CHAPTER 11

ARTIFICIAL INTELLIGENCE FOR SOCIAL GOOD: THE WAY FORWARD

Nuria Oliver
Institute of Humanity-centric AI (ELLIS Unit Alicante Foundation)

Summary

The chapter explores both opportunities and challenges linked to the implementation of artificial intelligence (AI) methods to address the sustainable development goals (SDGs). AI has the potential to improve our ability to measure and identify weaknesses, priorities and areas of improvement related to the SDGs, while accelerating their achievement. To realise such potential, five types of barriers need to be addressed (institutional, technical, ethical, financial and environmental) to effectively leverage the power of data-driven AI methods and accelerate their positive impact on the SDGs.

'It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us, we were all going direct to Heaven, we were all going direct the other way – in short, the period was so far like the present period, that some of its noisiest authorities insisted on its being received, for good or for evil, in the superlative degree of comparison only.'

Charles Dickens, *A Tale of Two Cities*

1. Introduction

As in Dickens' words, it is indeed the best of times and the worst of times. We live in a time of prosperity, but we also face tremendous global challenges that threaten our existence as a species – from poverty and hunger to climate change and the destruction of entire ecosystems. Effectively tackling these challenges requires an ambitious and coordinated commitment from most nations in the world. Hence, since the mid-1990s, starting with the Copenhagen Declaration on Social Development in 1995 and the six International Development Goals that followed in 1996, the United Nations (UN) has periodically established far-reaching goals for the world, aimed at addressing the most pressing issues of our times at a global scale and in a coordinated manner.

In 2015, when the Millennium Development Goals approached their target date, the UN defined a new set of global goals and the 2030 Agenda for Sustainable Development, which was adopted by all EU Member States in 2015. The resulting global goals are known as the 17 SDGs, a call to action by all countries (developed and developing) with the aim of eradicating the world's poverty and other deprivations, together with improving health and education, fostering economic growth, reducing inequality, preserving our environment and combating climate change.

In parallel to the establishment of such an ambitious agenda for the world, a global movement gained traction on the role that data and AI could play in this context from two perspectives: first, to help us better measure the level of achievement of the SDGs and identify weaknesses, priorities and areas for improvement; and second, to enable and accelerate the achievement of the SDGs. In November 2014, the UN published the report 'A World that Counts: mobilizing the data revolution for sustainable development'[1], authored by

1 https://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf

the Independent Expert Advisory Group on a Data Revolution for Sustainable Development as requested by the UN Secretary-General. The report outlines both the opportunities and the risks that the 'data revolution' presents for sustainable development, and proposes five key recommendations for actions, including investing resources in capacity development, sharing technology and innovations for the common good, developing a global consensus on principles and standards and creating the Global Partnership for Sustainable Development Data[2], which was created in 2015 'to help stakeholders across countries and sectors fully harness the data revolution for sustainable development, using this new knowledge to improve lives and protect the planet'. The UN subsequently organised three editions of the World Data Forum, in 2017, 2018 and 2020 in South Africa, Dubai and Switzerland, respectively. The 2018 Forum wrapped up with the launch of the Dubai Declaration, which aims to increase financing for better data and statistics for sustainable development.

Moving from a global context to the European arena, the European Commission established six priorities for the 2019-2024 period[3], which include the twin green and digital transitions, captured by the 'European Green Deal' and 'Europe fit for the digital age' priorities, respectively. The European Commission considers these two transitions to be deeply interrelated, as it is evident that digital technologies are playing and will continue to play a crucial role in enabling Europe to move to a clean and circular economy, restore biodiversity and reduce pollution. Thus, data and AI are considered to be not only key pillars of the digital transition, but also of the green transition.

The European Commission recognises that data – and more importantly, the ability to use it, analyse it (prominently by means of data-driven AI methods) and draw insights from it – are essential for sustainable growth and innovation. The European vision for AI entails developing and using trustworthy AI systems, that is, systems that are safe, ethical, transparent, unbiased and under human control. Such a vision is articulated in several strategic documents, including an ethical framework to achieve trustworthy AI[4], a set of policy and investment recommendations to boost Europe's competitiveness in AI[5], a new European regulation on data governance to facilitate data sharing across the Member States – placing citizens at its centre and contributing to the creation of a European single data market[6], and a new European regulation of AI systems based on a classification of their risk, which can range from unacceptable (and thus banned) to minimal risk[7].

A key question posed by many scientists, policy makers, practitioners, activists and citizens today is whether these data and AI revolutions that we are immersed in will contribute to achieving **sustainable development**, i.e. development that not only meets the needs of the present but ensures the ability of future generations to meet their own needs.

In this chapter, I provide an overview of both the tremendous opportunities that data-driven AI methods offer to help us address the 17 SDGs and the challenges and limitations posed by AI that might hinder the realisation of such potential.

2 https://www.data4sdgs.org/
3 https://ec.europa.eu/info/strategy/priorities-2019-2024_en
4 https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
5 https://www.europarl.europa.eu/italy/resource/static/files/import/intelligenza_artificiale_30_aprile/ai-hleg_policy-and-investment-recommendations.pdf
6 https://data.europa.eu/en/highlights/data-governance-act-open-data-directive
7 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

2. AI and the 17 SDGs

AI is the discipline within computer science or engineering that has the objective of the development of computationally – i.e. non-biological – intelligent systems, taking human intelligence as a reference. In the same way that human intelligence is complex and diverse, there are many areas of knowledge within AI that aim to emulate specific aspects of human intelligence, such as computer vision, speech recognition, natural language processing, planning, reasoning, knowledge representation, learning theory and decision-making.

Historically, there have been four views in the literature as to how to achieve AI or what AI means: (1) AI means **acting humanly**, i.e. acting like a person – the Turing test is a classic example of such view of AI; (2) AI means **thinking humanly**, i.e. thinking like a person, which is the object of study of cognitive science; (3) AI means **thinking rationally**, i.e. modelling thinking as a logical process, where conclusions are reached based on symbolic logic; and (4) AI means **acting rationally**, i.e. performing actions to achieve one's goal, based on one's understanding and beliefs about the world.

In terms of how to build AI systems, there have been two basic schools of AI since its emergence in the 1950s: first, the top-down or symbolic-logic school, and second, the bottom-up or data-driven school. According to the symbolic-logic school, to achieve AI, human knowledge would be collected and codified, deriving new knowledge from such initial knowledge using the rules of logic. The methods in this school are based on symbolic representations of problems, logic and search. This approach to AI was the dominant paradigm from the birth of the discipline in the 1950s until the late 1980s. The canonical example of the top-down school are expert systems, which were the first successful example of commercialisation of AI systems.

The methods developed in the bottom-up, data-driven school, are inspired from biology: biological intelligent beings learn from their interactions with their environment, from experience. Hence, bottom-up approaches to AI focus on developing methods that learn from data as opposed to modelling symbolic descriptions of the environment. The canonical example of a bottom-up method are neural networks.

Bottom-up, data-driven methods in AI have experienced an unprecedented exponential growth in the past 15 years mainly due to three factors:

- the existence of massive amounts of non-structured data (referred to as 'big data') which is the result of both our interactions with the digital world and the increased digitisation of the physical world;
- the availability of large-scale computing at low cost, as a consequence of Moore's law;
- the development of sophisticated machine learning algorithms, inspired by the neural networks from the 1950s, but significantly more complex, called deep neural networks or deep learning (LeCun et al., 2015), which have the flexibility and the power to learn from large-scale data by leveraging high performance computing.

Because of these three factors, we have witnessed tremendous achievements in data-driven machine learning algorithms applied to numerous areas, including computer vision, audio processing, natural language processing, time series analysis, recommendations, reinforcement learning and control, robotics and uncertainty quantification. Thus, it should not come as a surprise that these methods are at the core of most of the AI-enabled systems that we use in our phones, our homes, our cities and our cars.

The domains most likely to be disrupted, transformed and enriched by these new AI approaches are data-rich, challenges related to identifying patterns and trends in non-structured data (images, videos, audio, text, sensor data, etc.), challenges that require making predictions about future phenomena and/or would benefit from data-driven decisions.

Thus, these advances are valuable to address many of the challenges related to the 17 SDGs. In fact, in recent years several research papers have been published (Vinueasa et al., 2018) and initiatives have been launched to identify projects that investigate the use of AI in the context of the 17 SDGs. Examples of such initiatives include the SDG AI Repository managed by the UN's International Telecommunication Union (ITU) agency[8]; the database of the AI for Sustainable Development Goals (AI4SDGs) Think Tank[9] and the database of the University of Oxford's Research Initiative AIxSDGs[10], which lists 108 projects.

But what are the concrete opportunities that AI offers in the context of the SDGs? How can AI methods help us to achieve such an ambitious global agenda? What are the challenges associated with leveraging AI for social good?

The following section provides an overview of the challenges and opportunities for AI in each of the SDGs, except SDG 17, which refers to the importance of establishing partnerships and collaborations across regions, countries and institutions in pursuit of all the goals by 2030. Therefore, it is not included in the discussion.

SDG 1 – No poverty

After declining for 20 years, global extreme poverty rose again in 2020 due to a variety of factors, including the impact of the COVID-19 pandemic and climate change[11]. The World Bank estimates that up to 1.9 billion people in the world today live below the societal poverty line, which combines the USD 1.9/day absolute poverty line with a country-dependent component based on the median consumption or income in the country. Most of the poor live in rural areas and poverty is a long-lasting reality in many parts of the world. However, obtaining granular, high-quality data on poverty to inform policy making is still a challenge.

AI techniques have been used to automatically analyse satellite (Jean et al., 2016), mobile (Syndsoy et al., 2016; Soto et al., 2011; Blumenstock and Cadamuro, 2015) or digital transaction and real state online advertisements (Cruz et al., 2019) to automatically infer poverty or socio-economic levels in developing and developed countries.

Beyond leveraging AI methods to assess poverty, AI-powered, evidence-based decision-support systems could inform public decisions relative to poverty eradication programmes both to measure the success of such programmes and to guide resource allocation depending on the estimated current and predicted levels of poverty in different regions. Moreover, data-driven AI is emerging as a driver to improve the overall quality of life[12].

8 https://www.itu.int/en/ITU-T/AI/Pages/ai-repository.aspx
9 https://ai-for-sdgs.academy/about
10 https://www.aiforsdgs.org/
11 https://www.worldbank.org/en/topic/poverty/overview
12 E. O. of the President National Science and T. C. committee on technology. Preparing for the future of AI. https://obamawhitehouse.archives.gov/sites/default/fil es/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf, 2016.

SDG 2 – Zero hunger

Hunger refers to the generalised lack of access to food by a community or a population in a sustained manner. Hunger still prevails in many developing countries, and it is often exacerbated by extreme weather events (e.g. severe draughts), poverty and/or wars. The early detection of hunger is typically an effective strategy to prevent or mitigate it (Holley, 2018).

Weather (United States Agency for International Development, 2010), satellite, demographic (Quinn et al., 2010) and socio-economic (Okori and Obua, 2011) data have been analysed using AI techniques to make an early detection of hunger in developing countries, such as Uganda. Other authors have used machine learning techniques to predict food demand in areas impacted by natural disasters (Xiaoyan et al., 2010).

There are several examples where AI techniques have been used to predict the yield of crops from climate and agriculture data (Gandhi and Armstrong, 2016; Zhu et al., 2018), sometimes combined with satellite data (Badr et al., 2016). Invasive species and plagues have been automatically recognised in images by deep neural networks (Fedor et al., 2009; Mohanty et al., 2016) and machine learning techniques have been proposed to identify and recommend crops depending on the characteristics of the soil (Kulkarni et al., 2018).

Several international organisations, including UN agencies, the World Bank, NGOs (such as Mercy Corps, Save the Children and Oxfam) and data institutions (such as the UN Centre for Humanitarian Data, the Integrated Food Security Phase Classification, IPC, or the Famine Early Warning Systems Network, FEWS) have partnered in the Famine Action Mechanism (FAM)[13], a global initiative to end famine. FAM was launched in 2018 and focuses on three data-driven areas of collaboration to anticipate and address food security crises: food security crisis risk analysis, anticipatory and early action financing, and programming. The Group on Earth Observations Global Agricultural Monitoring Initiative (GEOGLAM)[14] is an open community created to increase market transparency and improve food security by producing and sharing relevant, timely and actionable data on agricultural conditions and outlooks of food production at different scales (national, regional and global).

Finally, there are two emergent, relevant concepts: precision agriculture (Zhang et al., 2002) and smart farming (Sundmaeker et al., 2016; Wolfert et al., 2017), which focus on leveraging data captured by a variety of pervasive sensors and state-of-the-art technology to optimise the yield of crops while preserving resources. According to the Food and Agriculture Organization (FAO) (2017), smart farming refers to the use of modern digital technology – including internet-of-things sensors, autonomous drones (Faulkner and Cebul, 2014), robots to feed cattle (Grobart, 2012) and AI techniques to analyse the data captured by the sensors – to improve agricultural production systems. The European Commission has established Horizon 2020 programmes to promote smart farming.

SDG 3 – Good health and well-being

Having good health is a human right and a key contributor to growth and prosperity[15]. The levels of health and well-being in a population are a proxy indicator of the nation's level of progress. Unfortunately, we are still far from achieving good health everywhere on the planet.

13 https://www.worldbank.org/en/programs/famine-early-action-mechanism

14 https://earthobservations.org/geoglam.php

15 https://irp-cdn.multiscreensite.com/be6d1d56/files/uploaded/SDSN%20Health%20Solutions.pdf, 2019.

AI for public policy making during the COVID-19 pandemic: the Valencian experience

In March 2020, the Valencian government created the Data Science against COVID-19 taskforce, composed of over 20 volunteer scientists from several Valencian research institutions (universities and research centres), working on four areas to support the Government's decision-making during the pandemic:

- large-scale human mobility modelling via the analysis of large-scale data derived from the mobile network infrastructure;
- computational epidemiological models to predict the evolution of the pandemic curve, not only under the current conditions but also under different scenarios of non-pharmaceutical interventions;
- machine learning-based predictive models of hospital and intensive care occupancy;
- a large-scale, online citizen survey via the COVID-19 Impact Survey, which is one of the largest surveys in the world about COVID-19, with over 700 000 answers.

The work of this taskforce has received national and international visibility and recognition, including winning first prize at the 500k XPRIZE Pandemic Response Challenge competition, sponsored by Cognizant. It is the first time that a team from Spain (ValenciaIA4COVID) has won an XPRIZE competition. As part of the XPRIZE challenge, the team developed a novel deep learning-based epidemiological model able to predict the number of COVID-19 cases in 236 countries and regions in the world and a non-pharmaceutical intervention prescriptor, recommending up to 10 different public policies that would have the optimal trade-off between the cost of the public policies and their impact on containing the number of COVID-19 cases. This work also received the best paper award at ECML-PKDD 2021.

This initiative is an example of the use of AI for social good, by means of a collaboration between academia and the scientific community, society at large (through the citizen survey) and a government to achieve evidence-driven decision-making.

The intersection between data, AI and health is rich and full of opportunity, as highlighted by many authors (Singh, 2019; Guo and Li, 2018). Broadly speaking, AI methods are redefining healthcare from at least three perspectives.

The first perspective is by accelerating the discovery and design of effective treatments and vaccines, enabling the prediction of expected results and side-effects, in addition to automating the discovery of new pharmacological compounds (Ong et al., 2020; Schneider, 2018) and protein folding (Senior et al., 2020).

The second is by assisting in clinical decision-making related to, e.g., the diagnosis of cancer (Esteva et al., 2017; Fauw et al., 2018), COVID-19 (Oh et al., 2020) or tuberculosis (Doshi et al., 2017) in radiological tests, potentially providing expert feedback and diagnoses to patients where human medical experts might not be available; improving pregnancy, post-partum (Rodriguez et al., 2016; Poon et al., 2009) and infant care and thus preventing deaths (Malak et al., 2018; Adegbosin et al., 2019); and predicting the efficacy of treatments (Pham et al., 2017) or the probability of needing intensive care (Kaji et al., 2019).

The third is by supporting policy-making related to public health – including mental health (Walsh et al., 2017) and infectious diseases, such as malaria (Wasolowski et al., 2012), influenza (Kagashe et al., 2017), Ebola (Wasolowski et al., 2014) and COVID-19 (Oliver et al., 2020) – via the analysis of multi-dimensional data captured by the mobile network infrastructure, social media platforms and pervasive sensors.

Moreover, the increased availability of wearable devices at affordable prices (e.g. activity wristbands, smartwatches, etc.) enables the collection of large-scale, longitudinal data about daily activities, sleep habits and physiological signals, which, analysed via machine learning techniques, could be extremely valuable in the early diagnosis of disease and the realisation of personalised, preventive and predictive medicine (Clifton et al., 2013).

In fact, precision (predictive, personalised, preventive) medicine will not be achievable without the use of AI techniques applied to genomic, behavioural, contextual (e.g. pollution, weather) and medical data (Collins and Varmus, 2015).

On the negative side, in addition to numerous ethical and technical challenges discussed below, we need to consider that data-driven AI methods are at the core of the digital services and social media platforms that we use today, so they can personalise the user experience, recommend relevant content and increase engagement. Unfortunately, these services, which are designed to maximise our engagement, could lead to an excessive (and possibly addictive) use by their users with potential negative consequences for our well-being (Zheng and Lee, 2016).

SDG 4 – Quality education

Education is a key pillar for sustainable development and prosperity. While the world has made progress in reducing the education gap, particularly for women and girls, there are still today over 260 million children of primary and secondary school age worldwide who do not attend any school, 130 million children who can barely read and write despite attending school and 75 million children aged 3-18 years old who live in countries facing violence and war, needing educational support.

AI has the potential to contribute to education in several ways. First, by enabling a personalisation of the learning experience, moving from a generalist, one-to-many education model to an individual, one-to-one model. Intelligent tutoring systems (ITS) via software agents, chatbots or social robots can personalise both the content and the strategies used to teach students, to maximise their learning. In addition, AI-powered intelligent educational interfaces enable the early detection of students with physical or mental disabilities (David and Balakrishnan, 2010) and provide the necessary tools to help them to learn more effectively (Abdul Hamid et al., 2018). Second, data-driven AI methods are used to enable more efficient academic management (e.g. automatically create the schedules for teachers, support teachers in grading, provide 24/7 support via chatbots, etc.) and to evaluate the quality of the education (Nieto et al., 2019).

Conversely, the potential risks of the use of AI in education would need to be further studied. Such risks include the violation of privacy, the subliminal manipulation of the students' behaviours via personalised algorithms, different kinds of discrimination and potential negative effects on the students' physical and mental health along with their behavioural development (Zanett et al., 2019).

SDG 5 – Gender equality

Gender equality is a fundamental human right. It is also a foundational element to achieving a more sustainable, peaceful and prosperous world. In recent decades, significant advances towards gender equality have taken place: today, more girls attend school and fewer girls are forced into early marriage, more women occupy leadership positions and legislation is being approved to advance gender equality. Despite this progress, women continue to be underrepresented in leadership positions, one every five women and girls aged 15-49 reports having experienced physical or psychological violence within a 12-month period and discriminatory cultural and social norms and laws remain pervasive.

AI methods can be used to automatically identify gender bias (Feldman and Paeke, 2021), analyse the role of women in meetings through speech recognition, natural language processing and conversation analysis and automatically identify differences in gender representation, coverage and gender biases in newspaper coverage or commercial films via text, image, video and speech analysis (Jang et al., 2019; Kagan et al., 2020).

Data-driven AI decision-making systems are not exempt from limitations, including gender discrimination and bias, as later described in this chapter. Hence, while AI can help us better diagnose and fight against gender inequality, it might also contribute to perpetuating or even exacerbating pre-existing patterns of inequality. Thus, it is of paramount importance to ensure that the AI tools that we use provide non-discrimination guarantees.

SDG 6 – Clean water and sanitation

Access to clean water and proper sanitation are necessary to ensure adequate living conditions. However, according to a report by the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF)[16], in 2019 more than 785 million people did not have access to at least basic water services and more than 884 million people did not have access to clean, safe water to drink. Moreover, more than 2 000 million people worldwide did not have access to basic sanitation and approximately 3 000 million people lack proper facilities to safely wash their hands at home. Here, sub-Saharan Africa is the most affected region in the world with 75 % of the population lacking basic handwashing facilities.

16 https://www.unicef.org/reports/progress-on-drinking-water-sanitation-and-hygiene-2019

The consequences of such statistics are daunting: annually there are 1700 million cases of diarrhoea among children younger than 5 years old, causing the death to 440000 children; 3 million cases of cholera resulting in 95000 deaths and 11 million cases of typhoid fever which cause 129000 deaths. Parasitic worms found in contaminated soil and associated with a lack of proper sanitation facilities infect around 1500 million people worldwide. In addition, a lack of access to adequate sanitation facilities for girls reaching puberty makes them significantly more likely to miss school than boys.

SDG 6 aims at achieving universal access to drinking water at affordable prices and to proper sanitation services; improving the quality of water, reducing pollution; making an efficient and sustainable use of water resources, to be managed in an integrated manner; protecting and re-establishing water ecosystems and increasing the international cooperation in the context of water and sanitation.

Data-driven AI methods have been used to optimise and predict the efficacy of water desalination plants (Dargam et al., 2020), to predict groundwater levels in coastal aquifers (Yoon et al., 2011) and/or their salinity (Sahoura et al., 2020), to model groundwater level changes in agricultural regions (Sahoo et al., 2017), to detect and track major sources of water contamination (Wu et al., 2021) – including drinking water networks (Dogo et al., 2019), to forecast wastewater quality indicators (Granata et al., 2017), to automatically detect water leakage (Kang et al., 2018) and cyber-attacks (Chandy et al., 2017) in water distribution systems and hence avoid wasting water, to predict water consumption in cities and thus better anticipate demand (Brentan et al., 2017) and to forecast water levels in multiple temperate lakes (Zhu et al., 2020), which is a vital indicator of the health of the lake ecosystems and their management. Precipitation and extreme water-related events which could lead to floods can also be modelled using state-of-the-art data-driven AI methods, as described below in the context of SDG 16 (climate action).

SDG 7 – Affordable and clean energy

Access to affordable and clean energy is undoubtedly essential to progress, yet it is one of the biggest challenges that the world faces today. It is estimated that 13% of the world's population – mostly located in sub-Saharan Africa and India – does not have access to electricity. In 2016, 3000 million people in the world depended on highly polluting fuels to perform basic, daily tasks, such as cooking. In addition to the environmental and climate impact, the burning of solid fuels (e.g. wood, charcoal, coal, dung and crop residues) poses a serious public health issue, as it fills the houses and huts with smoke that causes pneumonia, heart disease, lung cancer, stroke and chronic obstructive pulmonary disease.

This SDG aims to achieve by 2030 universal access to modern and affordable energy, a significant increase in the production of renewable energy, double the world's rate of energy efficiency, investment in R&D related to clean energy and investment in the necessary infrastructure to provide modern and sustainable energy in developing economies.

AI has a fundamental role to play in the context of SDG 7. In fact, many of its ambitious goals will not be achievable without the help of data-driven AI techniques. Smart grids depend on AI methods to, e.g., predict demand and optimise the maintenance and functioning of the grid (Raza and Khosravi, 2015), to significantly increase the grid's reliability and efficiency via the automatic detection of failures (Mishra and Rout, 2018) and cyberattacks (Karimipour et al., 2019) and the prediction of load (Hosei and Hosein, 2017). Semi-autonomous or fully autonomous robots are and will be used to inspect and maintain renewable energy plants (Iqbal et al., 2019), such that they could be placed in remote or dangerous locations yet with optimal energy production prospects.

The application of AI in the nuclear engineering domain has been limited to date. However, in recent years several authors have proposed using machine learning and deep neural networks to predict the behaviour of nuclear reactors, perform predictive maintenance of nuclear infrastructures or improve fire hazard models (Fernandez et al., 2017).

Finally, there are numerous examples of how data-driven AI methods are key enablers to creating efficient renewable energy (wind, solar, geothermal, hydro, ocean, bioenergy and hybrid) systems by providing accurate predictions of the expected behaviour of the renewable energy source and hence enabling the optimisation of the energy generation systems (Jha et al., 2017).

SDG 8 – Decent work and economic growth

The creation of high-quality jobs is still a challenge in most countries in the world. While the global unemployment rate is estimated to be 5.5%, there are many countries in both the developing and the developed world where having a job is no guarantee of being above the poverty line or having a decent life.

AI is having and will have a profound impact in the labour market and economic growth. The adoption of AI will affect a wide range of professions, including those that require high levels of qualifications (Mitchell and Brynjolfsson, 2017) in the medical (Barlow, 2016), finance (Dunis et al., 2016), legal and education (Woolf et al., 2013) sectors.

There are numerous studies that have analysed the impact of AI on the labour market, both in terms of the displacement of entire jobs (Frey and Osborne, 2017) or the automatisation of certain tasks within jobs (Arntz et al., 2016). Most authors concur in predicting a significant level of job or task displacements due to AI automatisation. According to Arntz et al. (2016), the percentage of jobs that are susceptible to being displaced by AI range from 6% in South Korea to 59% in Germany, with an average value for Europe between 45% and 60% (Bowles, 2014). This transformation of the job market could lead to an increase in the polarisation of labour (Autor et al., 2010) and migrations to urban centres, which would contribute to geographic and social inequality (Frank et al., 2018). At the same time, AI techniques have been used and will be used to help reduce inequalities, as explained in the SDG 10 section.

SDG 9 – Industry, innovation and infrastructure

Sustainable economic growth relies on the availability of high-quality infrastructure for, e.g., transport, energy, water supply and communication and ambitious investments in innovation to contribute to prosperity, guarantee competitiveness and enable the ability to tackle future challenges.

Data-driven AI techniques are particularly valuable for monitoring, analysing and predicting failures in existing infrastructure by, for example, analysing aerial images using deep learning and machine learning techniques (Bao et al., 2019; Rafiei and Adeli, 2017; Ren et al., 2020; Xu et al., 2019; Gopalakrishnan et al., 2017) or detecting energy consumption anomalies and the production of pollutants in industry (Xu et al., 2015) and the construction of infrastructure. Digital twins are increasingly used as a digital representation of the physical world, including digital twins to predict the behaviour of large infrastructure, such as bridges (Ye et al., 2019).

Another clear area of impact of AI related to SDG 9 are transportation systems, including the use of data-driven AI methods to predict and better regulate transport flows (Zhao et al., 2019); Yao et al., 2019; Pan et al., 2019; Li et al., 2017), to assist in planning more efficient public transport routes (Saracco, 2018) and to deploy autonomous vehicles for passenger and freight land (Niestadt, 2019), rail (Schut and Wisniewski, 2015) or even aerial transport.

SDG 10 – Reduced inequality

In the last 25 years, total global inequality (i.e. the inequality across all individuals in the world) has been declining[17], meaning that the average incomes in developing economies are increasing at a faster rate than those in developed countries. However, inequality within countries has worsened, such that 71% of the world's population live in countries where inequality has increased. With the 21st century, we are witnessing an unprecedented concentration of income and wealth in the hands of the very few: in 2018, the 26 richest people in the world had as much wealth as the bottom half of the world's population.

While AI has been attributed to contributing to inequality[18] due to algorithmic bias and the 'winner-takes-all' phenomenon associated with technological development, data-driven AI methods can be used to reduce inequality. For example, AI algorithms can improve child welfare systems by automatically identifying when children might be in need of welfare (Schwartz et al., 2017), can foster financial inclusion by building alternative credit scores (San Pedro et al., 2015) or by shedding light on the factors for mobile money adoption (Centellegher et al., 2018), and can drive measurable positive change in the lives of minorities and vulnerable groups[19] and ensure fair decision-making (Zemel et al., 2013).

SDG 11 – Sustainable cities and communities

For the past few centuries, the world has experienced a process of urbanisation, that is, the displacement of the population from rural to urban areas, leading to the creation and growth of towns and cities. More than half of the world's population today lives in urban areas and by 2030 it is expected that this figure will raise to about 5 000 million people[20]. People tend to migrate from rural to urban areas looking for a better life and more opportunities to prosper. Indeed, cities have the potential to contribute to higher levels of well-being, education, resource efficiency and economic growth. However, urbanisation is not exempt from challenges, including inequality and poverty, overcrowding, criminality, energy consumption and environmental impact, pollution, waste generation and lack of appropriate living standards.

Data-driven AI techniques have been used to improve urban planning by estimating urban density from aerial images (Lu et al., 2010), informing decisions related to road (Krol, 2016) and public transport (Mukai et al., 2008; Froehlich et al., 2009), planning traffic, detecting traffic incidents (Dia and Rose, 1997; Dia, 2001) and predicting future traffic conditions (Huang et al., 2014; More et al., 2016) or mobility needs (Held, 2018).

Urban intelligent transport systems are only possible thanks to data-driven AI methods, which lead to safer, more inclusive and efficient public transport (Liao et al., 2018; Yao et al., 2018).

17 https://openknowledge.worldbank.org/bitstream/handle/10986/25078/9781464809583.pdf?sequence=24&isAllowed=y
18 https://www.cs.dartmouth.edu/~ccpalmer/teaching/cs89/Resources/Papers/AIs%20White%20Guy%20Problem%20-%20NYT.pdf
19 https://d4bl.org/about.html
20 https://www.unfpa.org/urbanization

AI pervades modern commercial vehicles, which include AI systems to increase safety in intersections, to detect incoming traffic and pedestrians (Enzweiler and Gavrila, 2009), to avoid collisions by, e.g., detecting inattentive drivers (Mandal et al., 2017), predicting driver manoeuvres (Oliver and Pentland, 2000; Jain et al., 2016), predicting pedestrian behaviour (Wu et al., 2018) or warning drivers when invading other lanes (Kim et al., 2016), and to assist drivers in adverse weather conditions (Tuma et al., 2020).

Smart cities depend on AI. There are numerous initiatives worldwide to realise the vision of achieving smart cities, including projects that analyse data captured by internet-of-things devices to measure and optimise energy consumption, recycling levels, pollution and refuse collection in cities (see, e.g., the Urbo[21] project by Telefonica). Urban safety is a critical area that contributes to the quality of life in cities. Machine learning methods have been applied to automatically detect and predict crime hotspots in cities (Bogomolov et al., 2014). The World Council on City Data provides the Open City Data Portal[22], which enables comparison of different metrics across multiple cities.

The newly created Urban AI[23] is a think tank that proposes ethical modes of governance and sustainable uses of AI in the context of cities. Its focus is to develop and deploy AI systems that embrace the diversity of cultures in the world, to contribute to making cities sustainable and vibrant and to preserve our social contract.

SDG 12 – Responsible consumption and production

This SDG aims at making an efficient use of energy and resources, improving access to basic services, building and maintaining infrastructures that are environmentally respectful and creating well-paid jobs with good working conditions.

It is related to many of the other SDGs, including the goals related to poverty, hunger, gender equality, clean water and sanitation, affordable and clean energy, decent work, industry innovation, climate action and reduced inequality.

Thus, only areas of AI impact that complement those described in the sections corresponding to the rest of SDGs are highlighted here.

In terms of contributing to a sustainable and responsible use of natural resources, beyond the impact of AI in the context of renewable energy and agriculture, data-driven AI methods can be used to forecast consumption patterns yielding more efficient production systems with minimal excess production, to automatically create land-use maps to provide a more accurate picture of the state and actual use of natural resources (Talukdar et al., 2020) or to estimate the impact of logging in forests to optimise the logging processes and ensure their sustainability (Hethcoat et al., 2019).

According to the UN Environment Programme[24], approximately one third of the food produced in the world for human consumption is wasted or gets lost, accounting for almost USD 1 000 million globally. Hence, reducing food waste is an important endeavour.

21 https://smartcity.telefonica.com
22 https://www.dataforcities.org/data-portal
23 https://urbanai.fr/
24 https://www.unep.org/thinkeatsave/get-informed/worldwide-food-waste (retrieved in July 2021)

Household waste can be minimised thanks to machine learning methods applied to internet-of-things captured data (Dubey et al., 2020) and the factors that determine household food waste behaviours can be automatically modelled and understood via data-driven AI methods (Setti et al., 2016). Regarding other types of waste generation, machine learning methods can be applied to, for example, predicting solid waste in municipalities and hence enabling more efficient waste planning (Kannangara et al., 2018).

AI enables smart production systems (Petrillo et al., 2020) that, e.g., minimise energy consumption, anticipate demand, detect manufacturing failures, automate tasks and perform systematic evaluations to detect areas of improvement. Digital twins can also be used to optimise production systems via machine learning methods (Min et al., 2019).

Finally, socially responsible consumption and disposal behaviour can be inferred automatically via machine learning algorithms (Song et al., 2018), This information could be used to foster and reinforce consumer behaviours that contribute to sustainability.

SDG 13 – Climate action

The potential of AI to help address the climate emergency is unquestionable (Rolnick et al., 2022). In fact, we will not be able to combat climate change without the help of AI.

Data-driven AI methods are used to model climate and weather, identify patterns and make accurate predictions based on the analysis of multi-dimensional weather and climate datasets (Haidar and Verma, 2018; Ham et al., 2019).

Deep learning models have been used to represent sub-grid processes in climate models (Rasp et al., 2018), to predict global temperature changes (Ise and Oba, 2019) and weather (Weyn et al., 2020) and to model weather phenomena, such as rainfall (Sonderby et al., 2020). In addition to being used to build more accurate climate models and predictions, AI methods can also be applied to improve state-of-the-art weather modelling systems by enabling, e.g., the separation of noise in climate observations (Barnes et al., 2019) or the automatic labelling of climate data (Chattopadhyay et al., 2020).

Extreme weather events are increasing in frequency and intensity due to climate change. AI has also proven to be a valuable ally to predict extreme weather events and their impacts, such as heavy rain (Lee et al., 2020), hail (Gagne II et al., 2019), wildfires (Radke et al., 2019), floods (Pastor-Escuredo et al., 2014) and earthquakes (Wang et al., 2020) and to enable a more efficient, prompt response to natural disasters. Autonomous drones have been used to monitor heat and prevent fires (Allison et al., 2016) and to search for survivors in floods and earthquakes (Arntz et al., 2016). In this domain, the AI for Disaster Response (IADR)[25] project at Qatar Computing Research Institute (QCRI) provides a free online tool that analyses social media messages related to emergencies, humanitarian crises and disasters. It uses machine learning to tag up thousands of messages per minute automatically, acting as an early warning system.

Beyond the direct application of AI techniques to model and predict climate, AI methods may be applied to industries or sectors that have a negative environmental impact to enable the reduction of greenhouse gas (GHG) emissions.

25 http://aidr.qcri.org

According to a report commissioned by Microsoft from PwC[26], the use of AI in environmentally related use cases could contribute up to USD 5.2 billion to the global economy by 2030 while reducing GHG emissions by 4%, which is equivalent to the 2030 estimated annual emissions of Japan, Canada and Australia combined.

Examples of such scenarios include using AI methods to yield more efficient energy generation – particularly in highly polluting sectors, such as the petrochemical sector (Han et al., 2019) – and to better manage the electric grid by means of accurate energy consumption forecasts (Almalaq and Edwards, 2017).

Data-driven AI approaches could also be used to accurately predict both carbon emissions and the factors contributing to them (Huang et al., 2019), thus enabling prompt action.

Moreover, there are major private and public institutional programmes aimed at exploring the use of AI to help combat climate change. In Europe, the Cordis database of funded research reveals over 100 funded projects related to AI and climate change, covering topics that range from the detection of extreme events to using AI to accelerating the transition of cities to carbon neutrality by means of AI. The European Space Agency has launched the Digital Twin Earth[27] to accelerate the identification of solutions to predict the impact of climate change. The European Lab for Learning and Intelligent Systems (ELLIS), one of Europe's leading AI associations, has launched a research programme on machine learning for Earth and climate sciences that aims to 'model and understand the Earth system via machine learning methods'.

In the private sector, most technology companies have deployed initiatives aimed at using AI to help combat climate change.

For example, the Canadian AI company ElementAI has launched a climate programme[28] as a cross-company initiative to support private and public sector efforts that tackle the climate crisis and help to build a sustainable and resilient future; Microsoft's AI for Earth initiative[29] is a 5-year USD 50 million endeavour to put Microsoft's cloud and AI tools in the hands of those working to solve global environmental challenges; in October 2020, Facebook announced a partnership with Carnegie Mellon University[30] to assist scientists in using AI tools to develop renewable energy and combat climate change; and Google's 'AI for social good' programme recently issued an open call[31] to organisations around the world to submit their ideas for how they could use AI to help address societal challenges. Among the 20 organisations that are supported by Google, there are projects related to using AI to estimate emissions of fossil fuel in power plants.

Conversely, data-driven AI systems have a significant CO_2 footprint contribution which would need to be systematically measured and mitigated, as described in the next section.

26 https://www.pwc.co.uk/sustainability-climate-change/assets/pdf/how-ai-can-enable-a-sustainable-future.pdf
27 https://www.esa.int/ESA_Multimedia/Images/2020/09/Digital_Twin_Earth
28 https://www.elementai.com/ai-for-climate
29 https://www.microsoft.com/en-us/ai/ai-for-earth
30 https://www.cnet.com/news/facebook-plans-to-use-ai-to-help-fight-climate-change/
31 https://ai.google/social-good/impact-challenge/

SDG 14 and SDG 15 – Life below water and life on land

Healthy oceans, seas and land are essential to ensure the necessary living conditions on our planet. However, the quality of the waters and terrestrial ecosystems has significantly worsened in the past decades.

Regarding waters, the acidity of the oceans – which is key for climate regulation and to sustain entire ecosystems – is expected to increase by 100% to 150% by the end of the 21st century, according to the current trends. Moreover, each year at least 14 million tons of plastic end up in the oceans[32], which is 80% of all marine debris, threatening life in the oceans, human health, food safety and quality and contributing to climate change. In terms of land, each year tens of millions of hectares of forests and natural terrestrial environments disappear because of logging, wildfires, desertification due to climate change and human intervention.

Advances in computer vision (object detection in images and videos, image classification) together with other data-driven AI methods can be used to automatically monitor the quality of our oceans and our land.

For example, deep learning methods have been used to estimate the volume of plastic debris in coastal areas (Martin et al., 2018), detect oil spills (Jiao et al., 2019) or estimate the CO_2 flux (which plays an important role in ocean acidification) in the oceans by analysing aerial images (Chen et al., 2019).

Similarly, deforestation (de Bem et al., 2020), forest quality (Zhao et al., 2019), aboveground biomass (Madhab Ghosh and Behera, 2018) and the risk of wildfires (Oulad Sayad et al., 2019) can be automatically estimated via deep neural networks applied on aerial images alone or combined with other data sources.

Illegal wildlife trade can be automatically detected by analysing social media data via machine learning methods (Di Minin et al., 2019) and wildlife species can be automatically classified using deep neural networks on aerial or motion-activated camera images (Tabak et al., 2019).

AI also enables **smart fishing**, which combats overfishing and fosters sustainable fishing by the automatic classification of species, biomass estimation, prediction of the quality of the water and of the behaviours of aquatic animals; together with **precision agriculture** and **smart farming**, as described in the SDG 2 section.

SDG 16 – Peace, justice and strong institutions

Conflicts, insecurity, weak institutions and limited access to justice are clear barriers for sustainable development. While overall the world population is healthier, better connected and wealthier than ever before, there are numerous places in the world where people's lives are severely impacted by wars and insecurity, a lack of access to fair justice systems and the violation of human rights.

32 https://www.iucn.org/resources/issues-briefs/marine-plastics

According to the International Committee of the Red Cross (ICRC)[33], it was estimated that in 2018 roughly 2 000 million people in the world were affected by conflict, violence or fragility and by 2030 these people will most likely endure extreme-poverty living conditions. Approximately 120 million people worldwide depend on humanitarian aid. A recent report by the UN refugee agency (UNHCR) estimates that a record number of 80 million people in the world were displaced in 2020 by wars and violence, including almost 30 million refugees.

Data-driven AI techniques can be used to accelerate and promote peace, safety, justice and stronger institutions. For example, institutional corruption can be detected automatically by data-driven machine learning algorithms applied to financial transactions (Chang-Tien and Siriat, 2004; West and Bhattacharya, 2016; Hajek and Henriques, 2017), public tender processes (Lismont et al., 2018) and government corruption (Adam and Fazekas, 2018). In addition, institutions may significantly increase their efficiency by applying AI techniques that enable the complete or partial automatisation of administrative tasks and processes (Etscheid, 2019).

Mathematical tools have been used to detect and predict crime for decades, and today many of such techniques include data-driven AI methods. Machine learning methods can be used to identify illegal drug trafficking (Baveja et al., 1997; Li et al., 2019) and crime hotspots in cities (Bogomolov et al., 2014); and semantic and natural language processing techniques have been applied to social media content to detect extremist behaviours (Johansson et al., 2017).

Without a doubt, the domain where AI is playing a crucial role is in the detection and prevention of cybercrime (Siddiqui et al., 2018), which, increasingly leverages AI methods as well.

33 https://www.icrc.org/en/document/global-trends-war-and-their-humanitarian-impacts-0

3. Limitations and barriers

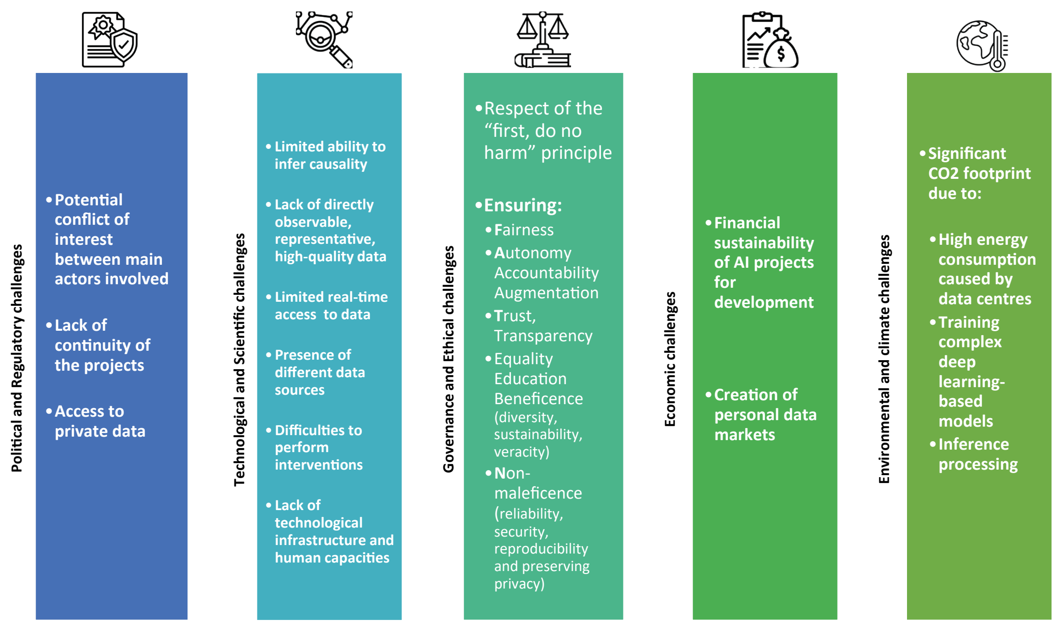
As previously described, the opportunities for data-driven AI methods to help us better measure and accelerate the achievement of the SDGs are paramount. Data- or evidence-informed decision-support systems, where machine-learning algorithms play a central role, have been referred to by some authors as 'human AI' systems (Letouze and Pentland, 2018). The concept of human AI systems is aligned with, but broader than, the 'human-centric AI' approach adopted by the European Commission.

Human-centric AI refers to designing and deploying AI systems that are aligned with core human values. Human AI systems add to this concept a vision where humans – alone or supported by AI systems – can make more informed, evidence-based decisions thanks to AI, supporting behaviours and decisions that are likely to yield positive outcomes while discouraging those that would not.

However, despite this immense opportunity, today's reality is far from this vision. To date, there have been few successful examples of real-world systems which **systematically** leverage large-scale data and AI methods to support humans in making better decisions for the public good. More than a decade into the (big) data revolution and half a decade into the 17 SDGs, major barriers remain, including difficulties related to the access and analysis of valuable data, which in many cases are privately held. In addition, there is lack of well-defined ethical principles, potential legal and regulatory barriers, technical limitations, competing commercial interests and the non-negligible carbon footprint of today data and computation-greedy AI systems.

In this context, there are five types of challenges and barriers that should be considered to ensure that AI is positively used for sustainable development in a safe and ethical manner. Most of these barriers are extensively described in (Letouze et al., 2019), a summary of which is presented next.

Figure 11-1: **Five types of challenges and barriers**



Political and regulatory challenges

The use of data-driven AI methods to support the achievement of the 17 SDGs would typically require the collaboration of three groups of actors: private organisations, public institutions and citizens. These three parties have potentially conflicting interests, constraints and priorities.

Thus, tackling barriers in this political dimension requires striking a balance between the private, public (e.g. governments) and individual interests, which implies understanding their underlying dynamics (Letouze et al., 2015).

The first group of stakeholders are private organisations, which in most legal frameworks are the legal owners or custodians of a significant portion of the data of interest, such as mobile network, financial transaction, satellite, energy consumption, employment or social media data.

The second group of stakeholders consists of the institutions that require access to the data to derive meaningful insights from it in the context of one or more of the 17 SDGs. Such institutions could be governmental – e.g. ministries, regional or local governments and national statistical offices (NSOs), academia or civil society organisations. In the case of NSOs, there is a strong movement related to using non-traditional data sources to compute official statistics, for example to estimate population density or poverty in a more efficient and frequent manner. However, there are very few examples of such a use in a systematic manner. While the potential value of data to help NSOs to build a more accurate picture of reality is clear, appropriate consultation and technological and governance safeguards are of critical importance to minimise the risks related to potentially breaking the citizens' trust, alienating private companies, breaching individual or group privacy and/or impacting the reputation of the institutions involved, particularly if the use of the data yields unintended negative consequences.

An example of reputational impact is the negative press received by a project launched by the Spanish National Statistics Institute, where they analysed aggregate insights derived from mobile network data from the three largest telcos in Spain without the knowledge or explicit consent of mobile users[34].

Incidentally, the project later became instrumental during the COVID-19 pandemic as it enabled teams of experts, working in collaboration with Spanish policy-makers, to model large-scale human mobility[35] and thus measure the compliance and impact of the confinement measures on the population's behaviour and the spread of COVID-19.

Another barrier in this regard relates to a potential lack of continuity of the projects, particularly if there are no guarantees that the necessary data and/or resources will be available over time. Specific regulations and multi-year partnerships could help address these concerns.

Finally, there are the individuals whose data is already analysed for many (commercial) purposes, in principle with their consent but possibly – or probably – not with their understanding. Key principles and rights – such as fairness, transparency, autonomy, veracity, reproducibility, reliability, control and privacy – would need to be demonstrably preserved. In Europe, the new proposal for a regulation on AI[36] addresses such principles and rights. It is a pioneering example of a legal framework

34 https://www.elconfidencial.com/tecnologia/2019-10-29/ine-operadoras-recopilacion-datos-moviles-proteccion-leyes_2304120/

35 https://infocoronavirus.gva.es/es/grup-de-ciencies-de-dades-del-covid-19-de-la-comunitat-valenciana

36 https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

on AI, formulating a risk-based regulation that positions Europe in a leading role globally with a human-centric approach.

Given the evident monetary value of the data, many authors and projects have proposed the creation of personal data markets (Staiano et al., 2014), where individuals would have control over their own data and decide whom to share it with, for which purposes and at what cost.

In terms of data privacy, the European General Data Protection Regulation (GDPR) and ePrivacy Directive place a premium on obtaining consent from users and require data controllers to implement the necessary measures to allow users to know and keep track of which data and for which purposes is being captured. Note that the GDPR allows for the lawful processing and sharing of privately held data in certain use cases, including the computation of statistics that are no longer considered personal data, which enables the analysis of data for research and policymaking purposes without requiring consent. Ideally, data-centric initiatives for sustainable development would be opted out by default (as opposed to opted in) and should be **opted out** at any time with ease by users, regardless of the intended purpose. A key challenge in this regard concerns obtaining such user consent: data subjects would need to be convinced that it is not only safe, but also in their interests to agree to make their data available for the purposes of public good – either by opting in or not opting out. Given concerns expressed in many countries, significant efforts still need to be undertaken to show evidence of the value – i.e. the positive social impact – that would result from the data analysis and to ensure that the technology and the methods behind it are sound and safe to generate public trust.

Thus, the next set of challenges concern developing the necessary technology and science to enable turning data into reliable, accurate and actionable knowledge.

Technological and scientific challenges

As previously illustrated, data-driven AI methods have tremendous potential to positively contribute to the achievement of the 17 SDGs. However, they are not exempt from technical limitations and risks that could yield negative (un)intended consequences impacting the lives of millions of people.

A particularly important type of risks concerns the computational violation of individual privacy that would result from the analysis of data via data-driven AI methods, even if the data is fully anonymised. Several research works have shown that human individual behaviours are unique. Thus, it is possible to de-identify an individual even when using anonymised and coarsened data (Blondel et al., 2013). Additional research efforts have focused on understanding the limits of human privacy and how it could be protected (Rocher et al., 2019). However, according to the current state-of-the-art, anonymising personal data is not sufficient to ensure the protection of individual privacy.

Differential privacy (Dwork and Roth, 2014) is a promising technical approach to preserving privacy. It consists of performing a statistical analysis of the datasets that may contain personal data, such that when observing the output of the data analysis, it is impossible to determine whether any specific individual's data was included or not in the original dataset. The behaviour of an algorithm applied to a differentially private dataset is guaranteed not to change when an individual is present or not in the dataset. This guarantee holds for any individual and for any dataset. Hence, regardless of the specific details of an individual's data (even if such an individual is an outlier), the guarantee of differential privacy should still hold.

Beyond privacy, there are additional technical and scientific challenges related to the analysis of data via AI methods that need to be addressed, including: the frequent lack of ground truth[37] that would enable the proper validation of supervised, data-driven AI models applied to tackle the 17 SDGs; difficulties with real-time access and analysis of the data, despite the fact that in many impactful use cases within the SDGs real-time access would be imperative (e.g. helping in the early detection of pandemics; predicting a natural disaster or supporting an immediate, proportionate response to natural disasters or emergencies, etc.); complexities derived from having to combine datasets from different sources; the difficulty of inferring causality – but rather identifying correlations – with the implications that this limitation may have for policy- and decision-making; the potential lack of representativeness of the available data, its generalisation capabilities and inherent biases; the lack of certification standards to guarantee the quality of the algorithms applied to the data, including non-discrimination guarantees; limited transparency, explainability and interpretability of complex machine-learning (notably deep learning-based) algorithms that might be applied to tackle a certain SDG; questions about the quality and veracity of the data; and difficulties in ensuring the reproducibility of results as they heavily depend on the data used to train the AI models and the parameter setting using when training.

Further technical challenges derive from the lack of the necessary technological infrastructure and human capacities to systematically store, analyse and effectively apply the insights derived from the data analysis.

Thus, appropriate investments in technical infrastructure and human resources are necessary to successfully realise the potential that data-driven AI methods have in the context of the 17 SDGs. Importantly, such resources would need to be allocated **prior** to the inception of any project. Given that the underlying reality is extremely complex and dynamic, projects would need multi-disciplinary teams of experts, including local talent, devoted fully to the projects on a continuous basis and located in the countries/regions where the projects are deployed.

Many of the scenarios where AI could enable and accelerate the achievement of the SDGs are in areas of consequential importance in people's lives, such as healthcare, education or immigration. Thus, the third critical set of barriers to overcome relate to the governance and ethical challenges derived from using data-driven methods to support human decision-making.

Governance and ethical challenges

Numerous governance challenges and ethical dilemmas emerge when applying data-driven AI methods to support decision-making processes and systems with impact on the lives of millions of people.

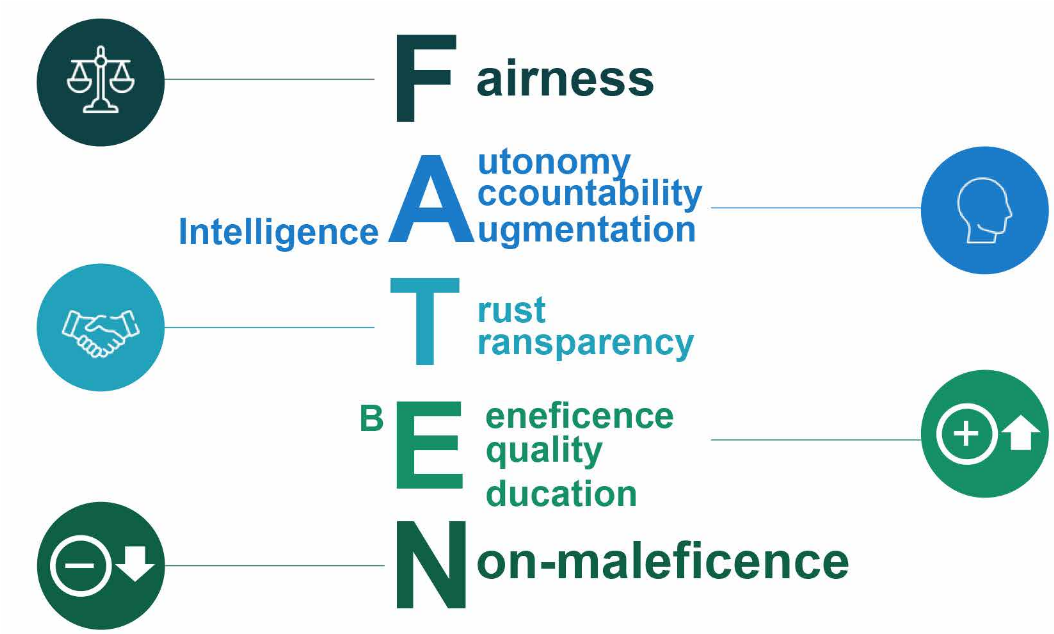
In this context, the 'first, do no harm' principle used in humanitarian scenarios is particularly relevant. Today, we have a much better understanding of the risks – even in the case of well-intentioned projects – that AI poses to human autonomy, privacy, equality, dignity, fairness and transparency than we did a decade ago. How can we be sure that applying AI to support the achievement of the 17 SDGs will do no harm? Will data-driven decisions used in this context be outside of our control? Who is accountable for such decisions, particularly in cases where they may be the result of analysing multiple datasets by complex software and social systems developed by potentially different parties? Will these systems include the necessary security mechanisms to prevent cyberattacks? What about the malicious use of

37 Ground truth refers to data obtained by direct observation (i.e. empirical evidence) as opposed to obtained by inference (e.g. by a machine learning algorithm). In supervised and semi-supervised machine learning, ground truth is needed to train and validate the models.

the data to serve the interests of non-democratic governments or organised crime? These are complex questions to tackle. Thus, ethical principles and standards of governance for data-driven initiatives for public good need to be both clearly defined and meticulously complied with.

In the past decade, many proposals have been published related to the ethical guidelines and principles to apply to the broad use of AI in society. Such proposals include the principles of the Menlo Report (Dittrich and Kenneally, 2012); the ethical principles included in the national AI strategies of over 50 countries in the world; the report by the European Commission for the development of trustworthy AI (European Commission, 2019); the OECD[38] principles for the development of AI; and the ethics in AI initiatives within professional organisations, such as the Institute of Electrical and Electronics Engineers (IEEE)[39] and the Association for Computing Machinery (ACM)[40]. Most of the previously proposed principles might be grouped using the FATEN (Oliver, 2019) acronym, which is an extension of the four basic principles of medical ethics (Gillon, 1994).

Figure 11-2: **The FATEN principles**



38 OECD, OECD Principles on AI, OECD, Paris, France, 2019.

39 The Institute of Electrical and Electronics Engineers, Ethically Aligned Design, IEEE, Piscataway, NJ.

40 Association for Computing Machinery, Code of Ethics and Professional Conduct, ACM, New York, NY, 2018.

F for fairness, i.e. without discriminating. Data-driven AI systems might discriminate for several reasons, including biases in the data used to train the algorithms, an inappropriate choice of an algorithm or model for the problem at hand, and a biased interpretation of the results. In the past 5 years, many highly impactful cases of algorithmic discrimination in social good areas have been made public, such as in the areas of criminal justice (Angwin et al., 2016), credit granting (Blattner and Nelson, 2021), human resources and hiring[41], education[42] and healthcare (Ledford, 2019). The detection and measurement of algorithmic bias and the development of fair machine-learning algorithms are fertile areas of research, as illustrated by the newly created ACM Conference on Fairness, Accountability and Transparency (ACM FAccT)[43], the ELLIS research programme on human-centric machine learning[44] or the newly created Institute of Humanity-centric AI[45] in Spain, which is one of the 34 ELLIS units launched since December of 2019.

A for autonomy, **accountability and intelligence augmentation**. The principle of autonomy is at the core of Western ethics. According to this principle, every person should be able to freely choose their own thoughts and actions. However, using data-driven AI methods today we can build computational models of our personalities, interests, tastes, needs, strengths/weaknesses and behaviour that could be – and probably are – used to subliminally influence our decisions, choices and actions.

Thus, we should ensure that AI systems that have a direct or indirect impact on people's lives always respect the principles of human autonomy and dignity. The letter A in FATEN also stands for accountability, i.e. having clarity with respect to the attribution of responsibility related to the consequences of using AI methods.

Finally, A stands for intelligence augmentation – rather than replacement: AI systems should be used to support and augment human decision-making and not to replace humans altogether. This view is fully aligned with the previously described human AI concept.

T for trust and transparency. Trust is a fundamental pillar in our relationships, not only with other humans but also with/between institutions. Trust is typically established in the context of a specific purpose. We might trust an institution or an individual to be custodians of our money, but not necessarily of our children, for example. Trust emerges when three conditions are met:

- competence, i.e. the ability to successfully carry out the committed task;
- reliability, i.e. sustained competence over time;
- honesty and transparency. Hence, the T in FATEN is also for transparency.

A data-driven decision-making system is transparent when non-experts can observe it and easily understand it. Data-driven decision-making systems might not be transparent for at least three reasons (Burnell, 2016):

- **intentionally**, to protect the intellectual property of the system's creators;

41 https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G (retrieved in July 2021)

42 https://www.brookings.edu/blog/the-avenue/2019/09/26/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases/ (retrieved in July 2021)

43 https://facctconference.org/

44 https://ellis.eu/programs/human-centric-machine-learning

45 https://ellisalicante.org

- due to the **digital illiteracy** of their users, which prevents them from understanding how the models work;
- **intrinsically**, given that certain data-driven AI approaches – particularly deep learning methods – are extremely complex and difficult to interpret.

Transparent, interpretable and explainable AI models are necessary in most of the use cases related to the SDGs.

E for bEneficence and equality. The principle of bEneficence refers to maximising the positive impact in the use of data-driven decision-making algorithms with sustainability, diversity and veracity. We cannot obviate the environmental cost of technological development, particularly when it comes to AI algorithms, given their need for large amounts of data to learn from and massive amounts of computation needed to process and be trained by such data. As this is a fundamental challenge, it is described later in more detail.

Diversity is also of paramount importance, from at least two perspectives. First, by ensuring that the teams developing data-driven AI systems that are used for sustainable development are diverse, which is not the case today. Diversity is needed to maximise the probability of finding innovative solutions to the immense challenges that we face – as diverse teams tend to be more innovative than non-diverse teams[46] – and of developing inclusive solutions that would be relevant in the communities where they will be deployed. Second, by incorporating diversity criteria into the algorithms we design, we can minimise the prevalence of filter bubbles and echo-chamber effects (Geschke et al., 2019) which might contribute – at least partially – to the polarisation of public opinion.

We also need to ensure the **veracity** of the data that is and will be used for sustainable development scenarios. Today, we can algorithmically generate fake text, audio, photos and videos by means of deep neural networks (**deep fakes**) that are indistinguishable to humans from real content. If we are using data to inform decisions that impact the lives of millions of people, we need to ensure that such data is indeed truthful and a reflection of the underlying reality that the models are attempting to model.

E also stands for equity. The development and wide adoption of the internet and the World Wide Web during the Third and Fourth Industrial Revolutions has undoubtedly been key to democratising the access to information. However, the original principles of universal access to knowledge and the democratisation of technology are in danger today due to the extreme dominance of technology giants in the USA (Apple, Amazon, Microsoft, Facebook and Alphabet/Google) and China (Tencent, Alibaba, Baidu). Together, these non-European technology companies have a market value of more than USD 5 trillion and a US market share of more than 90% in internet searches (Google), more than 70% in social networking (Facebook) and 50% in e-commerce (Amazon). This market dominance leads to data dominance. In fact, most of these technology companies are data companies that earn thousands of millions of dollars by analysing and monetising the data they collect about their users. Note that a significant portion of the valuable human behavioural data that could be used in the context of the 17 SDGs is generated and captured by the services that these technology companies offer to their customers – services that address many aspects of our lives, including our entertainment, work, health and wellbeing, sports, education, transportation, travel, social connections, communication, shopping, information and product needs.

46 https://www.forbes.com/sites/forbesinsights/2020/01/15/diversity-confirmed-to-boost-innovation-and-financial-results/?sh=2e02e09bc4a6 (retrieved in July 2021)

In addition, in the 21st century we are observing a polarisation in the distribution of wealth, as described in the context of SDG 10. According to the Global Wealth Report 2019 by Credit Suisse (Shorrocks and Hechler-Fayd'herbe, 2019), the 100 richest people in the world are richer than the poorest 4000 million. This accumulation of wealth in the hands of very few has been at least partially attributed to technology and the Fourth Industrial Revolution. With the agrarian revolution in the Neolithic and for thousands of years afterwards, wealth was associated with ownership of land. Following the First Industrial Revolution, wealth was a result of owning capital assets, such as machines and factories. Today, one could argue that data – and more importantly, the ability to leverage it and make sense of it – is the asset that generates the most wealth, generating what is known as the data economy. Thus, if our goal is to maximise the positive impact of this abundance of data, we should develop and promote new models of data ownership, management, exploitation and regulation. Data used for sustainable development could contribute to both better measuring and reducing inequality (see the SDG 10 section).

N for non-maleficence. This means minimising the negative impact that might result from the use of data-driven AI methods. Within this principle, we include being prudent in the development of AI-based systems and highlight the need to:

- provide reliability and reproducibility guarantees
- maximise data security
- always preserve people's privacy, as previously discussed.

Once agreed upon, the ethical principles will need to be published, implemented and complied with in practice through appropriate governance. The roles and responsibilities of each of the three actors – namely, companies, public and non-profit institutions, and people – need to be clearly defined, understood and accepted.

Given the multi-disciplinary nature of data-driven projects for public good, a combination of experts from different disciplines – ranging from AI to social sciences and humanities experts – is required for the projects to succeed. This multi-disciplinary nature adds complexity, but it is necessary and particularly beneficial when it comes to the definition of, and compliance with, ethical principles since the teams would include ethicists.

Moreover, external oversight bodies are also desirable to ensure that the ethical principles are complied with. Data stewards[47] have been proposed in recent years for this purpose. Data stewards are individuals or groups of individuals within an organisation who are responsible for the quality and governance of data in data-driven projects that take place in their organisations, including initiatives for social good. Alternative options include the creation of external oversight ethics boards and/or the appointment of a chief ethics officer with oversight and auditing responsibilities to ensure that projects with social impact are aligned with the pre-defined ethical principles and human values of the societies where they are developed.

Another approach to ensure compliance with the ethical principles agreed upon is by requiring the use of open processes, code and systems, by deploying regulation that requires the ethical principles to be followed and/or by fostering knowledge sharing, including collaborations with academia and civil society organisations.

In addition, understanding the cultural and social characteristics of the societies where the projects are deployed is a must. Therefore, working with local institutions and the civil society of the

47 Verhulst, Steefaan G., The Three Goals and Five Functions of Data Stewards: Data Stewards: a new Role and Responsibility for an AI and Data Age, Medium and The Data Stewards Network, New York, NY, 2018.

countries where the projects will take place is absolutely necessary, as previously highlighted.

In sum, any use of data-driven AI methods for sustainable development should be open, transparent, accountable and always respectful of human values and rights. The results of the projects should be auditable regarding their purpose, accuracy, reproducibility, veracity and fairness, particularly given the fact that the use of AI in the context of the 17 SDGs is an overly broad, ambitious, long-term and multi-institutional endeavour.

Even when the political, technological and ethical challenges are addressed, projects that leverage data-driven AI for public good might fail if they lack a sustainable financial model. Hence, the fourth type of challenges is of an economic nature

Economic challenges

Many initiatives that have applied data-driven AI methods to support the achievement of the 17 SDGs have been in the form of pilots. Questions inevitably arise about the generalisation capability and the financial sustainability of such projects.

Several companies that have been at the forefront of the 'data and AI for social good' movement over the past 10-15 years – particularly telecommunication operators such as Telefonica and Orange – have also invested in developing their own related commercial offerings. Recently, technology companies have joined the movement of leveraging their data for purposes related to social good and sustainable development, including Facebook[48] and Google[49]. In developed countries, the granularity, volume and richness of human behavioural data collected by technology companies is undisputed.

The commercial solutions developed by these companies provide user and client pre-computed indicators derived from aggregate customer data, such as population density and mobility estimations. These estimations are also valuable in the context of the 17 SDGs.

Given this overlap between commercial and public interest purposes, companies might resist the development of solutions for sustainable development as they could cannibalise their existing data-driven services. However, there are important considerations to be made in the context of data-driven projects to support the 17 SDGs. As described in the previous section, many of the data-driven AI systems used in the context of the 17 SDGs would need to comply with strict regulations, scientific rigor, ethical frameworks and governance models appropriate to the fact that they will be used for public-good purposes. Such requirements might not apply to the same extent in the case of proprietary, commercial services.

Thus, the value proposition related to projects for sustainable development would need to be defined such that it would be complementary to, and not in competition with, the existing commercial products offered by these companies. Moreover, a sustainable financial model is needed for the projects to succeed beyond their pilot phase. Even if they are for social good, they do not necessarily need to be for free, depending on the use case. This economic dimension is thoroughly discussed in the report by the European Commission's High-level Expert Group on Business-to-Government Data Sharing[50] and in Letouze et al. (2019).

48 https://dataforgood.fb.com/

49 https://cloud.google.com/data-solutions-for-change/

50 https://digital-strategy.ec.europa.eu/en/library/meetings-expert-group-business-government-data-sharing

An additional question related to economic challenges is whether people should be able to sell their own data on a personal data market. The cases for and against such a model can be and have been argued convincingly (Speikermann et al., 2015).

Finally, as previously explained, data-driven AI methods require massive amounts of data and computation, with a potentially significant CO_2 footprint. Thus, the final set of challenges concern the environmental and climate impact of the development and wide deployment of AI in our societies.

Environmental and climate challenges

AI has tremendous potential to help us address the climate emergency (SDG 13), as previously described. However, AI is also a non-negligible contributor to GHG emissions (Garcia-Martin et al., 2019) given the high energy needs of today's data-driven methods. This is for a variety of reasons.

First, a significant factor in the carbon emissions due to the development and deployment of AI systems stems from the energy consumption caused by data centres, given that data centres are a key element in the AI pipeline, hosting the vast amounts of data needed to train and use sophisticated machine learning models. On the positive side, while the demand and size of data centres has been growing steadily in the past years, their energy consumption has not grown proportionally, thanks to the development of energy-efficient infrastructure and hardware (Lei and Masanet, 2021), the use of renewable energy sources and even the application of AI methods to reduce their energy consumption[51].

Nonetheless, a report by the European Commission[52] estimates a 28 % growth in the energy consumption of data centres in Europe between 2018 and 2030. The report includes several recommendations to minimise the GHG emissions attributable to data centres, including recommendations relative to information/awareness raising measures, transparency initiatives, the development of standards and guidelines for energy-efficient cloud computing, soft-certification schemes, the inclusion of energy-consumption labels, the establishment of regulations for the non-material component of data centres and cloud services and of minimum criteria for energy-efficiency in newly built data centres in the EU, policy-awareness raising and the definition of green public procurement criteria and knowledge sharing.

Secondly, we need to consider the GHG emissions due to training complex data-driven AI (deep learning-based) models. OpenAI researchers Dario Amodey and Danny Hernandez estimate that since 2012 the amount of computing power used to train the largest data-driven AI models has been increasing exponentially, with a 3.4-month doubling time (faster than Moore's Law 2-year doubling period)[53]. A recent study (Strubell et al., 2019) found that the carbon footprint of training just one state-of-the-art deep-learning model to perform natural language processing tasks was equivalent to the amount of carbon dioxide that the average American produces in 2 years. In fact, the energy costs associated with training sophisticated machine learning algorithms has traditionally been the most expensive task when using AI to solve real-world problems.

51 https://research.google/pubs/pub42542/ (retrieved in July 2021)

52 https://digital-strategy.ec.europa.eu/en/library/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market (retrieved in July 2021)

53 https://openai.com/blog/ai-and-compute/ (retrieved in July 2021)

Third, we have the GHG emissions caused by inference processing, i.e. using a trained data-driven AI model on new, unseen data, which has grown tremendously, representing 80-90% of the cost of neural networks, according to Nvidia.

While the growth in AI-related energy consumption is partly mitigated by hardware-aware models (Marculescu et al., 2018) and energy-efficient hardware that has been specifically designed to train data-driven AI models (deep neural networks) – such as FPGAs and ASICs, there is an urgent need to implement systematic and accurate measurements of the carbon footprint of AI systems to ensure that their positive impact is larger than their environmental cost, creating what some authors refer to as green AI (Schwartz et al., 2020). Note that understanding the carbon footprint of AI entails more than measuring the energy consumption of data centres, model training and inference activities. In fact, given the broad set of use cases where AI is having an impact and the complex, multi-layered proprietary production process of AI systems, assessing the carbon footprint of AI is certainly challenging.

Thus, several authors have recently focused on assessing the carbon footprint related to AI research and have built tools to ease its measurement (see e.g. the **experiment impact-tracker** (Henderson et al., 2020) and the **machine learning emissions** calculator (Lacoste et al., 2019) projects), given that research methods and results are generally openly available via scientific publications. Even in this case, there is a lack of systematic carbon emission measurements of AI research (Cowls et al., 2021).

While these recent works are promising and reflect an increased interest in ensuring that the GHG emissions due to AI are minimised, current practices both in research and industry are far from what these research papers propose.

From the areas of opportunity highlighted in Section 2 and the challenges just described, several recommendations emerge to accelerate the positive impact of AI on the SDGs (and thus on our societies and the planet itself) while minimising its potential negative impact.

4. Recommendations

In this section, I formulate key recommendations related to each of the barriers described above to accelerate the achievement of the SDGs thanks to AI.

Data. Data is a fundamental asset for the SDGs. First, as a digital representation of an underlying reality that we need to measure so that we can assess the level of achievement of each SDG. Second, as a key element to enable the development of data-driven AI methods to find patterns, make predictions, detect outliers, automate tasks, etc. Thus, first and foremost, we should develop ambitious programmes to enable access to high-quality, relevant data, and invest in secure frameworks that provide access to data and/or actionable insights derived from the data, even when the data is privately held. Support for more effective and accessible use of existing datasets is also important, as many existing datasets are not properly leveraged due to difficulty of access. Finally, data gaps would need to be identified and actions to fill them would need to be taken.

R&D. The opportunities in the intersection of AI and the SDGs are immense. However, most of these opportunities still entail significant investment in research. Hence, ambitious and sustained investment in research and innovation on the topic of AI for sustainable development would be of paramount importance if we want to leverage the potential of AI to help us to achieve the SDGs. Moreover, many of the promising results have been achieved in small-scale studies with offline data. There is a lack of large-scale, real-world evidence of the systematic and sustained use of AI to support the achievement of the SDGs. Therefore investments to leverage research and pilot results and deploy them in the wild over long time periods are necessary.

Vulnerability analysis. As societies increasingly rely on AI systems, it becomes important to carry out vulnerability analyses of such dependencies and to deploy redundancy and backup systems to be able to gracefully recover in case of failures, malfunctioning or hacking of the AI systems.

Governance. Promote corporate governance and engagement models – including the appointment of data stewards, chief ethics officers and oversight boards – in public administrations, NGOs and private companies working on AI for SDG projects. Adopt and evaluate compliance with ethical frameworks to ensure that the use of the AI systems deployed to support the SDGs are aligned with the FATEN framework previously described.

Openness and transparency. Develop open, participatory systems and standards to enable data- and knowledge sharing across companies, sectors and countries with inputs and oversights from relevant stakeholders.

Education. Invest ambitiously in education, capacity building and outreach to obtain the support and contributions from all private and public actors (including citizens) in Europe and beyond. The development of local capacities would be of paramount importance to ensure the sustainability and actual impact of the projects.

Multi-disciplinary projects and diverse teams. Foster multi-disciplinary projects where AI experts collaborate closely with domain experts and policy makers to maximise the opportunities to have impact.

Best practices and centres of excellence. The recently created NAIXUS (https://ircai.org/global-network-of-ai-excellence-centers/) global network of AI excellence centers is a promising example of a multi-institutional, international effort to bring Sustainable Development to the AI research agenda. Support local and regional centres of excellence that leverage data and AI for the SDGs in key cities in Europe. Identify and share best practices.

Incentives and regulation. Implement incentives, remove regulatory barriers and define enabling regulations with the aim of accelerating the use of AI for sustainable development following ethical principles that are complied with and accounted for. Invest in the necessary infrastructure and capacities to audit the compliance of AI systems with such ethical principles.

Sustainable AI. Invest in and incentivise sustainable AI systems. Develop regulations that require the systematic measurement and publishing of their carbon footprint.

Funding. Provide necessary funding to enable a financial model for AI for sustainable development projects. Foster public-private long-term collaborations to accelerate the achievement of the SDGs by leveraging AI methods.

5. Conclusion

We live in a time of prosperity, but we also face tremendous global challenges that threaten our own existence as a species – from poverty and hunger to climate change and the destruction of entire ecosystems. Effectively tackling these challenges requires an ambitious and coordinated commitment from most nations in the world, as reflected by the 17 SDGs. AI – and specifically data-driven AI methods – has the potential to significantly accelerate the achievement of the SDGs. However, to realise such a potential, we need to address five types of barriers related to the use of AI in this context: institutional, technical, ethical, financial and environmental. It is therefore of paramount importance to invest ambitiously in tackling such barriers so we can effectively leverage the power of AI to help us improve living conditions in our planet. An opportunity that we must not miss, as it might be our best (and last) chance to ensure not just the sustainability of our societies and our planet but our own survival. As Theodore Roosevelt said: 'A revolution is sometimes necessary'. As there is no planet B, I invite you to join the 'AI for sustainable development' revolution.

References

Abdul Hamid, S. S., Admodisastro, N., Manshor, N., Kamaruddin, A., Ghani, A. A. A. (2018), *Dyslexia adaptive learning model: student engagement prediction using machine learning approach*, In International Conference on Soft Computing and Data Mining, pp. 372-384. Springer, Cham.

Adam, I., and M. Fazekas (2018), *Are emerging technologies helping win the fight against corruption in developing countries?*, Pathways for Prosperity Commission, Background Paper Series, (21).

Adegbosin, A. E., Stantic, B., Sun, J. (2019), 'Predicting under-five mortality across 21 low and middle-income countries using deep learning methods', *medRxiv*, 19007583.

Allison, R. S., Johnston, J. M., Craig, G., Jennings, S. (2016), 'Airborne optical and thermal remote sensing for wildfire detection and monitoring', *Sensors*, 16(8), 1310.

Almalaq, A., Edwards, G. (2017), *A review of deep learning methods applied on load forecasting*, In 2017 16th IEEE international conference on machine learning and applications (ICMLA), pp. 511-516. IEEE.

Angwin, J., Larson, J., Mattu, S., Kirchner, L. (2016), *Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks*, ProPublica.

Arntz, M., T. Gregory, U. Zierahn, (2016), *The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis*, OECD Social, Employment and Migration Working Papers, No. 189, OECD Publishing, Paris, https://doi.org/10.1787/5jlz9h56dvq7-en.

Autor, D. (2010), *The polarization of job opportunities in the US labor market: Implications for employment and earnings*, Center for American Progress and The Hamilton Project(6), pp. 11-19.

Badr, G., Klein, L. J., Freitag, M., Albrecht, C. M., Marianno, F. J., Lu, S., … Hamann, H. F. (2016), 'Toward large-scale crop production forecasts for global food security', *IBM Journal of Research and Development*, 60(5/6), pp. 5-1.

Bao, Y., Tang, Z., Li, H., Zhang, Y. (2019), 'Computer vision and deep learning–based data anomaly detection method for structural health monitoring', *Structural Health Monitoring*, 18(2), pp. 401-421.

Barlow, M. (2016), *AI and medicine: Data-driven strategies for improving healthcare and saving lives*, O'Reilly Media.

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(22), pp. 13389-13398.

Baveja, A., Caulkins, J. P., Liu, W., Batta, R., Karwan, M. H. (1997), 'When haste makes sense: Cracking down on street markets for illicit drugs', *Socio-Economic Planning Sciences*, 31(4), pp. 293-306.

Blattner, L., Nelson, S. (2021), *How costly is noise? data and disparities in consumer credit*, arXiv preprint arXiv:2105.07554.

Blondel, Vicent D., De Montjoye Y-A, Hidalgo, C. A., V. Michel, (2013), *Unique in the Crowd: The privacy bounds of human mobility*, Belgium: "Communauté française de Belgique" on Information Retrieval in Time Evolving Networks

Blumenstock, J., Cadamuro, G., On, R. (2015), 'Predicting poverty and wealth from *mobile phone metadata', Science, 350(6264),* pp. 1073-1076.

Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., Pentland, A. (2014), *Once upon a crime: towards crime prediction from demographics and mobile data*, In Proceedings of the 16th international conference on multimodal interaction (pp. 427-434)

Bowles, J. (2014), *The computerisation of European jobs*, Bruegel, https://www.bruegel.org/2014/07/the-computerisation-of-european-jobs/

Brentan, B. M., Luvizotto Jr., E., Herrera, M., Izquierdo, J., Pérez-García, R. (2017), 'Hybrid regression model for near real-time urban water demand forecasting', *Journal of Computational and Applied Mathematics*, (309), pp. 532-541, https://doi.org/10.1016/j.cam.2016.02.009

Burrell, J. (2016), 'How the machine 'thinks': Understanding opacity in machine learning algorithms', *Big Data & Society*, 3(1), 2053951715622512.

Centellegher, S., Miritello, G., Villatoro, D., Parameshwar, D., Lepri, B., Oliver, N. (2018), 'Mobile money: Understanding and predicting its adoption and use in a developing economy', *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4), pp. 1-18.

Chandy, S. E., Rasekh, A., Barker, Z. A., Campbell, B., Shafiee, M. E. (2017), *Detection of cyber-attacks to water systems through machine-learning-based anomaly detection in SCADA data*, In World Environmental and Water Resources Congress 2017, pp. 611-616.

Chang-Tien, K. Y. L., Sirirat, S. (2004), *Survey of fraud detection techniques in networking, sensing and control*, In IEEE International Conference(2), pp. 749-754).

Chattopadhyay, A., Hassanzadeh, P., Pasha, S. (2020), 'Predicting clustered weather patterns: A test case for applications of convolutional neural networks to spatio-temporal climate data', *Scientific reports*, 10(1), pp. 1-13.

Chen, S., Hu, C., Barnes, B.B, Wanninkhof, R., Cai, W.J, Barbero, L., Pierrot, D., (2019), 'A machine learning approach to estimate surface ocean pCO_2 from satellite measurements', *Remote Sensing of Environment*, (228), pp 203-226, https://doi.org/10.1016/j.rse.2019.04.019

Clifton, L., Clifton, D. A., Pimentel, M. A., Watkinson, P. J., Tarassenko, L. (2013), 'Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors', *IEEE journal of biomedical and health informatics*, 18(3), pp. 722-730.

Collins, F. S., Varmus, H. (2015), 'A new initiative on precision medicine', *New England journal of medicine*, 372(9), pp. 793-795.

Cowls, J., Tsamados, A., Taddeo, M., Floridi, L. (2021), 'The AI gambit: leveraging artificial intelligence to combat climate change – opportunities, challenges, and recommendations', *Ai & Society*, pp. 1-25.

Cruz, E., Vaca, C., Avendaño, A. (2019), *Mining top-up transactions and online classified ads to predict urban neighborhoods socioeconomic* *status*, In 2019 IEEE International Conference on Big Data (Big Data), pp. 4055-4062. IEEE.

Dargam, F., Perz, E., Bergmann, S., Rodionova, E., Sousa, P., Souza, F. A. A., ... Bonachela, P. Z. (2020), *Supporting operational decisions on desalination plants from process modelling and simulation to monitoring and automated control with machine learning*, In International Conference on Decision Support System Technology, pp. 150-164.

David, J.M., and K. Balakrishnan (2010), 'Machine Learning Approach for Prediction of Learning Disabilities in School-aged Children', *Int. Journal of Computer Applications*, 9(11), pp. 975-8887

De Bem, P. P., de Carvalho Junior, O. A., Fontes Guimarães, R., Trancoso Gomes, R. A. (2020), 'Change detection of deforestation in the Brazilian Amazon using landsat data and convolutional neural networks', *Remote Sensing*, 12(6), pp. 901.

De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., ... Ronneberger, O. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*, 24(9), pp. 1342-1350.

Di Minin, E., Fink, C., Hiippala, T., Tenkanen, H. (2019), 'A framework for investigating illegal wildlife trade on social media with machine learning', *Conservation Biology*, 33(1), pp. 210.

Dia, H. (2001), 'An object-oriented neural network approach to short-term traffic forecasting', *European Journal of Operational Research*, 131(2), pp. 253–261

Dia, H., Rose, G. (1997), 'Development and evaluation of neural network freeway incident detection models using field data', *Transportation Research Part C: Emerging Technologies*, 5(5), pp. 313-331.

Dogo, E. M., Nwulu, N. I., Twala, B., Aigbavboa, C. (2019), 'A survey of machine learning methods applied to anomaly detection on drinking-water quality data', *Urban Water Journal*, 16(3), pp. 235-248.

Doshi, R., Falzon, D., Thomas, B. V., Temesgen, Z., Sadasivan, L., Migliori, G. B., Raviglione, M. (2017), 'Tuberculosis control, and the where and why of artificial intelligence', *ERJ Open Research*, 3(2).

Dubey, S., Singh, P., Yadav, P., Singh, K. K. (2020), 'Household waste management system using IoT and machine learning', *Procedia Computer Science*, 167, pp. 1950-1959.

Dunis, C., Middleton, P., Karathanasopolous, A., Theofilatos, K. (2016), *Artificial intelligence in financial markets*, London: Palgrave Macmillan.

Dwork, C., Roth, A. (2014), 'The algorithmic foundations of differential privacy', *Found. Trends Theor. Comput. Sci.*, 9(3-4), pp. 211-407.

Enzweiler, M., Gavrila, D. M. (2008), 'Monocular pedestrian detection: Survey and experiments', *IEEE transactions on pattern analysis and machine intelligence*, 31(12), pp. 2179-2195.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S. (2017), 'Dermatologist-level classification of skin cancer with deep neural networks', *Nature*, 542(7639), pp. 115-118.

Etscheid, J. (2019),'Artificial Intelligence in Public Administration. Int. Conf. on Electronic Government (EGOV)', *Lecture Notes in Comptuer Science*, 11685, pp 248-261.

European Commission, Directorate-General for Communications Networks, Content and Technology, (2019). *Ethics guidelines for trustworthy AI*, Publications Office. https://data.europa.eu/doi/10.2759/177365

FAO (2017), *Smart farming is key for the future of agriculture*, https://www.fao.org/family-farming/detail/es/c/897026

Faulkner, A., Cebul, K., McHenry, G. (2014), *Agriculture gets smart: the rise of data and robotics*, Cleantech Group: San Francisco, CA, USA.

Fedor, P., Vaňhara, J., Havel, J., Malenovský, I., Spellerberg, I. (2009), 'Artificial intelligence in pest insect monitoring', *Systematic Entomology*, 34(2), pp. 398-400.

Feldman, T., Peake, A. (2021), *End-To-End Bias Mitigation: Removing Gender Bias in Deep Learning*, arXiv preprint arXiv:2104.02532.

Fernandez, M. G., Tokuhiro, A., Welter, K., Wu, Q. (2017), 'Nuclear energy system's behavior and decision making using machine learning', *Nuclear Engineering and Design*, 324, pp. 27-34.

Frank, M. R., Sun, L., Cebrian, M., Youn, H., Rahwan, I. (2018), 'Small cities face greater impact from automation', *Journal of the Royal Society Interface*, 15(139), 20170946.

Frey, C. B., Osborne, M. A. (2017), 'The future of employment: How susceptible are jobs to computerisation?', *Technological forecasting and social change*, 114, pp. 254-280.

Froehlich, J., Neumann, J., and N. Oliver (2009), *Sensing and predicting the pulse of the city through shared bicycling*, IJCAI'09: Proceedings of the 21st international joint conference on Artificial Intelligence, pp. 1420–1426

Gagne II, D. J., Haupt, S. E., Nychka, D. W., Thompson, G. (2019), 'Interpretable deep learning for spatial analysis of severe hailstorms', *Monthly Weather Review*, 147(8), pp. 2827-2845.

Gandhi, N., Armstrong, L. (2016), *Applying data mining techniques to predict yield of rice in humid subtropical climatic zone of India*, In 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1901-1906. IEEE.

García-Martín, E., Rodrigues, C. F., Riley, G., Grahn, H. (2019), 'Estimation of energy consumption in machine learning', *Journal of Parallel and Distributed Computing*, 134, pp. 75-88.

Geschke, D., Lorenz, J., Holtz, P. (2019), 'The triple-filter bubble: Using agent based modelling to test a meta theoretical framework for the emergence of filter bubbles and echo chambers', *British Journal of Social Psychology*, 58(1), pp. 129-149.

Gillon, R. (1994), 'Medical ethics: four principles plus attention to scope', *Bmj*, 309(6948), pp. 184.

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., Agrawal, A. (2017), 'Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection', *Construction and building materials*, 157, pp. 322-330.

Granata, F., Papirio, S., Esposito, G., Gargano, R., De Marinis, G. (2017), 'Machine learning algorithms for the forecasting of wastewater quality indicators', *Water*, 9(2), pp. 105.

Grobart, S. (2012), *Big dairy enters the era of big data*, https://www.bloomberg.com/news/articles/2012-10-18/big-dairy-enters-the-era-of-big-data

Guo, J., Li, B. (2018), 'The application of medical artificial intelligence technology in rural areas of developing countries', *Health equity*, 2(1), pp. 174-181.

Haidar, A., Verma, B. (2018), 'Monthly rainfall forecasting using one-dimensional deep convolutional neural network', *IEEE Access*, 6, pp. 69053-69063.

Hajek, P., Henriques, R. (2017), 'Mining corporate annual reports for intelligent detection of financial statement fraud–A comparative study of machine learning methods', *Knowledge-Based Systems*, 128, pp. 139-152

Ham, Y. G., Kim, J. H., Luo, J. J. (2019), 'Deep learning for multi-year ENSO forecasts', *Nature*, 573(7775), pp. 568-572.

Han, Y., Wu, H., Jia, M., Geng, Z., Zhong, Y. (2019), 'Production capacity analysis and energy optimization of complex petrochemical industries using novel extreme learning machine integrating affinity propagation', *Energy Conversion and Management*, (180), pp. 240-249.

Held, M., Küng, L., Çabukoglu, E., Pareschi, G., Georges, G., Boulouchos, K. (2018), *Future mobility demand estimation based on sociodemographic information: A data-driven approach using machine learning algorithms*, In 18th Swiss Transport Research Conference (STRC 2018). STRC.

Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., Pineau, J. (2020), 'Towards the systematic reporting of the energy and carbon footprints of machine learning', *Journal of Machine Learning Research,* 21(248), pp. 1-43.

Hethcoat, M.G., Edwards, D.P., Carreiras, J.M.B, Bryant, R.G., França, F.M and S. Quegan (2019), 'A machine learning approach to map tropical selective logging', *Remote Sensing of Environment*, (221), pp 569-582, https://doi.org/10.1016/j.rse.2018.11.044

Holley, P. (2018), *The World Bank's latest tool for fighting famine: Artificial intelligence*, The Washington Post, 23.

Hosein, S., Hosein, P. (2017), *Load forecasting using deep neural networks*, In 2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1-5. IEEE. https://irp-cdn.multiscreensite.com/be6d1d56/files/uploaded/SDSN%20Health%20Solutions.pdf 2019

Huang, W., Song, G., Hong, H., Xie, K. (2014), 'Deep architecture for traffic flow prediction: deep belief networks with multitask learning', IEEE Transactions on *Intelligent Transportation Systems*, 15(5), pp. 2191-2201.

Huang, Y., Shen, L., Liu, H. (2019), 'Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China', *Journal of Cleaner Production*, 209, pp. 415-423.

Iqbal, J., Al-Zahrani, A., Alharbi, S. A., Hashmi, A. (2019), 'Robotics inspired renewable energy developments: prospective opportunities and challenges', *IEEE Access*, 7, pp. 174898-174923.

Ise, T., Oba, Y. (2019), 'Forecasting climatic trends using neural networks: an experimental study using global historical data', *Frontiers in Robotics and AI*, 32.

Jain, A., Singh, A., Koppula, H. S., Soh, S., Saxena, A. (2016), *Recurrent neural networks for driver activity anticipation via sensory-fusion architecture*, In 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 3118-3125.

Jang, J. Y., Lee, S., Lee, B. (2019), 'Quantification of gender representation bias in commercial films based on image analysis', *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), pp. 1-29.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., Ermon, S. (2016), 'Combining satellite imagery and machine learning to predict poverty', *Science*, 353(6301), pp. 790-794.

Jha, S. K., Bilalovic, J., Jha, A., Patel, N., Zhang, H. (2017), 'Renewable energy: Present research and future scope of Artificial Intelligence', *Renewable and Sustainable Energy Reviews*, 77, pp. 297-317.

Jiao, Z., Jia, G., Cai, Y. (2019), 'A new approach to oil spill detection that combines deep learning with unmanned aerial vehicles', *Computers & Industrial Engineering*, 135, pp. 1300-1311.

Johansson, F., Kaati, L., Sahlgren, M. (2017), *Detecting linguistic markers of violent extremism in online environments*, In Artificial intelligence: Concepts, methodologies, tools, and applications, pp. 2847-2863. IGI Global.

Kagan, D., Chesney, T., Fire, M. (2020), 'Using data science to understand the film industry's gender gap', *Palgrave Communications*, 6(1), pp. 1-16.

Kagashe, I., Yan, Z., Suheryani, I. (2017), 'Enhancing seasonal influenza surveillance: topic analysis of widely used medicinal drugs using Twitter data', *Journal of medical Internet research*, 19(9), pp. 7393.

Kaji, D. A., Zech, J. R., Kim, J. S., Cho, S. K., Dangayach, N. S., Costa, A. B., Oermann, E. K. (2019), 'An attention based deep learning model of clinical events in the intensive care unit', *PloS one*, 14(2), e0211057.

Kang, J., Park, Y. J., Lee, J., Wang, S. H., Eom, D. S. (2017), 'Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems', *IEEE Transactions on Industrial Electronics*, 65(5), pp. 4279-4289.

Kannangara, M., Dua, R., Ahmadi, L., Bensebaa, F. (2018), 'Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches', *Waste Management*, Volume 74, pp 3-15, https://www.sciencedirect.com/science/article/abs/pii/S0956053X17309406?via%3Dihub

Karimipour, H., Dehghantanha, A., Parizi, R. M., Choo, K. K. R., Leung, H. (2019), 'A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids', *IEEE Access*, 7, pp. 80778-80788.

Kenneally, E., Dittrich, D. (2012), *The Menlo Report: Ethical principles guiding information and communication technology research*, Available at SSRN 2445102.

Kim, H., Lee, Y., Woo, T., Kim, H. (2016), *Integration of vehicle and lane detection for forward collision warning system*, In 2016 IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin), pp. 5-8.

Król, A. (2016), 'The Application of the Artificial Intelligence Methods for Planning of the Development of the Transportation Network', *Transportation Research Procedia*, 14, pp. 4532–4541.

Kulkarni, N. H., Srinivasan, G. N., Sagar, B. M., Cauvery, N. K. (2018), *Improving crop productivity through a crop recommendation system using ensembling technique*, In 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), pp. 114-119.

Lacoste, A., Luccioni, A., Schmidt, V., Dandres, T. (2019), *Quantifying the carbon emissions of machine learning*, arXiv preprint arXiv:1910.09700.

LeCun, Y., Bengio, Y., Hinton, G. (2015), 'Deep learning', *Nature* 521, pp. 436–444. https://doi.org/10.1038/nature14539

Ledford, H. (2019), 'Millions of black people affected by racial bias in health-care algorithms', *Nature*, 574(7780), pp. 608-610.

Lee, K., Choi, C., Shin, D. H., Kim, H. S. (2020), 'Prediction of heavy rain damage using deep learning', *Water*, 12(7), pp. 1942.

Lei, N., Masanet, E. R. (2021), 'Global Data Center Energy Demand and Strategies to Conserve Energy, *Data Center Handbook: Plan, Design, Build, and Operations of a Smart Data Center*, pp. 15-26.

Letouzé, E. and A. Pentland (2018), 'Towards a Human Artificial Intelligence for Human Development', *ITU Journal: ICT Discoveries, Special Issue*, 2.

Letouzé, E., Oliver, N. (2019), *Sharing is Caring Four Key Requirements for Sustainable Private Data Sharing and Use for Public Good*, Data-pop alliance and Vodafone Institute for Society and Communications, London.

Letouzé, E., Vinck, P., Kammourieh, L. (2015), *The law, politics and ethics of cell phone data analytics*, Data-Pop Alliance.

Li, J., Xu, Q., Shah, N., Mackey, T. K. (2019), 'A machine learning approach for the detection and characterization of illicit drug dealers on instagram: model evaluation study', *Journal of medical Internet research*, 21(6), e13803.

Liao, S., Zhou, L., Di, X., Yuan, B., Xiong, J. (2018), *Large-scale short-term urban taxi demand forecasting using deep learning*, In 2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC, pp. 428-433.

Lismont, J., Cardinaels, E., Bruynseels, L., De Groote, S., Baesens, B., Lemahieu, W., Vanthienen, J. (2018), 'Predicting tax avoidance by means of social network analytics', *Decision Support Systems*, 108, pp. 13-24.

Lu, Z., Im, J., Quackenbush, L., Halligan, K. (2010), 'Population estimation based on multi-sensor data fusion', *International Journal of Remote Sensing*, 31(21), pp. 5587-5604.

Madhab Ghosh, S., and M.D. Behera (2018), 'Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest', *Applied Geography*, 96, pp 29-40, https://doi.org/10.1016/j.apgeog.2018.05.011

Malak, J. S., Zeraati, H., Nayeri, F. S., Safdari, R., Shahraki, A. D. (2019), 'Neonatal intensive care decision support systems using artificial intelligence techniques: a systematic review', *Artificial Intelligence Review,* 52(4), pp. 2685-2704.

Mandal, B., Li, L., Wang, G. S., Lin, J. (2016), 'Towards detection of bus driver fatigue based on robust visual analysis of eye state', *IEEE Transactions on Intelligent Transportation System*s, 18(3), pp. 545-557.

Marculescu, D., Stamoulis, D., Cai, E. (2018), *Hardware-aware machine learning: modeling and optimization*, In Proceedings of the International Conference on Computer-Aided Design, pp. 1-8).

Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M. F., Duarte, C. M. (2018), 'Use of unmanned aerial vehicles for efficient beach litter monitoring', *Marine pollution bulletin*, 131, pp. 662-673.

Min, Q., Lu, Y., Liu, Z., Su, C., Wang, B. (2019), 'Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry', *International Journal of Information Management*, 49, pp 502-519, https://www.sciencedirect.com/science/article/abs/pii/S0268401218311484?via%3Dihub

Mishra, M., Rout, P. K. (2018), 'Detection and classification of micro-grid faults based on HHT and machine learning techniques' *IET Generation, Transmission & Distribution*, 12(2), pp. 388-397.

Mitchell, T., Brynjolfsson, E. (2017), 'Track how technology is transforming work', *Nature*, 544(7650), pp. 290-292.

Mohanty, S. P., Hughes, D. P., Salathé, M. (2016), 'Using deep learning for image-based plant disease detection', *Frontiers in plant science*, 7, pp. 1419.

R Montes, F.J. Melero, I Palomares, S Alonso et al, Artificial Intelligence and Digital Technologies for the Sustainable Development Goals, *Spanish Royal Academy of Engineering*.

More, R., Mugal, A., Rajgure, S., Adhao, R. B., Pachghare, V. K. (2016), *Road traffic prediction and congestion control using Artificial Neural Networks*, In 2016 International Conference on Computing, Analytics and Security Trends (CAST), pp. 52-57. IEEE.

Mukai, N., Watanabe, T., Feng, J. (2008), *Route optimization using Q-learning for on-demand bus systems*, In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp. 567-574. Springer.

Niestadt, M., Debyser, A., Scordamaglia, D., Pape, M. (2019), *Artificial intelligence in transport: Current and future developments, opportunities and challenges*, European Parliamentary Research Service.

Nieto, Y., García-Díaz, V., Montenegro, C., Crespo, R. G. (2019), 'Supporting academic decision making at higher educational institutions using machine learning-based algorithms', *Soft Computing*, 23(12), pp. 4145-4153.

Oh, Y., Park, S., Ye, J. C. (2020), 'Deep learning COVID-19 features on CXR using limited training data sets', IEEE transactions on medical imaging, 39(8), pp. 2688-2700.

Okori, W., Obua, J. (2011), *Machine learning classification technique for famine prediction, In Proceedings of the world congress on engineering*, 2(1).

Oliver, N. (2019), *Governance in the Era of Data-driven Decision-making Algorithms*, in Women shaping Global Economic Governance, Center for Economic Policy Research Press (CEPR)

Oliver, N. and A.P. Pentland (2000), 'Graphical models for driver behavior recognition in a smart car', *Proc. of the IEEE Intelligent Vehicles Symposium*, pp 7-12.

Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M. others. (2020), 'Mobile Phone Data for Informing Public Health Actions across the COVID-19 Pandemic Life Cycle', *Science Advances*, 6.

Ong, E., Wong, M. U., Huffman, A., He, Y. (2020), 'COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning', *Frontiers in immunology*, pp. 1581.

Pan, Z., Liang, Y., Wang, W., Yu, Y., Zheng, Y., Zhang, J. (2019), *Urban traffic prediction from spatio-temporal data using deep meta learning*, In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1720-1730.

Pastor-Escuredo, D., Morales-Guzmán, A., Torres-Fernández, Y., Bauer, J. M., Wadhwa, A., Castro-Correa, C., … Luengo-Oroz, M. (2014), *Flooding through the lens of mobile phone activity*, In IEEE Global Humanitarian Technology Conference (GHTC 2014), pp. 279-286.

Petrillo, A., Travaglioni, M., De Felice, F., Cioffi, R., Piscitelli, G. (2020), 'Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions' *Sustainability*, 12(2), pp. 492. https://doi.org/10.3390/su12020492

Pham, T., Tran, T., Phung, D., Venkatesh, S. (2017), 'Predicting healthcare trajectories from medical records: A deep learning approach', *Journal of biomedical informatics*, 69, pp. 218-229.

Poon, L. C., Kametas, N. A., Maiz, N., Akolekar, R., Nicolaides, K. H. (2009), 'First-trimester prediction of hypertensive disorders in pregnancy', *Hypertension*, 53(5), pp. 812-818.

Quinn, J. A., Okori, W., Gidudu, A. (2010), *Increased-specificity famine prediction using satellite observation data*, In Proceedings of the First ACM Symposium on Computing for Development, pp. 1-6.

Radke, D., Hessler, A., Ellsworth, D. (2019), *FireCast: Leveraging Deep Learning to Predict Wildfire Spread*, In IJCAI, pp. 4575-4581.

Rafiei, M. H., Adeli, H. (2017), 'A novel machine learning-based algorithm to detect damage in high-rise building structures', *The Structural Design of Tall and Special Buildings*, 26(18), e1400.

Rasp, S., Pritchard, M. S., Gentine, P. (2018), 'Deep learning to represent subgrid processes in climate models', *Proceedings of the National Academy of Sciences*, 115(39), pp. 9684-9689.

Raza, M. Q., Khosravi, A. (2015), 'A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings', *Renewable and Sustainable Energy Reviews*, 50, pp. 1352-1372.

Ren, Y., Huang, J., Hong, Z., Lu, W., Yin, J., Zou, L., Shen, X. (2020), 'Image-based concrete crack detection in tunnels using deep fully convolutional networks', *Construction and Building Materials*, 234, 117367.

Rocher, L., Hendrickx, J. M., De Montjoye, Y. A. (2019), 'Estimating the success of re-identifications in incomplete datasets using generative models', *Nature communications*, 10(1), pp. 1-9.

Rodríguez, E. A., Estrada, F. E., Torres, W. C., Santos, J. C. M. (2016), *Early prediction of severe maternal morbidity using machine learning techniques*, In Ibero-American Conference on Artificial Intelligence, pp. 259-270.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... Bengio, Y. (2022), 'Tackling climate change with machine learning', *ACM Computing Surveys (CSUR)*, 55(2), pp. 1-96.

Sahoo, S., Russo, T. A., Elliott, J., Foster, I. (2017), *Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US Water,* Resour Res 53, pp. 3878–3895.

Sahour, H., Gholami, V., Vazifedan, M. (2020), 'A comparative analysis of statistical and machine learning techniques for mapping the spatial distribution of groundwater salinity in a coastal aquifer', *Journal of Hydrology*, 591, 125321.

San Pedro, J., Proserpio, D., Oliver, N. (2015), *MobiScore: Towards Universal Credit Scoring from Mobile Phone Data*, In: Ricci F., Bontcheva K., Conlan O., Lawless S. (eds) User Modeling, Adaptation and Personalization. UMAP. Lecture Notes in Computer Science, 9146. Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-20267-9_16

Saracco, R. (2018). *Digital twins take the train.*

Sayad, Y. O., Mousannif, H., Al Moatassime, H. (2019), 'Predictive modeling of wildfires: A new dataset and machine learning approach', *Fire safety journal*, 104, pp. 130-146.

Schneider, G. (2018), 'Automating drug discovery', *Nature reviews drug discovery*, 17(2), pp. 97-113.

Schut, D., Wisniewski, J. (2015), A global vision of railway development, Paris: International Union of Railways (UIC)

Schwartz, I. M., York, P., Nowakowski-Sims, E., Ramos-Hernandez, A. (2017), 'Predictive and prescriptive analytics, machine learning and child welfare risk assessment: The Broward County experience', *Children and Youth Services Review,* 81, pp. 309-320.

Schwartz, R., Dodge, J., Smith, N. A., Etzioni, O (2020), *Green AI*, Communications of the ACM, 63 (12), pp 54–63, https://doi.org/10.1145/3381831

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., ... Hassabis, D. (2020), 'Improved protein structure prediction using potentials from deep learning', *Nature,* 577(7792), pp. 706-710.

Setti, M., Falasconi, L., Segrè, A., Cusano, I., & Vittuari, M. (2016), 'Italian consumers' income and food waste behavior', *British Food Journal.*

Shorrocks, A., and N. Hechler-Fayd'herbe (2019), Global Wealth Report 2019: Global wealth rises by 2.6% driven by US China, despite trade tensions, Zurich, Switzerland: Credit Suisse

Siddiqui, M. Z., Yadav, S., Husain, M. S. (2018), 'Application of artificial intelligence in fighting against cyber crimes: a review', Int. J. Adv. Res. *Comput. Sci*, 9(2), pp. 118-122.

Singh, J. A. (2019), 'Artificial intelligence and global health: opportunities and challenges', *Emerging Topics in Life Sciences*, 3(6), pp. 741-746.

Sønderby, C. K., Espeholt, L., Heek, J., Dehghani, M., Oliver, A., Salimans, T., … Kalchbrenner, N. (2020), *Metnet: A neural weather model for precipitation forecasting*, arXiv preprint arXiv:2003.12140.

Song, S. Y., Kim, Y. K. (2018), 'Theory of virtue ethics: do consumers' good traits predict their socially responsible consumption?', *Journal of Business Ethics*, 152(4), 1159-1175.

Soto, V., Frias-Martinez, V., Virseda, J., Frias-Martinez, E. (2011), 'Prediction of socioeconomic levels using cell phone records', In *International conference on user modeling, adaptation, and personalization*, pp. 377-388.

Spiekermann, S., Acquisti, A., Böhme, R., Hui, K. L. (2015), 'The challenges of personal data markets and privacy', *Electronic markets*, 25(2), pp. 161-167.

Staiano, J., Oliver, N., Lepri, B., de Oliveira, R., Caraviello, M., Sebe, N. (2014), *Money walks: a human-centric study on the economics of personal mobile data*, In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 583-594.

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

Sundmaeker, H., Verdouw, C. N., Wolfert, J., Freire, L. P. (2016), *Internet of food and farm 2020*, In Digitising the industry, 49, pp. 129-150. River Publishers, en https://www.iof2020.eu/

Sundsøy, P., Bjelland, J., Reme, B. A., Iqbal, A. M., Jahani, E. (2016), D*eep learning applied to mobile phone data for individual income classification*, In Proceedings of the 2016 International Conference on Artificial Intelligence: Technologies and Applications, Bangkok, Thailand, pp. 24-25.

Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Sweeney, S. J., Vercauteren, K. C., Snow, N. P., Miller, R. S. (2019), 'Machine learning to classify animal species in camera trap images: applications in ecology', *Methods Ecol. Evol*, 10, pp. 585-590.

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y. A., Rahman, A. (2020), 'Land-use land-cover classification by machine learning classifiers for satellite observations – A review', *Remote Sensing*, 12(7), pp. 1135.

Tumas, P., Nowosielski, A., Serackis, A. (2020), 'Pedestrian detection in severe weather conditions', *IEEE Access*, 8, pp. 62775-62784.

USAID (2010), *Uganda food security outlook*. Famine Early Warning Systems Network. Technical Report

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., … Fuso Nerini, F. (2020), 'The role of artificial intelligence in achieving the Sustainable Development Goals', *Nature communications*, 11(1), pp. 1-10.

Walsh, C. G., Ribeiro, J. D., Franklin, J. C. (2017), 'Predicting risk of suicide attempts over time through machine learning', *Clinical Psychological Science*, 5(3), pp. 457-469.

Wang, Q., Guo, Y., Yu, L., Li, P. (2017), 'Earthquake prediction based on spatio-temporal data mining: an LSTM network approach', *IEEE Transactions on Emerging Topics in Computing*, 8(1), pp. 148-158.

Wesolowski, A., Buckee, C. O., Bengtsson, L., Wetter, E., Lu, X., Tatem, A. J. (2014), 'Commentary: containing the Ebola outbreak-the potential and challenge of mobile network data', *PLoS currents*, 6.

Wesolowski, A., Eagle, N., Tatem, A., Smith, DL., Noor, AM. et al. (2012), 'Quantifying the Impact of Human Mobility on Malaria', Science, 338(6104), pp. 267-272, DOI: 10.1126/science.1223467

West, J., Bhattacharya, M. (2016), 'Intelligent financial fraud detection: a comprehensive review', Computers & security, 57, pp. 47-66

Weyn, J. A., Durran, D. R., Caruana, R. (2020), 'Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere', *Journal of Advances in Modeling Earth Systems*, 12(9).

Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M. J. (2017), 'Big data in smart farming–a review', *Agricultural systems*, 153, pp. 69-80.

Woolf, B. P., Lane, H. C., Chaudhri, V. K., Kolodner, J. L. (2013), 'AI grand challenges for education', *AI magazine*, 34(4), pp. 66-84.

Wu, J., Ruenz, J., Althoff, M. (2018), *Probabilistic map-based pedestrian motion prediction taking traffic participants into consideration*, In 2018 IEEE Intelligent Vehicles Symposium (IV), pp. 1285-1292.

Wu, J., Song, C., Dubinsky, E. A., Stewart, J. R. (2021), 'Tracking Major Sources of Water Contamination Using Machine Learning', *Frontiers in microbiology*, 3623.

Xu, M., Cai, H., Liang, S. (2015), 'Big data and industrial ecology', *Journal of Industrial Ecology*, 19(2), pp. 205-210.

Xu, X., Qi, Y., Hua, Z. (2010), 'Forecasting demand of commodities after natural disasters', *Expert systems with applications*, 37(6), pp. 4313-4317.

Xu, Y., Bao, Y., Chen, J., Zuo, W., Li, H. (2019), 'Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images', *Structural Health Monitoring*, 18(3), pp. 653-674.

Yao, H., Tang, X., Wei, H., Zheng, G., Li, Z. (2019), *Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction*, In Proceedings of the AAAI conference on artificial intelligence, 33(01), pp. 5668-5675).

Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., ... Li, Z. (2018), *Deep multi-view spatial-temporal network for taxi demand prediction*, In Proceedings of the AAAI Conference on Artificial Intelligence, 32(1).

Ye, C., Butler, L., Calka, B., Iangurazov, M., Lu, Q., Gregory, A., … Middleton, C. (2019), *A digital twin of bridges for structural health monitoring*, In 12th International Workshop on Structural Health Monitoring 2019. Stanford University.

Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O., Lee, K. K. (2011), 'A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer', *Journal of hydrology*, 396(1-2), pp. 128-138.

Zanetti, M., Iseppi, G., Cassese, F. P. (2019), 'A "psychopathic" artificial intelligence: The possible risks of a deviating AI in education', *Research on Education and Media*, 11(1), pp. 93-99.

Zemel, R., Wu, Y., Swersky, K., Pitassi, T., Dwork, C. (2013), *Learning fair representations*, In International conference on machine learning, pp. 325-333.

Zhang, N., Wang, M., Wang, N. (2002), 'Precision agriculture – a worldwide overview', *Computers and electronics in agriculture*, 36(2-3), pp. 113-132.

Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., … Li, H. (2019), 'T-gcn: A temporal graph convolutional network for traffic prediction', *IEEE Transactions on Intelligent Transportation Systems*, 21(9), pp. 3848-3858.

Zhao, Q., Yu, S., Zhao, F., Tian, L., Zhao, Z. (2019), 'Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments', *Forest Ecology and Management*, 434, pp 224-234, https://doi.org/10.1016/j.foreco.2018.12.019

Zheng, X., and M.K.O. Lee (2016), 'Excessive use of mobile social networking sites: Negative consequences on individuals', *Computers in Human Behavior*, Volume 65, pp 65-76, https://doi.org/10.1016/j.chb.2016.08.011

Zhu, N., Liu, X., Liu, Z., Hu, K., Wang, Y., Tan, J., … Guo, Y. (2018), 'Deep learning for smart agriculture: Concepts, tools, applications, and opportunities' International *Journal of Agricultural and Biological Engineering*, 11(4), pp. 32-44.

Zhu, S., Hrnjica, B., Ptak, M., Choiński, A., Sivakumar, B. (2020), 'Forecasting of water level in multiple temperate lakes using machine learning models', *Journal of Hydrology*, 585, 124819