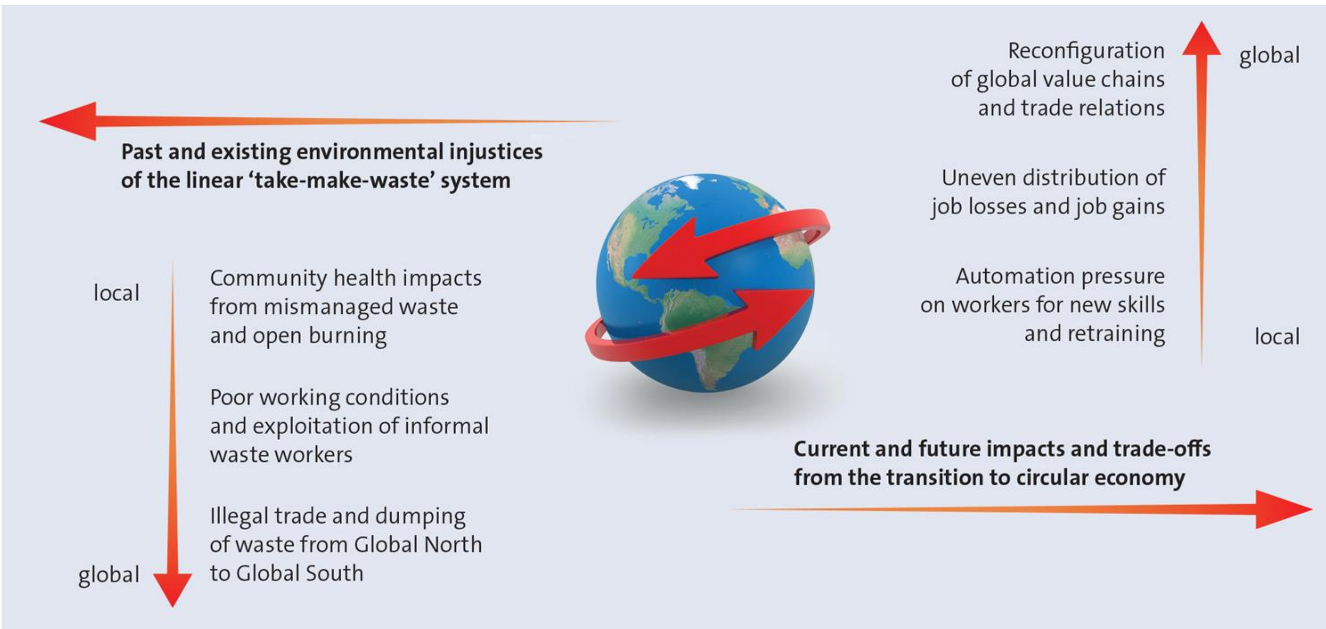
[318] Ministerial Declaration "Solidarity and Just Transition Silesia Declaration" as adopted by acclamation at the Leaders' Summit during the 24th Conference of the Parties (COP24) to the United Nations Framework Convention on Climate Change (UNFCCC), on 3 December 2018 in Katowice, Poland.

[319] As adopted during the 26th Conference of the Parties (COP26) to the United Nations Framework Convention on Climate Change (UNFCCC), on 4 November 2021 in Glasgow, Scotland.

[320] Regulation (EU) 2021/1056 of the European Parliament and of the Council of 24 June 2021 establishing the Just Transition Fund. The Fund focuses on economic diversification of the territories most affected by climate transition (i.e., with targeted economic incentives to specific areas) and on the re-skilling of workers and job seekers.

[321] Miller, AP., Hosanagar K. (2019). How targeted Ads and dynamic pricing can perpetuate bias. *Harvard Business Review*. https://hbr.org/2019/11/how-targeted-ads-and-dynamic-pricing-can-perpetuate-bias.

The concept of **environmental justice** is also relevant, as marginalised communities may have to bear the disproportionate burden of waste management and recycling facilities operating in the areas they work and live in. Moreover, the illegal dumping of waste into low-income communities or waste shipments to low- and middle-income countries result in severe health impacts on communities, waste workers and their families, as in the case of lead and air pollution from unsafe e-waste and battery recycling[322].



Picture 19

Past and existing environmental injustices and current and future trade-offs

Schroeder, P. and Barrie, J. (2022). Is going circular just? Environmental justice and just transition – key elements for an inclusive circular economy, Field *Actions Science Reports [Online]*, Special Issue 24, http://journals.openedition.org/factsreports/6864

322 Schroeder, P. and Barrie, J. (2022). Is going circular just? Environmental justice and just transition – key elements for an inclusive circular economy, *Field Actions Science Reports* [Online], Special Issue 24, http://journals.openedition.org/factsreports/6864

Thus, an equitable implementation of circular economy should prioritise addressing these disparities and prevent perpetuating environmental injustices. In the field of Law and regulation, robust and enforceable regulations must be in place to prevent the concentration of environmental hazards in vulnerable communities; for example, stronger efforts are needed to effectively integrate the concept of environmental justice into the decision-making processes of the authorities responsible for land use planning, permitting, and pollution control, in order to push for a more proportionate and fair distribution of environmentally practices. Environmental Impact Assessments could also consider including a "social impact assessment", by examining the potential differential impacts of projects, plans and programmes on disadvantaged and vulnerable communities[323].

Of course, one should not forget the possible **environmental damage** that AI in conjunction with circular economy could cause; as it was briefly addressed previously, circular processes like recycling of textiles often demand large amounts of water, while reconditioning, refurbishing, or remanufacturing of products may require significant energy inputs, particularly when the original manufacturing processes were not designed with circularity in mind; this extensive energy consumption might actually do more harm than good and have negative implications for carbon emissions, thus contributing to climate change. Same with waste production through circular economy, as GHG emissions and pollutants are produced during these procedures as well. At the same time, circular economy may inadvertently drive overconsumption, as the emphasis on reusing and recycling materials can create a false sense of security, potentially leading to increased production of waste. These **environmental "trade-offs"** should then be properly addressed by scientists and policymakers, as without careful planning and strategies that weigh the negatives versus the positives of circular economy practices, the promotion of environmental protection and sustainability through circular economy cannot be ensured.

---

[323] See Directive 2011/92/EU of the European Parliament and of the Council of 13 December 2011 on the assessment of the effects of certain public and private projects on the environment (codification) and Directive 2001/42/EC of the European Parliament and of the Council of 27 June 2001 on the assessment of the effects of certain plans and programmes on the environment.

As regards environmental and social damage, establishing frameworks for proper allocation of responsibility is crucial. **Accountability** frameworks are an essential aspect of AI implementation, as analysed before in the present thesis, and they become particularly important in the circular economy field. Where resources are continuously reused, reprocessed, remanufactured and recycled, chances are that potential environmental and social damages may occur by AI-driven decisions; so, the identification of the person(s) liable for the AI's shortcomings becomes very important within Circular Economy too.

All in all, the previous analysis makes it evident that the Circular Economy should not be perceived as a refined or enhanced approach to recycling but rather as a **transformational shift in current economic and consumption models towards sustainability,** which demands a careful and targeted approach through a technical, regulatory and ethical lens.

SECTION 4: Observations and Recommendations

4. A Governance Framework for AI

4.1. The Soft Law vs Hard Law Dilemma

When attempting to create a regulatory ecosystem for AI systems' production, training and deployment, it is important to keep in mind that the legal acquis has maybe not yet faced a challenge as complex and multifaceted as AI. In general, one could say that a proper governance framework should establish clear regulatory principles and standards that foster responsible AI deployment. Andrew Ng, co-founder of the online learning platform Coursera, has stated that "*AI is the new electricity. It will transform every industry and create huge economic value*[324]." However, emerging technologies have used soft law instruments[325] In the face of this challenge, several regulatory approaches have been suggested that range from "hard law", i.e., law in the traditional sense, to enacted "soft law", i.e., flexible and not directly enforceable mechanisms like guidelines and codes of conduct. For example, the OECD's Guidelines on AI could be characterised as a soft law approach; at the same time, the EU is close to adopting the first Regulation of AI (i.e., a hard law instrument).

In theory, soft law and its mechanisms are delimited in the area between non-legal positions and legally binding and judicially reviewable commitments[326], while its inclusion in the category of sources of law is a chronically controversial issue[327]. However, its capacity to incorporate the complexity of the [European] legal acquis, placing law in a broader social and political context and covering both non-binding rules with legal significance and binding rules, reflecting legal adaptability and flexibility, while

---

[324] Marchant, G E. et al. (2020). Governing Emerging Technologies Through Soft Law: Lessons for Artificial Intelligence. *JURIMETRICS*, VOL. 61, ISSUE NO. 1, p. 2

[325] See Hagemann, R., Huddleston, J. and Thierer, A. D. (2018). Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future. *Colorado Technology Law Journal*.

[326] Terpan F., (2015), Soft Law in the European Union, *European Law Journal*, 21: 68 -96. DOI: 10.1111/eulj.12090, 70

[327] For example, soft law is not mentioned as a source of international law in the Statute of the International Court of Justice, nor as a derivative of European law (TFEU, Art. 288).

at the same time reducing the cost and administrative burden of its adoption by the legal order concerned, has been positively highlighted[328].

In the field of AI, Buczynski et al. mention two primary uses of AI regulation: a) a technical and prudential approach, which can be defined as rules and best practices in operations, conduct, and governance; and b) a strategic approach, which includes fundamental principles (mainly of ethical nature) that provide guidance to the first use[329].

"**Soft law**" approaches involve a variety of instruments that establish substantive goals or norms that are not directly enforceable[330]. Hence, as an amalgam of technical approach and a principle-driven approach, it has been hailed as a competent responder to the ever-changing and ever-evolving nature of AI[331]. First of all, soft law development is **faster**: in contrast to hard law which is adopted through a step-by-step parliamentary procedure, soft law can be adopted on a non-governmental level, either by industry or separately on private stakeholder level. The internal procedures to adopt a soft guideline-led framework are usually a lot faster than the ones needed for the adoption of a Law or Regulation; of course, the same applies to when a revision or amendment of said instrument is needed, as the process is concluded relatively quickly and efficiently. AVs is an example, as numerous soft law guidelines for developing and operating AVs have been set out globally[332].

---

[328] Saurugger S. & Terpan F. (2021) Normative transformations in the European Union: on hardening and softening law, *West European Politics*, 44:1, 1-20, DOI: 10.1080/01402382.2020.1762440, p. 5

[329] Buczynski, W., Steffek, F., Cuzzolin, F., Jamnik, M., & Sahakian, B. (2022). Hard Law and Soft Law Regulations of Artificial Intelligence in Investment Management. *Cambridge Yearbook of European Legal Studies*, 24, 262-293. doi:10.1017/cel.2022.10, p. 266-267.

[330] Abbott, Kenneth W. and Snidal, Duncan (2000). *Hard and Soft Law in International Governance*, International Organizaion 54:421-456.

[331] Gutierrez, I., C., Marchant, G. (2021). *Soft law 2.0: Incorporating incentives and implementation mechanisms into the governance of artificial intelligence*, https://oecd.ai/en/wonk/soft-law-2-0.

[332] For example, see the Federal Automated Vehicles Policy of USA (September 2016), at: https://www.transportation.gov/AV/federal-automated-vehicles-policy-september-2016, the Guidelines for the Construction of the Internet of Vehicles Cybersecurity and Data Security Standard System of the Ministry of Industry and Information Technology of China (2022) at: https://www.miit.gov.cn/gyhxxhb/jgsj/wlaqglj/zcjd/art/2022/art_9b7b1bf309814505844193372032da9a.html (in Chinese), and the Guidelines for Testing Automated Driving Systems in Canada (2021), at: https://tc.canada.ca/en/road-transportation/innovative-technologies/connected-automated-vehicles/guidelines-testing-automated-driving-systems-canada.

Of course, soft law is also **flexible**; it can evolve in parallel with the subject it attempts to "regulate" and adapt itself accordingly, as well as lead to multiple different soft law concepts to be developed simultaneously (for example, guidelines as well as a set of principles). Another distinct characteristic of soft law is that **soft measures can be gradually "hardened**" into more formal regulatory instruments[333], and thus solidifying themselves in a legal ecosystem in an alternative, adaptable way. For example, G. Marchant observes that when a set of principles becomes well-established in an industry, courts can decide to hold companies liable for damage that occurs from their failure to follow those principles[334].

Quite similar is the concept of "harder soft governance" which was recently found in EU energy and climate law, when seemingly soft instruments are combined with "harder", otherwise of regulatory nature, elements, like: a) rigorous monitoring, b) concrete, quantitative progress reporting, c) mandatory justification in case of no response to recommendations, etc.[335]. In this case, a hybrid form of governance structure is created, which could, possibly, be the suitable regulatory approach for the unique field of AI[336].

However, there are some counter-arguments against soft law mechanisms, which include its ineffectiveness, incredibility, fragmentation and lack of enforcement provisions. To this end, it is important to make soft law as effective and credible as possible so it can address the governance challenges of AI systems[337]. This upping of effectiveness and credibility could be attempted in a twofold way: **a)** through incentives in order to facilitate the expansion of soft law instruments, like a boost in a

---

[333] Gersen, Jacob E. and Psoner, Eric A. (2008). Soft Law: Lessons from Congressional Practice. Stanford Law Review 61:573-627.

[334] Marchant, G. (2021). *Why Soft Law is the Best Way to Approach the Pacing Problem in AI*, at: https://www.carnegiecouncil.org/media/article/why-soft-law-is-the-best-way-to-approach-the-pacing-problem-in-ai.

[335] See Oberthür, S. (2019). Hard or Soft Governance? The EU's Climate and Energy Policy Framework for 2030. 7. 17-27. DOI: 10.17645/pag.v7i1.1796, Knodt M., Ringel M. & Müller R. (2020) 'Harder' soft governance in the European Energy Union, *Journal of Environmental Policy & Planning*, DOI: 10.1080/1523908X.2020.1781604, Schoenefeld J. J. & Jordan J. A (2020) Towards harder soft governance? Monitoring climate policy in the EU, *Journal of Environmental Policy & Planning*, 22:6, 774-786, DOI: 10.1080/1523908X.2020.1792861 and Schoenefeld J. J. & Knodt M. (2021) Softening the surface but hardening the core? Governing renewable energy in the EU, *West European Politics*, 44:1, 49-71, DOI: 10.1080/01402382.2020.1761732

[336] For the 'Race to AI Regulation" see Smuha, N. A., (2019). From a 'Race to AI' to a 'Race to AI Regulation' - Regulatory Competition for Artificial Intelligence, *Law, Innovation and Technology*, Vol. 13, Iss. 1, 2021, http://dx.doi.org/10.2139/ssrn.3501410

[337] Gutierrez, C. I., Marchant, G. E. (2021). *A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence,* http://dx.doi.org/10.2139/ssrn.3855171, p. 35.

company's reputation, a certification by an external independent authority, achieving interoperability of its systems, allocating research funding[338], etc., and **b)** by establishing compliance mechanisms in order to obtain a certification by an independent external entity or achieve a certain "ranking" etc. As regards the latter, possible issues of legitimacy could be dealt with certification/validation of, e.g., ethical quality standards via a neutral third party, like an ISO certification; as Coglianese puts it, standards-setting organisations periodically assess the state of their programs, some with the participation of experts from all sectors of society and by eliciting public commentary[339]. So, this form of **new, "softer" AI governance could be transformed from typical governmental oversight to a broad monitoring by stakeholders that ranges from businesses to private entities and independent auditors to academia and the public itself.**

To this end, it is likely that the complexity of the issues and the rapid pace of development will outstrip the capacities of the regulatory agencies to frame effective policies and standards[340]. So, the private sector takes up a rather responsible role, as it is tasked with shaping the space in which it operates in, in a certain degree.

4.2. Regulatory Sandboxes: a solution between innovation and regulation?

In a more practical sense, in order to establish a symbiotic relationship between hard and soft law in the field of AI, the AI Act itself[341] as well as various regulatory sectors worldwide have proposed the establishment of the so-called "**regulatory sandboxes**" to simultaneously facilitate and monitor the development of AI. As defined by the European Parliament's resolution in Art. 3 of the AI Act (point 44 g) a regulatory sandbox is "*a controlled environment established by a public authority that facilitates the safe*

[338] Marchant, G E. et al. (2020). Governing Emerging Technologies Through Soft Law: Lessons for Artificial Intelligence. *JURIMETRICS*, VOL. 61, ISSUE NO. 1, p. 138.
[339] *Id*., p. 145. See Coglianese, C. (2020). Environmental Soft Law as a Governance Strategy, 61 *JURIMETRICS* J. 19 (2020), p. 33–38.
[340] Lee, R., Jose, P.D. (2008). Self-interest, Self-restraint and Corporate Responsibility for Nanotechnologies: Emerging Dilemmas for Modern Managers. *Technology Analysis & Strategic Management* 20:113-125, p. 117.
[341] See recitals 71-73 and Art. 53-54.

development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan under regulatory supervision".

In the typology by the Datasphere Initiative, sandboxes are divided in two categories: the operational sandboxes, which are secure, collaborative data-spaces that pool data-sets and resources together, and regulatory sandboxes, that are time-limited collaborative endeavours involving regulators, service-providers and other relevant stakeholders to test innovative technology and data practices against regulatory frameworks[342]. OECD has defined regulatory sandboxes as "a *limited form of regulatory waiver or flexibility for firms, enabling them to test new business models with reduced regulatory requirements*[343]".

This **hybrid form of lax regulation and monitored innovation could be the most suitable semi-controlled environment for AI development.**

A Regulatory Sandbox

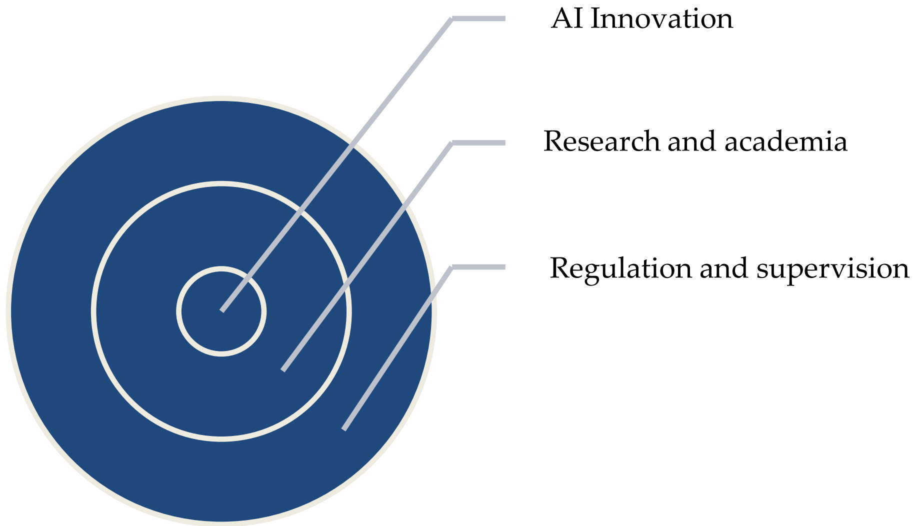


Figure 19

Created by the author

[342] Datasphere Initiative (2022). *Sandboxes for data: creating spaces for agile solutions across borders*, p. 10.
[343] Attrey, A., Lesher, M., Lomax, C. (2020), *The role of sandboxes in promoting flexibility and innovation in the digital age,* Going Digital Toolkit Note, No. 2, p. 7.

The benefits of the regulatory sandboxes' framework, which is still under development and in the early stages of universal adoption, include the following:

a) **Increased understanding of the complex technical innovations, in favour of regulation efficiency**. According to a 2020 report by the European Bank for Reconstruction and Development, 73 unique sandboxes in 57 countries were identified in the fintech sector[344], which offered further understanding of fintech innovations to policymakers in order to increase their capacity to regulate them. Hence, more adequate decision-making and supervision policies can be achieved. Where the legal situation is unclear or meaningful legislation has yet to be created, they can provide a stopgap solution to determine and impose the relevant rules in real time[345].

b) **Cooperation and exchange of empirical evidence and opinions** between the innovators and the regulators, which facilitates more efficient policy development by the regulators and better understanding of the policymakers' expectations by the innovators.

c) **A favourable environment for innovators** is created, which allows for testing of new technologies without having to meet all regulatory requirements normally applicable in a specific area[346].

d) **Strengthening of competition,** as smaller private entities can enter the sandbox environment and provide valuable insights in an equal manner to much bigger enterprises.

---

[344] International Bank for Reconstruction and Development. (2020). *Global Experiences from Regulatory Sandboxes*, at: https://documents1.worldbank.org/curated/en/912001605241080935/pdf/Global-Experiences-from-Regulatory-Sandboxes.pdf
[345] Datasphere Initiative (2022). *Sandboxes for data: creating spaces for agile solutions across borders*. p. 22.
[346] Madiega, T., Van De Pol, A.L. (2022). *Artificial intelligence act and regulatory sandboxes*, European Parliamentary Research Service, p. 3.

AI sandboxes are already being implemented, with the first presented in Spain[347] and more followed, for example Norway's Data Protection Authority's regulatory sandbox for AI[348]. As regards the Twin Transitions, the mobility sector and the biodiversity genomics field have been identified as potential spaces for cross-border regulatory sandboxes creation. Concerning the first, it suggests that cross-border sandboxes could test some of the emerging solutions in both mobility and related data (such as GHG emissions) and advance their use if their benefits are proved scalable[349]; as regards the second, regulatory sandboxes could become a testing space for the creation of a robust and inclusive biodiversity data-sharing and management ecosystem[350].

To regulators	To firms	To consumers
• Inform long-term policy making through learning and experimentation • Signal commitment to innovation and learning • Promote communication and engagement with market participants • Update regulations that might prohibit beneficial innovation	• Reduce time to market by streamlining the authorisation process • Reduce regulatory uncertainty, such as that new technologies and business models will be prohibited • Gather feedback on regulatory requirements and risks • Improve access to capital • Remove market-entry barriers for companies (especially SMEs and start-ups) by democratising the knowledge about legal frameworks around certain innovative products	• Promote introduction of new and potentially safer products • Increase access to financial products and services

Figure 20

The benefits of Regulatory Sandboxes

OECD (2023), Regulatory sandboxes in artificial intelligence, OECD Digital Economy Papers, No. 356, OECD Publishing, Paris, https://doi.org/10.1787/8f80a0e6-en, p. 15, as adapted from Parenti, R. (2020) *Regulatory Sandboxes and Innovation Hubs for FinTech*, Study for the committee on Economic and Monetary Affairs

347 European Commission. (2022). *First regulatory sandbox on Artificial Intelligence presented*, at: https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented

348 See more at the Authority's website, at: https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/.

349 Datasphere Initiative (2022). Sandboxes for data: creating spaces for agile solutions across borders, p. 39-40.

350 *Id*., at p. 44.

As regards the AI Act up until the text adopted by the European Parliament in June 2023, AI sandbox establishment in EU Member States is deemed as "*a next step to be made mandatory with established criteria* (…) *to ensure a legal framework that promotes innovation, is future-proof, and resilient to disruption*". More specifically, the AI Act mentions that the establishment of regulatory sandboxes aims to increase the authorities' understanding of technical developments, to improve supervisory methods and provide guidance to AI systems developers and providers to achieve regulatory compliance with the AI Act or where relevant, other applicable Union and Member States legislation, like the GDPR, as well as with the Charter of Fundamental Rights. This is particularly important, as the Act introduces both a regulatory compliance framework as well as an ethical and human rights dimension. Furthermore, this ecosystem should allow for identification of risks during the development and testing phase of AI and its suspension in cases of non-conformity with already established Law.

Another major takeaway from the AI Act is that, in order to avoid fragmentation, the establishment of **uniform rules** for the regulatory sandboxes' implementation and a framework for cooperation between the relevant authorities involved in their supervision is provided for; to this end, the new Art. 53a authorises the European Commission, in consultation with the to-be-established "AI office", to "*adopt* ***delegated acts*** *detailing the modalities for the establishment, development, implementation, functioning and supervision of the AI regulatory sandboxes, including the eligibility criteria and the procedure for the application, selection, participation and exiting from the sandbox, and the rights and obligations of the participants* (…)". According to Art. 290 TFEU, the European legislator may grant the power to the Commission to adopt non-legislative acts of general application ("delegated acts") to supplement or amend certain non-essential elements of the legislative act (here, the AI Act). So, the European Commission will essentially set the specific rules for the entirety of the regulatory sandboxes' functioning on EU-wide level.

It should be noted that according to Art. 53 (4), prospective sandbox providers in the AI regulatory sandbox are not exempt from **liability** under applicable Union and Member States liability legislation for any harm inflicted on third parties as a result of the experimentation taking place in the sandbox. Furthermore, an interesting provision is the new 1f paragraph of Art. 53, which establishes a presumption of conformity with the Act's provisions that were assessed within the sandbox concerning AI systems that successfully exit it; this means that insofar as the AI system complies with the requirements that will be set when exiting the sandbox, it shall be presumed to be in conformity with the AI Act. However, according to EURACTIV's L. Bertucci, the Spanish presidency highlights some concerns with this process, including losing control over the compliance process and causing negative impact on competition[351].

Meanwhile, Open Loop, a global programme connecting policymakers and companies in the field of AI, has set certain conditions of a successful AI regulatory sandbox[352], which include:

a) the generalisation of its findings for the benefit of society and for the sake of transparency,
b) the establishment of clear roles and responsibilities, objective entry criteria and exit requirements for participants (especially as regards conformity testing),
c) the provision of attractive benefits for participants in order to ensure participation and,
d) of course, the central to the concept of sandboxes "regulatory leeway" in the sense of flexibility in terms of regulation, like temporary legal exemptions or experimentation clauses[353].

As the new AI Act recital 28a provides, "*the fundamental right to a high level of environmental protection enshrined in the Charter and implemented in Union policies should also be considered when assessing the severity of the harm that an AI system can cause*". Hence, the **sustainability dimension** should also be considered as regards the above; for example, one of the sandbox's functioning criteria should be that no significant environmental damage occurs through the development and/or testing of the AI system. If so occurs, then the AI system should not pass the sandbox's conformity assessment when exiting it. Therefore, certain pre-established semi(?)-binding mechanisms that are activated inside a sandbox's operational

351 Bertucci, L. (2023). EU Council sets path for innovation measures in AI Act's negotiations, at: https://www.euractiv.com/section/artificial-intelligence/news/eu-council-sets-path-for-innovation-measures-in-ai-acts-negotiations/
352 Norberto Nuno de, A., Galindo, L., Zarra, A. (2023). *Artificial Intelligence Act: A Policy Prototyping Experiment: Regulatory Sandboxes*, at: https://openloop.org/reports/2023/04/Artificial_Intelligence_Act_A_Policy_Prototyping_Experiment_Regulatory_Sandboxes.pdf, p. 25-28.
353 German Federal Ministry for Economic Affairs and Energy. (2019). *Making space for innovation: The handbook for regulatory sandboxes*, at: https://www.bmwk.de/Redaktion/EN/Publikationen/Digitale-Welt/handbook-regulatory-sandboxes.pdf%3F__blob%3DpublicationFile%26v%3D2, p. 7.

phase, like an environmental conformity assessment, may be useful in order to secure trust in the sandbox's functions and overall effectiveness in the contemporary legal acquis.

Of course, regulatory sandboxes are a relatively new and underused tool, at least as regards AI[354]. However, their unique hybrid nature could perhaps be the most fitting solution for the challenging AI field, on condition that the thornier issues like accountability, transparency and the important ethical implications of AI use are clarified in advance. After all, certain values like human dignity and equality are too important to be confined to unenforceable, soft frameworks[355].

So, the author's opinion is that regulatory sandboxes are not "just a hype[356]" and **should be utilised in the greatest extent possible, provided that clear and universal** (at least in the EU) **rules as regards their establishment, development, implementation, functioning and supervision are set, along with accompanying concrete ethical provisions and** non-discouraging **accountability safeguards.**

4.3. The precautionary principle and AI

As primarily enshrined in International Law[357], the precautionary principle dictates that **preventive action** is a necessary condition for the effective protection of environmental assets[358]. Using the same

[354] On more on regulatory sandboxes and AI, see the recent report by OECD (2023), *Regulatory sandboxes in artificial intelligence*, OECD Digital Economy Papers, No. 356, OECD Publishing, Paris, https://doi.org/10.1787/8f80a0e6-en.
[355] Varošanec, I., (2022). On the Path to the Future: Mapping the Notion of Transparency in the EU Regulatory Framework for AI. *International Review of Law, Computers & Technology Vol*. 36(2), University of Groningen Faculty of Law Research Paper No. 18/2022, http://dx.doi.org/10.2139/ssrn.4066020, p. 18.
[356] See Moraes, T. (2023). *Regulatory sandboxes for Artificial Intelligence – hype or solution?* KU Leuven AI Summer School blogspot, at: https://www.law.kuleuven.be/ai-summer-school/blogpost/Blogposts/regulatory-sandboxes.
[357] United Nations, 1992 Rio Declaration on Environment and Development, UN Doc. a/ conf.151/26 (vol. I), 31 ilm 874 (1992). Principle number 15 mentions the precautionary principle with a view to protecting the natural environment.
[358] Kl. Pouikli, I. Tsakalogianni (2022). Criteria for identifying splitting of projects (salami-slicing) tactic in light of the C-11/20 Regione Puglia case and EIA practice in Greece, *Journal for European Environmental & Planning Law*, 19(3), 222-247. doi: https://doi.org/10.1163/18760104-19030005, p. 223.

rationale in the field of AI, preventative measures should be adopted first in order to address potential risks of technological innovation with prudent decision-making.

The precautionary principle, despite its significance, is not a static concept, and its interpretation varies across different jurisdictions and sectors. Some argue for a more rigid application, advocating for stringent regulations in the face of uncertain risks, while others propose a more flexible approach focusing on adaptive risk management.

In AI talk, governance and even Law, the precautionary principle is amiss; that is maybe due to the critique that excessive precaution might lead to missed opportunities and overregulation, which, especially in the technological field, may do great damage to progress or lead to extreme delays. In any case, significant values like AI transparency and explainability and the most fundamental ethical standards should be imposed ***ex ante***, while carefully weighing and balancing the need to empower technological innovation without risking creating a regulatory gap. **As regards the Twin Transitions, obligations like environmental monitoring of AI technologies and setting environmental standards for the licensing or design of carbon-intensive AI procedures** should be set, in order to prevent any negative environmental impacts.

4.4. Environmental Sustainability Mainstreaming and cross-sectoral consistency

After taking the above regarding regulating AI into account, it also seems essential to underline the importance of integrating the environmental sustainability factor into AI governance, irrespectively of the shape or form it assumes. As analysed in section 3 of the present thesis, the Twin Transition is already shaping our society and economy, so the governance system of AI will be largely incomplete if the climate – or more broadly, the environmental factor – is not embedded into it.

In general, AI tools can be integrated into the steps of policymaking process, as depicted below.

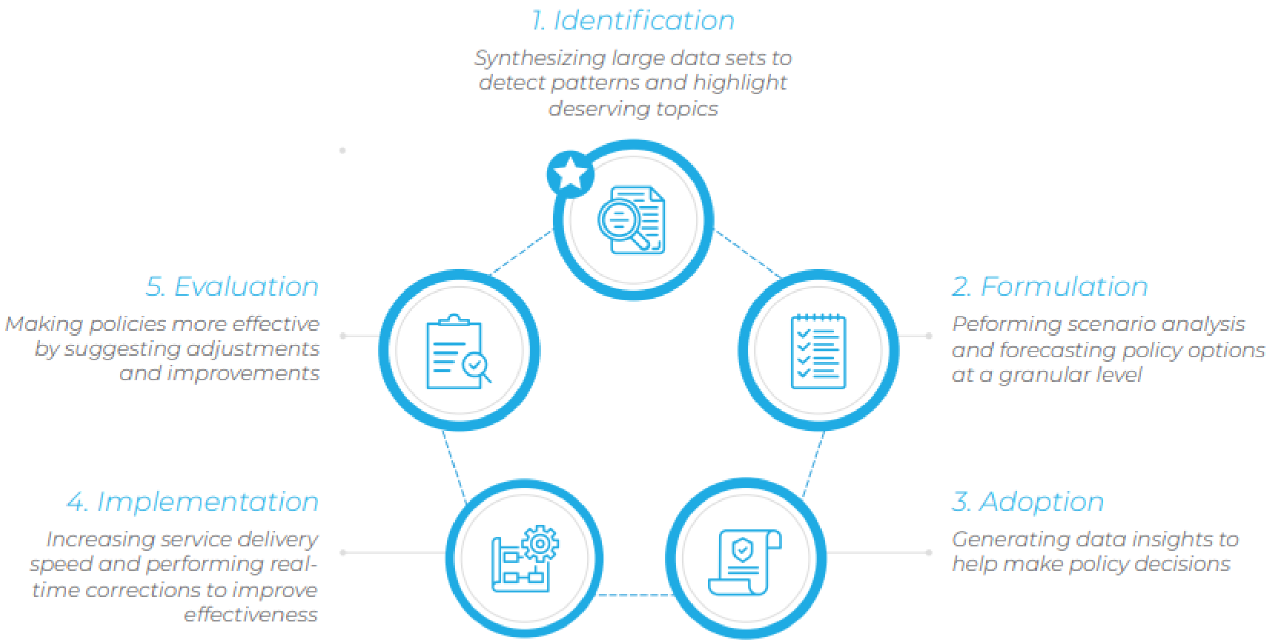


Figure 21
The role of AI in the policymaking cycle
GPAI (2022). *Biodiversity & Artificial Intelligence, Opportunities and Recommendations*, Report, November 2022, Global Partnership on AI, p. 52.

The challenge in utilising AI for policymaking is integrating this new technology into the policymaking against pressing environmental threats. In this vein, **Environmental Sustainability Mainstreaming** is regarded as the inclusion of the environmental factor, e.g., environmental protection, ecosystems' restoration, etc. as a top priority of systemic policy. It encompasses the idea of integrating environmental priorities into various sectors' programmes, such as reducing GHG emissions and biodiversity protection, which become cross-cutting concerns that affect economic development and social well-being and which require a synergistic implementation of actions that collectively contribute to sustainable development. By not treating environmental sustainability as a standalone issue, the risk of having various policies at odds with each other is greatly diminished. By ensuring that binding environmental obligations are set in agriculture, energy, transportation, urban planning, new technologies, etc. legislation, regulatory fragmentation and gaps in implementation can be reduced.

Several enabling factors are integral to the successful integration of sustainability considerations into the majority of public policies, which include public engagement, academia insight, access to related data and ensuring transparency, as well as the existence of robust legal frameworks that promote cross-sectoral consistency.

Of course, the above necessitate building political will and commitment at all levels, from national governments to local authorities, to prioritise environmental action in national, regional and local policies. At the same time, institutional arrangements like dedicated sustainability units or focal points in both the public and private sector are important, in order to coordinate and facilitate the integration of sustainability considerations vertically and horizontally.

Furthermore, the **must-needed collaboration** between actors through the promotion and exchange of innovative ideas, resources, and knowledge to address environmental challenges effectively is central to the concept of environmental sustainability mainstreaming and intrinsic to the success of the relationship between environmental and AI policy. As Miguel Luengo-Oroz, AI strategist and chief data scientist at the United Nations Global Pulse, after having attended both the Conference on Neural Information Processing Systems (NeurIPS) and COP 25 in December 2019, put it, "*I don't know if anyone else was involved in both… We* [spoke] *about AI and sustainability at COP25, together with other new emerging technologies. And then at NeurIPS, where there were a lot of researchers, they met few climate experts… We need real experts; experts who deeply understand both sides of the game*[359]." This **interdisciplinary aspect** is crucial in the governance of the Twin Transitions.

The mosaic of connections and dependencies created by AI will likely force a radical change—both cultural and operational—in governance and oversight[360]. AI-supported platforms and tools like sandboxes can encourage stakeholder engagement and enhance collaboration between private entities, scientists, academia, policymakers, and the public. By providing data-driven insights and facilitating

---

[359] Dhar, P. (2020). The carbon impact of artificial intelligence. *Nat Mach Intell* 2, 423–425, https://doi.org/10.1038/s42256-020-0219-9.
[360] Buczynski, W., Steffek, F., Cuzzolin, F., Jamnik, M., & Sahakian, B. (2022). Hard Law and Soft Law Regulations of Artificial Intelligence in Investment Management. *Cambridge Yearbook of European Legal Studies*, 24, 262-293. doi:10.1017/cel.2022.10, p. 288.

knowledge-sharing, AI can enable more inclusive and participatory decision-making processes that efficiently combine AI and the sustainability factor in a cross-sectoral manner.

In the end, **ensuring that environmental action is no longer confined to a niche area, but is instead embedded in every decision made at all levels of governance, and AI technologies as a whole, can be they key for ensuring a smooth Twin Transition.**

The observations of Section 4 are depicted below:

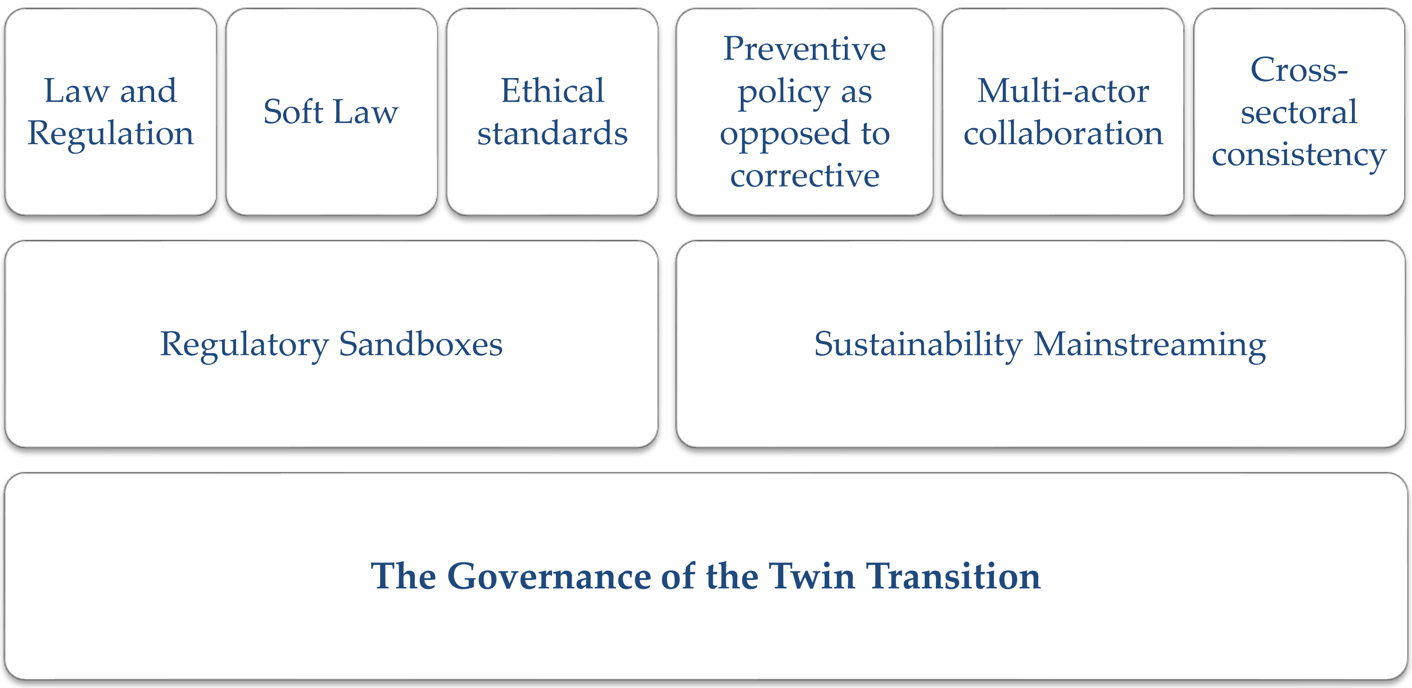


Figure 22

A proposal for a governance framework for the Twin Transition

Created by the author

Conclusions

Climate change has emerged as one of the most significant challenges that humanity faces in the 21st century; similarly, AI's complex, multifaceted and ever-evolving nature is already pushing beyond the boundaries of governance and regulation worldwide. AI can radically transform the bulk of economy and society: from energy, climate science, transportation, industry and urban planning to the food system and biodiversity protection and conservation, the possibilities are seemingly endless.

In Europe, at the heart of climate ambition, a collective, equitable, green and digital transition in order to achieve climate neutrality is one of the pinnacles of its policy[361]. This transitional intersection, the **Twin Transition**, demands innovative and comprehensive solutions that ease the clashes and smooth the edges of the interaction between the AI technological boom and the green transition. The synergy between AI and climate change provides us with a powerful arsenal to combat the existential threat of climate change, as AI is an attractive new field that offers unprecedented opportunities to industry, economy, society and each and every one of us.

AI is rapidly growing in development, usage, and, admittedly, in popularity as well. However, there is a real risk that the immense potential of AI becomes a double-edged sword, as its training, deployment and utilisation left unregulated and unsupervised, and society finds itself in a situation where a technology surpasses the ability of the Law to govern it[362]. As an invention of the human mind, AI also carries the negative features of its creator; the obvious advantages should not distract from the risks AI carries, from ethical concerns, human rights violations and systemic biases to legal "grey areas" and negative environmental impacts.

[361] COM (2020) 442, p. 10.

[362] Gutierrez, C.I. (2020), *Can Existing Laws Cope with the AI Revolution?* Brookings, at: https://www.brookings.edu/techstream/can-existing-laws-cope-with-the-ai-revolution.

In its current state, AI has already disrupted the legal ecosystem; as it becomes increasingly, and seamlessly, incorporated into business, society, and our personal lives, we can also expect a new regulatory climate to take shape[363]. The AI Act, the European Regulation on AI, the first in the world aiming to create binding rules for AI development and use, is an aspiring legal attempt and, when formally adopted, it will be added to the GDPR, the Digital Services Act[364], the Digital Markets Act[365] and the Data Governance Act[366]. As this "legislative pentagon" is completed, the EU will have developed a "digital constitutionalism" for an infosphere where its citizens may live and work better and more sustainably[367].

However, the challenging and unique nature of AI demands the adoption of new forms of legal devices and the familiarisation of the regulators and policy-makers with the intricacies of AI systems. In the words of Joseph Weizenbaum, in the artificial realms "*machines are made to behave in wondrous ways, often sufficient to dazzle even the most experienced observer. But once a particular program is unmasked, once its inner workings are explained in language sufficiently plain to induce understanding, its magic crumbles away; it stands revealed as a mere collection of procedures, each quite comprehensible*[368]". AI should not be seen as an indecipherable technological enigma, but rather as a product of human intellect.

To this end, instruments like regulatory sandboxes can provide a safe and comfortable environment for the development and testing of AI systems, provided that uniform rules are set that deal with monitoring mechanisms, accountability provisions, ethical safeguards, sustainability conditions and entry and exit prerequisites. By creating a controlled space for exchange of opinions, ideas and concerns between multiple actors of the public and private sector, scholars, innovators and the public, society and

---

363 McKinsey & Company (2023). *What is generative AI?* at: https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai.
364 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act).
365 Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act).
366 Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).
367 Floridi, L. The European Legislation on AI: a Brief Analysis of its Philosophical Approach. *Philos. Technol*. 34, 215–222 (2021). https://doi.org/10.1007/s13347-021-00460-9
368 *Supra* 27, at p. 36.

the environment can be protected from the potential risks of AI, while capturing its full potential in a responsible, ethical and regulatory monitored way. Fostering and promoting real awareness of the extent and causes of the problematic facets of AI, at the same time as ensuring interdisciplinarity and inclusiveness is one of the key aspects of AI governance, in order to not have policy and industry, or the Law and innovators clash.

Thomas L. Friedman called our era as "a new Promethean Moment". This providential moment is not fuelled by a specific discovery, as was the printing press or the steam engine, but by a technological grand cycle[369]. Today, amidst concerns about AI's unpredictable nature and potential harmful outcomes, its future shines with remarkable brightness. At the same time, according to the IPCC, there is a rapidly closing window of opportunity to secure a liveable and sustainable future for all, as the climate-policy choices and climate actions implemented in this decade will have impacts now and for thousands of years[370]. In the AI and climate change "marriage", in this **Twin Transition**, some risks, such as AI's carbon footprint, are not entirely avoidable, but they can certainly be minimised, to deliver the best strategies against climate change[371].

The global ever-evolving research and the timely adaptation of the legislation trying to keep up with it can offer an important opportunity to create a truly ethical, just and sustainable AI. In the end, **they way that we approach AI, much like climate change, will most definitely determine the world we want to (and will) live in.**

[369] Friedman, T. (2023). Our New Promethean Moment, New York Times, at: https://www.nytimes.com/2023/03/21/opinion/artificial-intelligence-chatgpt.html
[370] IPCC (2023) Sixth Assessment Report, Synthesis Report (SYR), *Summary for Policymakers*, p. 25.
[371] Cowls, J., Tsamados, A., Taddeo, M. et al. The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Soc 38*, 283–307 (2023). https://doi.org/10.1007/s00146-021-01294-x

Bibliography

Adadi, A., Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), *IEEE Access*, 2018, Vol. 6, p. 52138.

Abbott, K. W., Snidal, D. (2000). *Hard and Soft Law in International Governance*, International Organization 54:421-456.

AI HLEG. (2019). *A Definition of AI: Main Capabilities and Disciplines*.

Ali, A.M., Söffker D. (2018). Towards optimal power management of hybrid electric vehicles in real-time: A review on methods, challenges, and state-of-the-art solutions. *Energies*, 11(3).

Amodei, D., Hernandez, D. (2018) *AI and Compute*, at: https://openai.com/blog/ai-and-compute.

Anders, A. (2017). *Total Consumer Power Consumption Forecast*, Nordic Digital Business Summit Conference.

Attrey, A., Lesher, M., Lomax, C. (2020), *The role of sandboxes in promoting flexibility and innovation in the digital age,* Going Digital Toolkit Note, No. 2.

Balayn, A., Gürses, S. (2021). Beyond *Debiasing: Regulating AI and its inequalities*, European Digital Rights.

Bar, G. (2021) *Robot personhood – what is anthropocentric artificial intelligence good for*? in: Lai, L., Świerczyński, M. (eds.), Legal and Technical aspects of Artificial Intelligence, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw.

Barnes, E.A., Hurrell, J.W., Ebert Uphoff, I., Anderson, C., & Anderson, D. (2019). Viewing forced climate patterns through an AI Lens. *Geophysical Research Letters*, 46, 13,389–13,398, https://doi.org/10.1029/2019GL084944.

BCG Gamma - AI for the Planet (2022) *How AI Can Be a Powerful Tool in the Fight Against Climate* Change.

Bellamy, R. et al. (2018) *AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias,* https://doi.org/10.48550/arXiv.1810.01943.

Bergengruen V., Perrigo, B. for The Times (2021) *Facebook Acted Too Late to Tackle Misinformation on 2020 Election, Report Finds*, at: https://time.com/5949210/facebook-misinformation-2020-election-report/.

Bertucci, L. (2023). EU Council sets path for innovation measures in AI Act's negotiations, at: https://www.euractiv.com/section/artificial-intelligence/news/eu-council-sets-path-for-innovation-measures-in-ai-acts-negotiations/

Bertucci, L. for EURACTIV (2023). *EU lawmakers set to settle on OECD definition for Artificial Intelligence*, at: https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/?utm_source=substack&utm_medium=email

Bieger, J., Thórisson, K. R., Steunebrink, B. R., Thorarensen, T., and Sigurdardottir, J. S. (2016). *Evaluation of general-purpose artificial intelligence: why, what & how*, Evaluating General-Purpose AI.

Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts, 44th Parl., 1st Sess., 70-71 Elizabeth II, 2021-2022 (First Reading), at: https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading.

Bitterman, D., Aerts, H. & Mak, R. (2020). Approaching autonomy in medical artificial intelligence. *The Lancet Digital Health*. 2. e447-e449. 10.1016/S2589-7500(20)30187-4.

Boch, A., Hohma, E., Trauth, R. (2022). Towards an Accountability Framework for AI: Ethical and Legal Considerations, Technical University of Munich, Center for Technology in Society, Institute for Ethics in Artificial Intelligence, at: https://ieai.sot.tum.de/wp-content/uploads/2022/03/ResearchBrief_March_Boch_Hohma_Trauth_FINAL_V2.pdf.

Bommasani, R. et al. (2021) *On the Opportunities and Risks of Foundation Models*, arXiv:2108.07258v3, https://doi.org/10.48550/arXiv.2108.07258.

Bordt, S., Finck, M., Raidl, E., von Luxburg, U. (2022) Post-Hoc Explanations Fail to Achieve their Purpose in Adversarial Contexts, FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, p. 891–905, doi: 10.1145/3531146.3533153

Boucher, P. (2020) "*Artificial intelligence: How does it work, why does it matter, and what can we do about it*?", Panel for the Future of Science and Technology EPRS, European Parliamentary Research Service.

Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press.

Brittain, B. for Reuters (2023) *AI-created images lose U.S. copyrights in test for new technology*, at: https://www.reuters.com/legal/ai-created-images-lose-us-copyrights-test-new-technology-2023-02-22/.

Buczynski, W., Steffek, F., Cuzzolin, F., Jamnik, M., & Sahakian, B. (2022). Hard Law and Soft Law Regulations of Artificial Intelligence in Investment Management. *Cambridge Yearbook of European Legal Studies*, 24, 262-293. doi:10.1017/cel.2022.10.

Buiten, M. (2019). Towards Intelligent Regulation of Artificial Intelligence. *European Journal of Risk Regulation*, 10(1), 41-59. doi:10.1017/err.2019.8.

Burgard, W. (2022). Artificial Intelligence: Key Technologies and Opportunities. In S. Voeneky, P. Kellmeyer, O. Mueller, & W. Burgard (Eds.), The Cambridge Handbook of Responsible Artificial Intelligence: Interdisciplinary Perspectives (Cambridge Law Handbooks, pp. 11-18). Cambridge: Cambridge University Press. doi:10.1017/9781009207898.003.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Societ*y, 3(1). https://doi.org/10.1177/2053951715622512.

Cappello M. (2020), *Artificial intelligence in the audiovisual sector*, IRIS Special, European Audiovisual Observatory, Strasbourg.

Catlett, C., Cesario, E., Talia, D., Vinci, A. (2019). Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments, *Pervasive and Mobile Computing*, Volume 53, https://doi.org/10.1016/j.pmcj.2019.01.003.

Chesterman, S. (2020). Artificial Intelligence and the Limits of Legal Personality. *International & Comparative Law Quarterly*, 69(4), 819-844. doi: 10.1017/S0020589320000366.

CIO (2022) *The AI Sustainability Paradox – And How to Solve It*, at: https://www.cio.com/article/415787/the-ai-sustainability-paradox-and-how-to-solve-it.html.

Collingridge, D. (1980). *The Social Control of Technology*, New York: St. Martin's Press; London: Pinter.

COM (2020) 442, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS The EU budget powering the recovery plan for Europe.

COM (2022) 495 final 2022/0302 (COD) Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on liability for defective products.

COM (2022) 496 final, 2022/0303 (COD) Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL, 2022 Strategic Foresight Report, Twinning the green and digital transitions in the new geopolitical context, COM (2022) 289 final.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS The European Green Deal, COM/2019/640 final

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, *A New Industrial Strategy for Europe*, COM/2020/102 final.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, A New Circular Economy Action Plan for a Cleaner and more Competitive Europe, COM/2020/98 Final.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS - A New Industrial Strategy for Europe, 10.3.2020, COM(2020) 102 final. On more on the EU's circular economy policy see

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS *Closing the loop - An EU action plan for the Circular Economy,* COM/2015/0614 final.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS, *A New Circular Economy Action Plan for a Cleaner and more Competitive Europe,* COM/2020/98 Final.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS EU, *Strategy for Sustainable and Circular Textiles*, COM/2022/141 final.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS 'Fit for 55': delivering the EU's 2030 Climate Target on the way to climate neutrality, COM(2021) 550 final.

Conference of the Parties, Adoption of the Paris Agreement, Dec. 12, 2015. U.N. Doc. FCCC/CP/2015/L.9/Rev/1 (Dec. 12, 2015).

Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products

Council of Europe, European Commission for the Efficiency of Justice (2018) "European Ethical Charter on the Use of AI in Judicial Systems and their environment", as adopted at the 31st plenary meeting of the Commission.

Council of Europe. (1950). Convention for the Protection of Human Rights and Fundamental Freedoms. In Council of Europe Treaty Series 005. Council of Europe.

Cowls, J., Tsamados, A., Taddeo, M. et al. (2023). The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Society* 38, 283–307. https://doi.org/10.1007/s00146-021-01294-x.

Daly, A., Hagendorff, T., Hui, L., Mann, M., Marda, V., Wagner, B., & Wang, W. (2021). *AI, Governance and Ethics: Global Perspectives.* In H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, & G. De Gregorio (Eds.), Constitutional Challenges in the Algorithmic Society (pp. 182-201). Cambridge: Cambridge University Press. doi:10.1017/9781108914857.010.

Datasphere Initiative (2022). *Sandboxes for data: creating spaces for agile solutions across borders*.

Datatilsynet – The Norwegian Data Protection Authority (2018) *Artificial intelligence and privacy Report*, at: https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

DelSole, T., Banerjee, A. (2017) Statistical seasonal prediction based on regularized regression. *J. Clim.* 30(4), 1345–1361.

Dhar, P. (2020). The carbon impact of artificial intelligence. *Nat Mach Intell* 2, 423–425, https://doi.org/10.1038/s42256-020-0219-9.

Directive 2011/65/EU of the European Parliament and of the Council of 8 June 2011 on the restriction of the use of certain hazardous substances in electrical and electronic equipment.

Directive 2011/92/EU of the European Parliament and of the Council of 13 December 2011 on the assessment of the effects of certain public and private projects on the environment (codification) and Directive 2001/42/EC of the European Parliament and of the Council of 27 June 2001 on the assessment of the effects of certain plans and programmes on the environment.

Directive 2012/19/EU of the European Parliament and of the Council of 4 July 2012 on waste electrical and electronic equipment (WEEE) (recast)

Domingues, T., Brandão, T., & Ferreira, J. C. (2022). Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey. *Agriculture*, 12(9), 1350. https://doi.org/10.3390/agriculture12091350

Downes, L. (2009). *The Laws of Disruption: Harnessing the New Forces that Govern Life and Business in the Digital Age*, Basic Books.

Dymitruk, M. (2021) *The role of XAI in automated decision-making in the judiciary,* in: Lai, L., Świerczyński, M. (eds.), Legal and Technical aspects of Artificial Intelligence, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw.

EEA Report (2020). *Bio-waste in Europe — turning challenges into opportunities*, available at: https://www.eea.europa.eu/publications/bio-waste-in-europe

Ellen MacArthur Foundation. (2019). *Artificial intelligence and the circular economy - AI as a tool to accelerate the transition*, at: http://www.ellenmacarthurfoundation.org/publications.

Euronews. (2020). *Social scoring: Could that Facebook post stop you getting a loan or a mortgage?,* at: https://www.euronews.com/2020/06/24/social-scoring-could-that-facebook-post-stop-you-getting-a-loan-or-a-mortgage

European Commission, *Shaping Europe's Digital Future*, at: https://commission.europa.eu/system/files/2020-02/communication-shaping-europes-digital-future-feb2020_en_4.pdf.

European Commission. (2022). *First regulatory sandbox on Artificial Intelligence presented*, at: https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented.

European Commission. *Smart cities: Cities using technological solutions to improve the management and efficiency of the urban environment*, at: https://commission.europa.eu/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en.

European Environment Agency. (2016). *Circular Economy in Europe – Developing the knowledge base*, EEA Report.

European Environmental Agency. (2022). *Textiles and the environment: the role of design in Europe's circular economy*, at: https://www.eea.europa.eu/publications/textiles-and-the-environment-the.

European Parliament Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) (2018/C 252/25).

European Parliament, Policy Department for Economic, Scientific and Quality of Life Policies (2021) *Artificial Intelligence in smart cities and urban mobility.*

European Parliament. (2023). *Draft European Parliament Legislative Resolution 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)), https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_EN.html#_section1.

European Parliamentary Research Service, Panel for the Future of Science and Technology (2023). *Social approach to the transition to smart cities*, p. 41 and Annex VI for relevant EU legislation.

Fazackerley, A. (2023), *AI makes plagiarism harder to detect, argue academics – in paper written by chatbot*, The Guardian, at: https://www.theguardian.com/technology/2023/mar/19/ai-makes-plagiarism-harder-to-detect-argue-academics-in-paper-written-by-chatbot.

Federal Ministry of Transport and Digital Infrastructure, Ethics Commission, *Automated and Connected Driving*, at: https://bmdv.bund.de/SharedDocs/EN/publications/report-ethics-commission-automated-and-connected-driving.pdf?__blob=publicationFile.

Floridi, L. (2021). The European Legislation on AI: a Brief Analysis of its Philosophical Approach. *Philosophy & Technology*. 34, 215–222, https://doi.org/10.1007/s13347-021-00460-9.

Floridi, L., Nobre, A.-C. (2020) *The Green and the Blue: How AI may be a force for good*, OECD Forum, available at: https://www.oecd-forum.org/posts/the-green-and-the-blue-how-ai-may-be-a-force-for-good.

Fontes, C., Hohma, E., Corrigan, C., Lütge, C. (2022), AI-powered public surveillance systems: why we (might) need them and how we want them, *Technology in Society*, Volume 71, 102137, https://doi.org/10.1016/j.techsoc.2022.102137.

Forti, V. et al. (2020) *The Global E-waste Monitor 2020: Quantities, flows, and the circular economy potential,* at: https://collections.unu.edu/view/UNU:7737#viewAttachments.

Friedman, B., Nissenbaum (1996). H. Bias in Computer Systems, *ACM Transactions on Information Systems,* Volume 14 Issue 3.

Friedman, T. (2023). Our New Promethean Moment, New York Times, at: https://www.nytimes.com/2023/03/21/opinion/artificial-intelligence-chatgpt.html

Future of Life Institute. (May 2022) *General Purpose AI and the AI Act.*

Gailhofer, P. et al. (2021). *The role of Artificial Intelligence in the European Green Deal*, Study for the special committee on Artificial Intelligence in a Digital Age (AIDA), Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg.

Gasser, U., Virgilio, A. (2017). A Layered Model for AI Governance. *IEEE Internet Computing* 21 (6) (November): 58–62. doi:10.1109/mic.2017.4180835.

Genus, A., Stirling, A. (2018). Collingridge and the dilemma of control: Towards responsible and accountable innovation. *Research Policy*, Volume 47, Issue 1, doi: 10.1016/j.respol.2017.09.012,

German Federal Ministry for Economic Affairs and Energy. (2019). *Making space for innovation: The handbook for regulatory sandboxes,* at: https://www.bmwk.de/Redaktion/EN/Publikationen/Digitale-Welt/handbook-regulatory-sandboxes.pdf%3F__blob%3DpublicationFile%26v%3D2.

Gersen, Jacob E. and Psoner, Eric A. (2008). Soft Law: Lessons from Congressional Practice. Stanford Law Review 61:573-627.

Ghahramani A, Galicia P, Lehrer D, Varghese Z, Wang Z and Pandit Y. (2020). Artificial Intelligence for Efficient Thermal Comfort Systems: Requirements, Current Applications and Future Directions. *Front. Built Environ.* 6:49. doi: 10.3389/fbuil.2020.00049.

Ghosh, P., & P. Kumpatla, S. (2022). *GIS Applications in Agriculture.* IntechOpen. doi: 10.5772/intechopen.104786.

Global Partnership on AI Report in collaboration with Climate Change AI and the Centre for AI & Climate. (2021). *Climate Change and AI: Recommendations for Government Action*.

Gómez Mont C. (2021). *The Intersection between AI and the Climate and Biodiversity Crisis*, AI Ethics: Global Perspectives.

Google (2023). *Google Data Centers efficiency: How we do it*., at: https://www.google.com/about/datacenters/efficiency/internal/.

GPAI (2022). *Biodiversity & Artificial Intelligence, Opportunities and Recommendations*, Report, Global Partnership on AI.

Gröger, J, Liu, R., Stobbe, L., Druschke, J., Richter, N. (2021), *Green Cloud Computing, Lebenszyklusbasierte Datenerhebung zu Umweltwirkungen des Cloud Computing,* Fraunhofer-Institut für Zuverlässigkeit und Mikrointegration, available at: https://www.umweltbundesamt.de/sites/default/files/medien/5750/publikationen/2021-06-17_texte_94-2021_green-cloud-computing.pdf.

Guinée, J.B.; Heijungs, R.; Huppes, G.; Zamagni, A.; Masoni, P.; Buonamici, R.; Ekvall, T.; Rydberg, T. (2011). Life Cycle Assessment: Past, Present, and Future. *Environ. Sci. Technol*. 45, 90–96.

Gupta, D., & Shah, M. (2022). A comprehensive study on artificial intelligence in oil and gas sector. Environmental science and pollution research international, 29(34), 50984–50997. https://doi.org/10.1007/s11356-021-15379-z.

Gutierrez, C. I., Marchant, G. E. (2021). *A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence,* http://dx.doi.org/10.2139/ssrn.3855171.

Gutierrez, C., Aguirre, A., Uuk, R., Boine, C., and Franklin M., *A Proposal for a Definition of General Purpose Artificial Intelligence Systems*, Future of Life Institute Working Paper, at: https://futureoflife.org/wp-content/uploads/2022/11/SSRN-id4238951-1.pdf

Gutierrez, C.I. (2020), *Can Existing Laws Cope with the AI Revolution?* Brookings, at: https://www.brookings.edu/techstream/can-existing-laws-cope-with-the-ai-revolution.

Gutierrez, I., C., Marchant, G. (2021). *Soft law 2.0: Incorporating incentives and implementation mechanisms into the governance of artificial intelligence*, https://oecd.ai/en/wonk/soft-law-2-0.

Hagemann, R., Huddleston, J. and Thierer, A. D. (2018). Soft Law for Hard Problems: The Governance of Emerging Technologies in an Uncertain Future. *Colorado Technology Law Journal*.

Hale, K. (2021) *A.I. Bias Caused 80% Of Black Mortgage Applicants To Be Denied, available at: https://www.forbes.com/sites/korihale/2021/09/02/ai-bias-caused-80-of-black-mortgage-applicants-to-be-denied/.*

Harb. M., Malik, J., Circella, G., Walker, J. (2022). Glimpse of the Future: Simulating Life with Personally Owned Autonomous Vehicles and Their Implications on Travel Behaviors. *Transportation Research Record*. 2022;2676(3):492-506. doi:10.1177/03611981211052543.

Hardman, S., Chakraborty, D., & Kohn, E. (2021). *A Quantitative Investigation into the Impact of Partially Automated Vehicles on Vehicle Miles Travelled in California*. UC Office of the President: University of California Institute of Transportation Studies. http://dx.doi.org/10.7922/G2XK8CVB

Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*, Cambridge, Mass: MIT Press.

Heaven D. W. (2021) *AI is learning how to create itself*, MIT Technology Review, at: https://www.technologyreview.com/2021/05/27/1025453/artificial-intelligence-learning-create-itself-agi/.

Hekler EB, Klasnja P, Chevance G, Golaszewski NM, Lewis D, Sim I. (2019). Why we need a structured data paradigm. *BMC Med.* 17:133. doi: 10.1186/s12916-019-1366-x.

Hertwich, E. et al. (2019). Material efficiency strategies to reducing greenhouse gas emissions associated with buildings, vehicles, and electronics—a review. *Environmental Research Letters* 14, 4 (2019), 043004.

High-Level Expert Group on AI (2019) *Ethics Guidelines for Trustworthy AI*.

Hilburn, K.A., Ebert-Uphoff, I., Miller, S.D. (2021). Development and interpretation of a neural-network-based synthetic radar reflectivity estimator using GOES-R satellite observations. *J. Appl. Meteorol. Climatol.* 60(1), 3–21.

Hill, K. (2020), *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*, New York Times, at: https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html.

Hodal, K. (2012). *Death Metal: Tin Mining in Indonesia,* The Guardian, at: http://www.theguardian.com/environment/2012/nov/23/tin-mining-indonesia-bangka.

Hollands, R. G. (2008). *Will the real smart city please stand up? Intelligent, progressive or entrepreneurial*? City, 12(3), 303-320. doi:10.1080/13604810802479126.

Hunt, E. for the Guardian (2016) *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*, at: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter.

IDB-OECD (2021), *Responsible use of AI for public policy: Data science* toolkit, https://publications.iadb.org/publications/english/document/Responsible-use-of-AI-for-public-policy-Data-science-toolkit.pdf.

ILO. (2015). *Guidelines for a just transition towards environmentally sustainable economies and societies for all*, at: https://www.ilo.org/wcmsp5/groups/public/@ed_emp/@emp_ent/documents/publication/wcms_432859.pdf.

International Bank for Reconstruction and Development. (2020). *Global Experiences from Regulatory Sandboxes*, at: https://documents1.worldbank.org/curated/en/912001605241080935/pdf/Global-Experiences-from-Regulatory-Sandboxes.pdf

International Labour Office. (2018). *World Employment and Social Outlook 2018: Greening with jobs* – Geneva: ILO.

IPCC (2014). *Climate Change 2014: Synthesis Report*. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change' [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)] (IPCC, 2014) Geneva, Switzerland.

IPPC Secretariat (2021). *Scientific review of the impact of climate change on plant pests – A global challenge to prevent and mitigate plant pest risks in agriculture, forestry and ecosystems*. Rome, https://doi.org/10.4060/cb4769en.

IPCC (2023). *Sixth Assessment Report, Synthesis Report* (SYR), Summary for Policymakers.

Isern, J. Barranco, F., Deniz, D., Lesonen, J., Hannuksela, J., R. Carrillo, R. (2020). *Reconfigurable Cyber-Physical System for Critical Infrastructure Protection in Smart Cities via Smart Video-Surveillance,* https://doi.org/10.48550/arXiv.2011.14416.

ISO 14040:2006 Environmental management — Life cycle assessment — Principles and framework, at: https://www.iso.org/standard/37456.html

ISO 14044:2006 Environmental management — Life cycle assessment — Requirements and guidelines, at: https://www.iso.org/standard/38498.html.

Jacobs, T. (2023) *Unlocking the Value of Artificial Intelligence (AI) in Supply Chains and* Logistics at: https://throughput.world/blog/ai-in-supply-chain-and-logistics/.

Kaack, L., Donti, P., Strubell, E., Kamiya, G., Creutzig, F. et al. (2021). *Aligning artificial intelligence with climate change mitigation*. ffhal-03368037f.

Kather, J.N. (2023). *Artificial intelligence in oncology: chances and pitfalls*, J Cancer Res Clin Oncol, https://doi.org/10.1007/s00432-023-04666-6.

Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., & Zhang, W. (2018). The Rise of Artificial Intelligence under the Lens of Sustainability. *Technologies, 6(4),* 100. https://doi.org/10.3390/technologies6040100

Kommenda, N. for The Guardian (2021). *How your flight emits as much CO2 as many people do in a year,* at: https://www.theguardian.com/environment/ng-interactive/2019/jul/19/carbon-calculator-how-taking-one-flight-emits-as-much-as-many-people-do-in-a-year.

Kudina, O., Verbeek, P.-P. (2019). Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy. *Science, Technology, & Human Values*, 44(2), 291–314. doi: 10.1177/0162243918793711.

Kumar, B., Roy, S., Sinha, A., Iwendi, C., & Strážovská, Ľ. (2022). E-Commerce Website Usability Analysis Using the Association Rule Mining and Machine Learning Algorithm. Mathematics, 11(1), 25. https://doi.org/10.3390/math11010025.

Kumar, V. T. M. (ed.) (2014). *Governance for Smart Cities, Advances in 21st Century Human Settlements*, Springer Singapore, https://doi.org/10.1007/978-981-287-287-6.

Kuzior, A., Pakhnenko, O., Tiutiunyk, I., & Lyeonov, S. (2023). E-Governance in Smart Cities: Global Trends and Key Enablers. *Smart Cities, 6(4),* 1663–1689. https://doi.org/10.3390/smartcities6040078.

Lacoste, A., Luccioni, A. Schmidt, V., Dandres, T. (20219) *Quantifying the Carbon Emissions of Machine Learning,* https://doi.org/10.48550/arXiv.1910.09700.

Lacoste, A., Luccioni, A., Schmidt, V., and Dandres, T. (2019). *Quantifying the carbon emissions of machine learning*. Preprint arXiv:1910.09700 (2019), https://doi.org/10.48550/arXiv.1910.09700.

Langford, C. (2017). *Houston Schools Must Face Teacher Evaluation Lawsuit*, Courthouse News Service, at: https://www.courthousenews.com/houston-schools-must-face-teacher-evaluation-lawsuit/.

Lee, R., Jose, P.D. (2008). Self-interest, Self-restraint and Corporate Responsibility for Nanotechnologies: Emerging Dilemmas for Modern Managers. *Technology Analysis & Strategic Management* 20:113-125.

Lee, H. et al. (2023) *Summary for Policymakers* in: Synthesis Report of the IPCC Sixth Assessment Report.

Lester, C. .(2019) A Study on Driverless-Car Ethics Offers a Troubling Look Into Our Values, the New Yorker at: https://www.newyorker.com/science/elements/a-study-on-driverless-car-ethics-offers-a-troubling-look-into-our-values.

Liu, R. et al. (2019). *Impacts of the digital transformation on the environment and sustainability*, Öko-Institut, issue paper under Task 3 from the "Service Contract on future EU environment policy". at: https://ec.europa.eu/environment/enveco/resource_efficiency/pdf/studies/issue_paper_digital_transformation_20191220_final.pdf.

Luccioni et al. (2022) *Estimating the Carbon Footprint of BLOOM*, https://doi.org/10.48550/arXiv.2211.02001.

Lynn H. Kaack, Priya L. Donti, Emma Strubell, and David Rolnick (2020). Artificial intelligence and climate change, at: https://eu.boell.org/sites/default/files/2021-04/Artificial%20Intelligence%20and%20Climate%20Change_FINAL_14042021.pdf.

Madiega, T., Van De Pol, A.L. (2022). *Artificial intelligence act and regulatory sandboxes*, European Parliamentary Research Service.

Maind, S., Wankar, P. (2014). Research Paper on Basic of Artificial Neural Network, International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 2 Issue: 1 96 – 100.

Malmodin, J., & Lundén, D. (2018). The Energy and Carbon Footprint of the Global ICT and E&M Sectors 2010–2015. *Sustainability*, 10(9), 3027. https://doi.org/10.3390/su10093027.

Mamalakis, A., Ebert-Uphoff, I., Barnes, E. (2022). Explainable Artificial Intelligence in Meteorology and Climate Science: Model Fine-Tuning, Calibrating Trust and Learning New Science. In: Holzinger, A., Goebel, R., Fong, R., Moon, T., Müller, KR., Samek, W. (eds) *xxAI - Beyond Explainable AI.*, Lecture Notes in Computer Science(), vol 13200. Springer, Cham. https://doi.org/10.1007/978-3-031-04083-2_16.

Marchant E. G., Wallach, W. (2016). Introduction, in: *Emerging Technologies: Ethics, Law and Governance*, 1-12, Routledge, Taylor & Francis.

Marchant, G E. et al. (2020). Governing Emerging Technologies Through Soft Law: Lessons for Artificial Intelligence. *JURIMETRICS*, VOL. 61, ISSUE NO. 1.

Marchant, G. (2011). Addressing the Pacing Problem, in: *The Growing Gap Between Emerging Technologies* and Legal-Ethical Oversight, doi: 10.1007/978-94-007-1356-7_13.

Marchant, G. (2019). "Soft *Law" Governance of Artificial Intelligence*. UCLA: The Program on Understanding Law, Science, and Evidence (PULSE).

Marchant, G. (2021). *Why Soft Law is the Best Way to Approach the Pacing Problem in AI*, at: https://www.carnegiecouncil.org/media/article/why-soft-law-is-the-best-way-to-approach-the-pacing-problem-in-ai.

Martínez-Díaz, M., Al-Haddad, C., Soriguera, F., Antoniou, C., (2021) Platooning of connected automated vehicles on freeways: a bird's eye view, *Transportation Research Procedia,* Volume 58, 2021, p. 479-486, ISSN 2352-1465, https://doi.org/10.1016/j.trpro.2021.11.064.

Martínez-Díaz, M., Soriguera, F. (2018). Autonomous vehicles: theoretical and practical challenges. *Transp. Res. Procedia 33*, 275–282. https://doi.org/10.1016/j.trpro.2018.10.103

Maslej N., et al. (2023). *The AI Index 2023 Annual Report*, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA.

Mazur, S. (2020). *Smart Traffic Management: Optimizing Your City's Infrastructure Spend,* at: https://www.digi.com/blog/post/smart-traffic-management-optimizing-spend.

McCarthy, J., "What is Artificial Intelligence?" (12 November 2007), at: www.formal.stanford.edu/jmc/whatisai.pdf.

McCarthy, J., Minsky L., M., Rochester, N., and Shannon E., C., (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," August 31, 1955, at: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

McCoy, G., L., Brenna, C., Stacy S. C., Vold, K., Das, S., (2022). Believing in black boxes: machine learning for healthcare does not need explainability to be evidence-based, *Journal of Clinical Epidemiology*, Volume 142, p. 252-257, doi: /10.1016/j.jclinepi.2021.11.001.

McCrudden, C. (2008). Human Dignity and Judicial Interpretation of Human Rights, *European Journal of International Law*, Volume 19, Issue 4, Pages 655–724, https://doi.org/10.1093/ejil/chn043.

McKinsey & Company (2023). *What is generative AI?* at: https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai.

Mercedes-Benz Group (2023). *Certification for SAE Level 3 system for U.S. market*, at: https://group.mercedes-benz.com/innovation/product-innovation/autonomous-driving/drive-pilot-nevada.html.

Met Office (2022). *Embedding machine learning and artificial intelligence in weather and climate science and services: a framework for data science in the Met Office* (2022-2027). Exeter, UK.

Mikalef, P., Gupta, M., (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance, Information & Management, Volume 58, Issue 3, 2021, https://doi.org/10.1016/j.im.2021.103434.

Micklitz, H., Pollicino, O., Reichman, A., Simoncini, A., Sartor, G., De Gregorio D. (Eds.) (2021). Constitutional Challenges in the Algorithmic Society, Cambridge: Cambridge University Press. doi:10.1017/9781108914857.010.

Miller, AP., Hosanagar K. (2019). How targeted Ads and dynamic pricing can perpetuate bias. *Harvard Business Review*. https://hbr.org/2019/11/how-targeted-ads-and-dynamic-pricing-can-perpetuate-bias.

Miller, R. (2017). *Autonomous cars could drive a deluge of data center demand*, at: https://datacenterfrontier.com/autonomous-cars-could-drive-a-deluge-of-data-center-demand/.

Min, H. (2010) Artificial intelligence in supply chain management: theory and applications, *International Journal of Logistics Research and Applications*, 13:1, 13-39, DOI: 10.1080/13675560902736537

Misuraca, G., van Noordt, C. (2020). *Overview of the use and impact of AI in public services in the EU*, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19540-5, doi:10.2760/039619, JRC120399.

Mitropoulos, L., Kortsari, A., Mizaras, V., & Ayfantopoulou, G. (2023). Mobility as a Service (MaaS) Planning and Implementation: Challenges and Lessons Learned. *Future Transportation,* 3(2), 498–518. https://doi.org/10.3390/futuretransp3020029.

Montes, J. (2020). *A Historical View of Smart Cities: Definitions, Features and Tipping Points*, doi.org/10.2139/ssrn.3637617.

Moore, G. (1965). Cramming more components onto integrated circuits, *Electronics*, Volume 38, Number 8.

Moraes, T. (2023). *Regulatory sandboxes for Artificial Intelligence – hype or solution?* KU Leuven AI Summer School blogspot, at: https://www.law.kuleuven.be/ai-summer-school/blogpost/Blogposts/regulatory-sandboxes.

Muench, S., Stoermer, E., Jensen, K., Asikainen, T., Salvi, M. and Scapolo, F., (2022) *Towards a green and digital future*, EUR 31075 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-52451-9, doi:10.2760/977331, JRC129319.

Murray, D, Liao, J, Stankovic, L, Stankovic V. (2016). Understanding usage patterns of electric kettle and energy saving potential. *Appl Energy* 171:231–242. https://doi.org/10.1016/j.apenergy.2016.03.038

National Transportation Safety Board (2018). *Highway Accident Report–Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian*, Tempe, Arizona, NTSB/HAR-19/03, 2019, at: https://www.ntsb.gov/investigations/accidentreports/reports/har1903.pdf.

Nelson, M. for DigitalInfraNetwork (2023) *Next-gen submerged cooling installed in Danish data centre*, at: https://digitalinfranetwork.com/news/next-gen-submerged-cooling-installed-in-danish-data-centre/.

Nobre, GC., Tavares, E. (2017). Scientific literature analysis on big data and internet of things applications on circular economy: a bibliometric study. *Scientometrics* 111(1):463–492. https://doi.org/10.1007/s11192-017-2281-6

Norberto Nuno de, A., Galindo, L., Zarra, A. (2023). *Artificial Intelligence Act: A Policy Prototyping Experiment: Regulatory Sandboxes*, at: https://openloop.org/reports/2023/04/Artificial_Intelligence_Act_A_Policy_Prototyping_Experiment_Regulatory_Sandboxes.pdf.

O' Carroll, L. for the Guardian. (2023). *EU moves closer to passing one of world's first laws governing AI,* at: https://www.theguardian.com/technology/2023/jun/14/eu-moves-closer-to-passing-one-of-worlds-first-laws-governing-ai.

Ocean Studies Board, The National Academies of Sciences, Engineering, and Medicine et al. (2016). *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts*. National Academies Press.

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.

OECD DIGITAL ECONOMY PAPERS No. 349 (2023) *ADVANCING ACCOUNTABILITY IN AI GOVERNING AND MANAGING RISKS THROUGHOUT THE LIFECYCLE FOR TRUSTWORTHY AI*.

OECF Panel (2022) *What is the environmental footprint of artificial intelligence*?, Informed by experts from the OECD.AI Expert Group on AI Compute and Climate and the Global Partnership on AI (GPAI) Responsible AI Working Group (RAI).

Othman, K. (2021). Public acceptance and perception of autonomous vehicles: a comprehensive review. *AI Ethics 1*, 355–387, https://doi.org/10.1007/s43681-021-00041-8.

Our World in Data (2021), CO2 per capita, at: https://ourworldindata.org/grapher/co2-per-capita-marimekko?tab=table.

Pampoukis, H., (2022) *Legal aspects of AI*, at "UNESCO: Open Science - Artificial Intelligence. Young New Horizons, Visions and Frontiers" Conference.

Parish, F., Sirin, A. A., Charman, D., Joosten, H., Minaeva, Y., Silvius, M. (2008) *Assessment on peatlands, biodiversity and climate change: Main report*. Global Environment Centre, Kuala Lumpur andWetlands International, Wageningen

Paris Agreement to the United Nations Framework Convention on Climate Change, Dec. 12, 2015, T.I.A.S. No. 16-1104:

Pasquale, F. A. (2015). *The Black Box Society: The Secret Algorithms that Control Money and Information*, Harvard University Press.

Paul Hawken. (2015). Drawdown: The *Most Comprehensive Plan Ever Proposed to Reverse Global Warming*. Penguin Books.

Pearce, D. W., Turner, R. K. (1990). *Economica of Natural Resources and the Environment*, The John Hopkings University Press.

Philips, J., Hahn, C., Fontana, P., Yates, A., Greene, K., Broniatowski, D. and Przybocki, M. (2021), *Four Principles of Explainable Artificial Intelligence*, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, doi: 10.6028/NIST.IR.8312.

Piao, J., McDonald, M., Hounsell, N., Graindorge, M., Graindorge, T., Malhene, N. (2016) Public views towards implementation of automated vehicles in urban areas. *Transp. Res. Procedia 14*, 2168–2177.

Poole, D., Mackworth, A., Goebel, R. (1998). *Computational Intelligence, A Logical Approach*, New York Oxford University Press.

Polisena, V. (2021). *Philosophy of Artificial Intelligence: Ethics for Smart Cities,* AI Ethics.

Pouikli, Kl. Tsakalogianni, I. (2022). Criteria for identifying splitting of projects (salami-slicing) tactic in light of the C-11/20 Regione Puglia case and EIA practice in Greece, *Journal for European Environmental & Planning Law*, 19(3), 222-247. doi: https://doi.org/10.1163/18760104-19030005.

Pouikli, A. Delegkou, A., Tsakalogianni, I. (2022). Circular Economy for a climate-neutral Europe: Mapping the latest policy and legislative developments *Carbon & Climate Law Review,* Volume 16, Issue 4 (2022) pp. 281 - 287 DOI: https://doi.org/10.21552/cclr/2022/4/8.

PROPOSAL FOR A COUNCIL DIRECTIVE ON THE LEGAL PROTECTION OF COMPUTER PROGRAMS, COM/88/816 FINAL, OJ C 91, 12.4.1989.

PROPOSAL FOR A DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on energy efficiency (recast) COM/2021/558 final.

PROPOSAL FOR A REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, COM/2021/206 final.

Psiroukis, V. et al. (2021). Monitoring of free-range rabbits using aerial thermal imaging, *Smart Agricultural Technology*, Volume 1, 100002, https://doi.org/10.1016/j.atech.2021.100002.

Rana, P., Miller, D. (2019) Machine learning to analyze the social-ecological impacts of natural resource policy: insights from community forest management in the Indian Himalaya, *Environ. Res. Lett.* 14 024008, DOI: 10.1088/1748-9326/aafa8f.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu J. P. (2019). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, *Journal of Machine Learning Research,* 21 (2020) 1-67, arXiv:1910.10683v3, doi: 10.48550/arXiv.1910.10683.

Rajpurkar, P., Chen, E., Banerjee, O. et al. (2022). AI in health and medicine. *Nat Med* 28, 31–38, https://doi.org/10.1038/s41591-021-01614-0

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

REGULATION (EU) 2019/1020 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011.

REGULATION (EU) 2019/2144 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users, amending Regulation (EU) 2018/858 of the European Parliament and of the Council and repealing Regulations (EC) No 78/2009, (EC) No 79/2009 and (EC) No 661/2009 of the European Parliament and of the Council and Commission Regulations (EC) No 631/2009, (EU) No 406/2010, (EU) No 672/2010, (EU) No 1003/2010, (EU) No 1005/2010, (EU) No 1008/2010, (EU) No 1009/2010, (EU) No 19/2011, (EU) No 109/2011, (EU) No 458/2011, (EU) No 65/2012, (EU) No 130/2012, (EU) No 347/2012, (EU) No 351/2012, (EU) No 1230/2012 and (EU) 2015/166

REGULATION (EU) 2021/1056 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 24 June 2021 establishing the Just Transition Fund. The Fund focuses on economic diversification of the territories most affected by climate transition (i.e., with targeted economic incentives to specific areas) and on the re-skilling of workers and job seekers.

REGULATION (EU) 2021/1119 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 June 2021 establishing the framework for achieving climate neutrality and amending Regulations (EC) No 401/2009 and (EU) 2018/1999 ('European Climate Law').

REGULATION (EU) 2022/1925 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act).

REGULATION (EU) 2022/2065 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act).

REGULATION (EU) 2022/868 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).

Reichman, A., & Sartor, G. (2021). *Algorithms and Regulation*. In H. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor, & G. De Gregorio (Eds.), Constitutional Challenges in the Algorithmic Society (pp. 131-181). Cambridge: Cambridge University Press. doi:10.1017/9781108914857.009.

Reike, D., Vermeulen, W. J., & Witjes, S. (2018). The circular economy: New or Refurbished as CE 3.0? — Exploring Controversies in the Conceptualization of the Circular Economy through a Focus on History and Resource Value Retention Options. *Resources, Conservation and Recycling*, 135, 246-264.doi: 10.1016/j.resconrec.2017.08.027

Rezaei, A., Caulfeld, B. (2020) Examining public acceptance of autonomous mobility. *Trav. Behav. Soc*. 21, 235–246, https://doi.org/10.1016/j.tbs.2020.07.002.

Rio Declaration on Environment and Development, Jun. 13, 1992. 31 ILM 874 (1992).

Rizos, V., Tuokko, K. Behrens, A. (2017). *The Circular Economy A review of definitions, processes and impacts*, at: https://www.researchgate.net/publication/315837092_The_Circular_Economy_A_review_of_definitions_processes_and_impacts.

Roberts, H., Zhang, J., Bariach, B. et al. (2022). Artificial intelligence in support of the circular economy: ethical considerations and a path forward. *AI & Soc,* https://doi.org/10.1007/s00146-022-01596-8.

Rolnick, D. et al. (2019) *Tackling Climate Change with Machine Learning*, at: https://arxiv.org/abs/1906.05433.

Rong, S., Bao-wen, Z. (2018). *The research of regression model in machine learning field*, MATEC Web, Conf. 176 01033 (2018), doi: 10.1051/matecconf/201817601033

Rory Cellan-Jones for BBC News (2014) *Stephen Hawking warns artificial intelligence could end mankind*, at: https://www.bbc.com/news/technology-30290540.

Russell, S., Norvig, P. (2009). "Artificial Intelligence: A Modern Approach", Prentice Hall, 3rd edition, 2009.

SAE International J3016, International Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016:Sept2016), doi: https://doi.org/10.4271/J3016_202104.

Sandvig C., Hamilton K., Karahalios K, et al. (2014) *Auditing algorithms: Research methods for detecting discrimination on internet platforms*. In: Annual Meeting of the International Communication Association, Seattle, WA, pp. 1–23.

Saurugger S. & Terpan F. (2021) Normative transformations in the European Union: on hardening and softening law, *West European Politics*, 44:1, 1-20, DOI: 10.1080/01402382.2020.1762440.

Schaeffer, R. Sims, J. Corfee-Morlot, F. Creutzig, X. Cruz-Nunez, D. Dimitriu, and M. D'Agosto (2014). *Transport, in IPCC, Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Climate Change 2014: Mitigation of Climate Change, Chapter 8*. Geneva. O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S.

Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, and J. C. Minx (Eds.). Cambridge University Press, Cambridge

Schroeder, P. and Barrie, J. (2022). Is going circular just? Environmental justice and just transition – key elements for an inclusive circular economy, *Field Actions Science Reports* [Online], Special Issue 24, http://journals.openedition.org/factsreports/6864

Schwartz, R., Dodge, J., Smith, N., Etzioni, O. (2019). *Green AI.* CoRR, abs/1907.10597, http://arxiv.org/abs/1907.10597.

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A. and Hall, P. (2022). *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, https://doi.org/10.6028/NIST.SP.1270.

Scott, J. (2014). Extraterritoriality and territorial extension in EU law. *The American Journal of Comparative Law*, 62(1), 87–126.

Shardelow, C. 2021. *Avoiding the Basilisk: An Evaluation of Top-Down, Bottom-Up, and Hybrid Ethical Approaches to Artificial Intelligence*, University of Nebraska-Lincoln.

Shchitova, A. (2020). Definition of Artificial Intelligence for Legal Regulation, 2nd International Scientific and Practical Conference on Digital Economy (ISCDE 2020), *Advances in Economics, Business and Management Research,* volume 156.

Silveira, L. (2014). Discovery and Justification of Judicial Decisions: Towards More Precise Distinctions in Legal Decision-Making, *Law and Method*, ReM 2014-09, doi: 10.5553/REM/.000007.

Simpson, C. for Bloomberg (2012). *The Deadly Tin Inside Your Smartphone*, available at: https://www.bloomberg.com/news/articles/2012-08-23/the-deadly-tin-inside-your-smartphone.

Smuha, N. A., (2019). From a 'Race to AI' to a 'Race to AI Regulation' - Regulatory Competition for Artificial Intelligence, *Law, Innovation and Technology*, Vol. 13, Iss. 1, 2021, http://dx.doi.org/10.2139/ssrn.3501410

Solon, O. for the Guardian (2016). *Team of hackers take remote control of Tesla Model S from 12 miles away,* at: https://www.theguardian.com/technology/2016/sep/20/tesla-model-s-chinese-hack-remote-control-brakes.

Strubell, E., Ganesh, A., McCallum, A. (2019). *Energy and policy considerations for deep learning in NLP*, Proceedings of the 57 Annual Meeting of the Association for Computational Linguistics.

Surden, H. (2019), Artificial Intelligence and Law: An Overview, *Georgia State University Law Review*, Vol. 35.

Sustainable Infrastructure Alliance (2022), *The Roadmap to Sustainable Digital Infrastructure by 2030*, https://sdialliance.org/roadmap.

Tasioulas, J. (2019). First Steps Towards an Ethics of Robots and Artificial Intelligence. *Journal of Practical Ethics*, Volume 7, Number 1, June 2019, King's College London Law School Research Paper.

Terpan F., (2015), Soft Law in the European Union, *European Law Journal*, 21: 68 -96. DOI: 10.1111/eulj.12090.

The World Bank (2021). *Climate Smart Agriculture,* at: https://www.worldbank.org/en/topic/climate-smart-agriculture.

Towsey, M., et al. (2012). *A toolbox for animal call recognition*, Bioacoustics, 21:2, 107-125, DOI: 10.1080/09524622.2011.648753.

Treaty on the Functioning of the European Union, OJ C 326, 26.10.2012, ELI: http://data.europa.eu/eli/treaty/tfeu_2012/oj

Tuia, D., Kellenberger, B., Beery, S. et al. (2022). Perspectives in machine learning for wildlife conservation. *Nat Commun* 13, 792, https://doi.org/10.1038/s41467-022-27980-y.

Turing A. M. (1950) "I.—Computing Machinery and Intelligence", *Mind*, Volume LIX, Issue 236, 1950, Pages 433–460, https://doi.org/10.1093/mind/LIX.236.433.

UNESCO (2017). *Report of World Commission on the Ethics of Scientific Knowledge and Technology on Robotics Ethics*, http://unesdoc.unesco.org/images/0025/002539/253952E.pdf.

UNESCO (2022) *Being able to predict floods will save lives in West Africa,* https://www.unesco.org/en/articles/being-able-predict-floods-will-save-lives-west-africa.

United Nations, 1992 Rio Declaration on Environment and Development, UN Doc. a/ conf.151/26 (vol. I), 31 ilm 874 (1992).

United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Population Prospects 2019: Highlights*.

United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Urbanization Prospects: The 2018 Revision.*

United Nations, UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE [1992].

Varošanec, I., (2022). On the Path to the Future: Mapping the Notion of Transparency in the EU Regulatory Framework for AI. *International Review of Law, Computers & Technology Vol*. 36(2), University of Groningen Faculty of Law Research Paper No. 18/2022, http://dx.doi.org/10.2139/ssrn.4066020.

Vermeulen, W., Reike, D, & Witjes, S. (2019). *Circular Economy 3.0 - Solving confusion around new conceptions of circularity by synthesising and re-organising the 3R's concept into a 10R hierarchy,* at: https://www.researchgate.net/publication/335602859_Circular_Economy_30_-_Solving_confusion_around_new_conceptions_of_circularity_by_synthesising_and_re-organising_the_3R's_concept_into_a_10R_hierarchy .

Wallach, W., Allen, C. (2009). Moral *Machines: Teaching Robots Right from Wrong*, Oxford Academic, https://doi.org/10.1093/acprof:oso/9780195374049.001.0001.

Wang, N., Tian, M.Y. (2022). 'Intelligent Justice': AI Implementations in China's Legal Systems. In: Hanemaayer, A. (eds) *Artificial Intelligence and Its Discontents*. Social and Cultural Studies of Robots and AI. Palgrave Macmillan, Cham. doi: /10.1007/978-3-030-88615-8_10.

Weizenbaum, J. (1966) ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9, 1, 36–45. https://doi.org/10.1145/365153.365168.

William F Lamb et al. (2021). *Environmental Research Letters*, 16 073005, DOI: 10.1088/1748-9326/abee4e.

Wismans, L., de Romph, E., Friso, K., Zantema, K. (2014). Real Time Traffic Models, Decision Support for Traffic Management, *Procedia Environmental Sciences*, Volume 22, 220-235, https://doi.org/10.1016/j.proenv.2014.11.022.

Wolfson, A. for the Guardian (2018) Amazon's Alexa recorded private conversation and sent it to random contact, at: https://www.theguardian.com/technology/2018/may/24/amazon-alexa-recorded-conversation.

Woollard, F., Howard-Snyder, F. (2022). Doing vs. Allowing Harm, The Stanford Encyclopedia of Philosophy, Edward N. Zalta & Uri Nodelman (eds.).

World Health Organization, *Surveillance in Emergencies, World Health Organization*, 2021. https://www.who.int/emergencies/surveillance.

World Meteorological Organisation, World Climate Conference-1 (12-23 February 1979; Geneva, Switzerland), https://library.wmo.int/doc_num.php?explnum_id=3778.

Woudstra, F. (2020). *Ethical Guidelines for Transparent Development and Implementation of AI - an Overview*, Filosofie in actie Blog, at: https://www.filosofieinactie.nl/blog/2020/4/9/ethical-guidelines-for-transparent-development-and-implementation-of-ai.

Wu, C-J., et al. (2022) *Sustainable AI: Environmental Implications, Challenges and Opportunities*, arXiv:2111.00364v2 [cs.LG].

Yeung, S.K., Yay, T., Feldman, G. (2021). Action and Inaction in Moral Judgments and Decisions: Meta-Analysis of Omission Bias Omission-Commission Asymmetries. *Personality and Social Psychology Bulletin.* 48 (10): 1499–1515. doi:10.1177/01461672211042315

Zalewski, T. (2021) Definition of artificial intelligence, in: Lai, L., Świerczyński, M. (eds.), *Legal and Technical aspects of Artificial Intelligence*, Scientific Publishing House of the Cardinal Stefan Wyszynski University, Warsaw.

Zhao, Z. W, Zhou, K. et. al (2023, ongoing). *A Survey of Large Language Models*, arXiv:2303.18223v4, doi: 10.48550/arXiv.2303.18223.