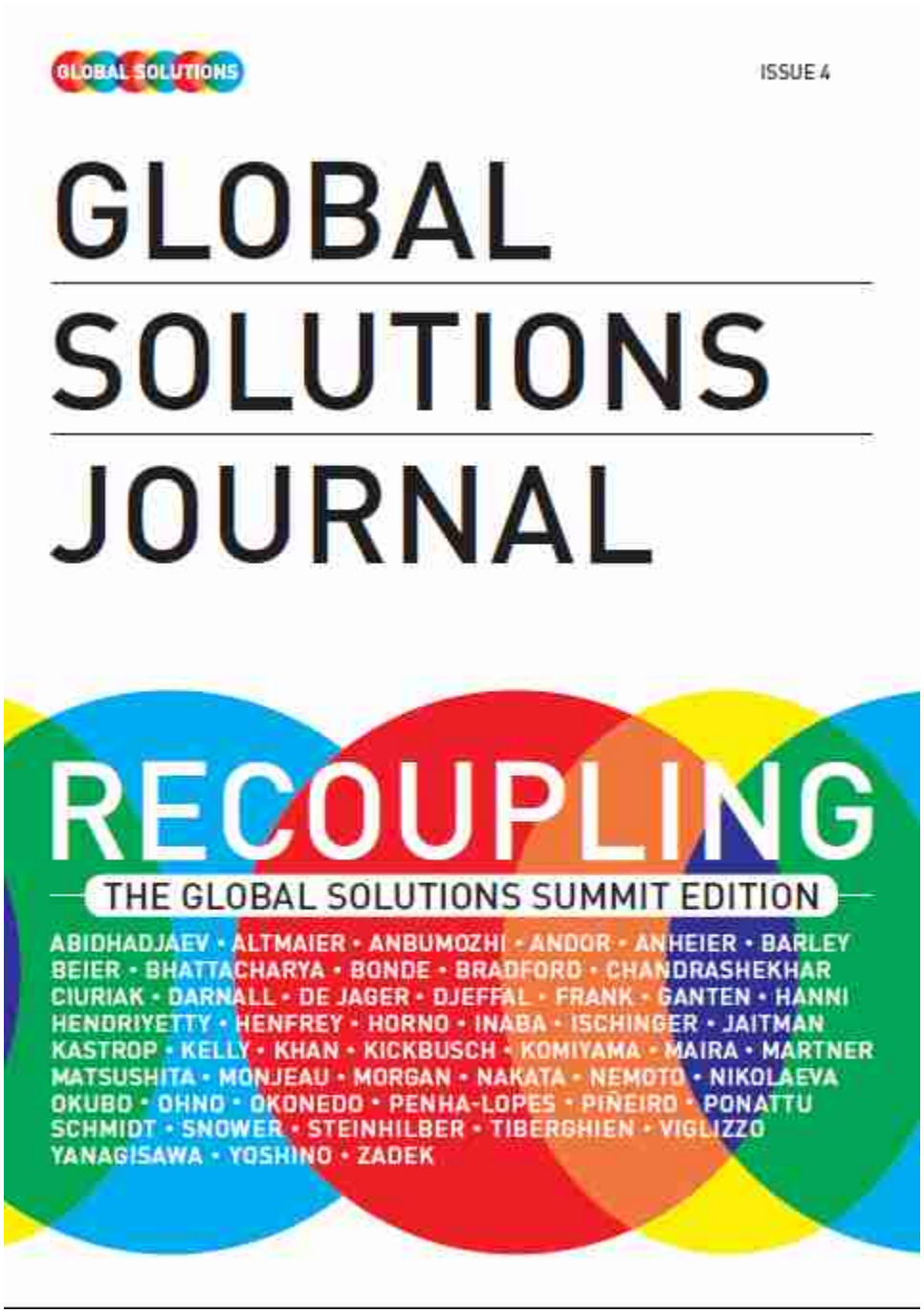

Djeffal, C. (2019). Sustainable development of artificial intelligence (SAID). *Global Solutions Journal*, 186–192.

Sustainable development of artificial intelligence (SAID)

The author:

Christian Djeffal
Associated Researcher, Alexander-von-Humboldt-Institute for Internet and Society (HIIG)

The institution:

The Alexander von Humboldt Institute for Internet and Society researches the development of the internet from a societal perspective. The aim is to better understand the digitalization of all spheres of life. As the first institute in Germany with a focus on internet and society, HIIG has established an understanding that emphasises the embeddedness of digital innovations in societal processes.

Artificial intelligence (AI) is changing our lives. However, it seems to be a solution and a problem at the same time. In order to understand opportunities and risks, to evaluate promises and warnings, we should ask the right questions and choose the right frame of analysis. It is time to think and talk about AI in terms of sustainable development.

I. WE NEED TO TALK ABOUT ARTIFICIAL INTELLIGENCE

AI applications play an increasing role in all parts of society. In hospitals, image recognition systems aid doctors in diagnosing diseases like skin cancer. They are in many cases more accurate than the doctors. In our private life, different AI applications allow us to "speak" with our mobile phones, to take better pictures and to sort them automatically. Fully automated public transportation is available in many countries and driving is increasingly automated. In factories, AI-empowered industry robots can handle an ever-increasing number of tasks. In offices, AI tools help to plough through data in order to make predictions or to identify certain important aspects of that data. Risk management tools are also used by tax administrations to identify tax

applications in need of further assistance. Face recognition and other controversial AI technologies are used to ensure public security and safety.

These examples highlight some of the important features of the concept of AI. According to Klaus Mainzer's definition, AI denotes systems that can deal with certain problems independently and efficiently. The term has famously been framed as a research question by four computer scientists in a grant proposal. Today, AI is considered a sub-discipline of computer science that is actively pursued by specialized academics and research institutions. AI is not a single technology but a set of technologies.[i] Those technologies are not limited to specific circumstances in their use. They can be used in many contexts and for various purposes. AI denotes a set of general-purpose technologies in the truest sense of the word. This might be one reason why AI has been compared to other general-purpose technologies like steam-machines, and the production of steel.

Being a general-purpose technology, one can create swords and ploughshares from AI.

The metaphors of swords and ploughshares certainly applies to AI, and this can even be seen in the practice of the United Nations. UN Global Pulse, an innovation initiative of the Secretary General, uses AI proactively for development and humanitarian action. The Pulse Lab Kampala developed a system that monitors local public radio stations all over the country. Small hardware is deployed all over the country, a text-to-speech application transforms the data and another AI application analyses and clusters the data. This is done in order to give the government a better understanding of the problems in the country, including floods, violence against women, malaria, and cholera. On the other side of the spectrum are so-called lethal autonomous weapons systems (LAWS), which are also fuelled by artificial intelligence. After several years of expert meetings of experts, states established a Group of Governmental Experts under the umbrella of the Convention on Certain Conventional Weapons. The controversial negotiations in this group show that new possibilities through AI can be hard to resolve.

»Being a general-purpose technology, one can create swords and ploughshares from AI.«

Governments, companies and research institutions are having heated debates about the future of artificial intelligence. They talk about ethics, about AI's impacts on human rights and about the ways in which AI could strengthen the economy. Depending on the respective focus, specific topics concerning AI are in the limelight. While stakeholders have identified many important issues, very little reflection has addressed the question of the right frame. In particular, policy makers

and other stakeholders have not explored the potential of framing AI in the terms of sustainable development.

II. WHY SUSTAINABLE AI DEVELOPMENT?

While there are sporadic attempts to highlight the potential of AI to contribute to the fulfilment of sustainable development goals, there is at the moment no comprehensive analysis of how different technologies are to be considered from the perspective of sustainable development. This is somewhat surprising, considering that the discourse on sustainable development is very far advanced but captures a basic conflict that represents current discussions concerning AI. That is the conflict between development and its potential costs. Sustainable development acknowledges the opportunities of progress, but also highlights its limitation due to impacts on societal concerns such as the environment. Reducing the main advantages of sustainable AI development to three words, it would be inclusiveness, governance and connectivity.

»Sustainable AI development is an inclusive framework.«

Inclusiveness

The discourse on sustainable development has attained a very high level of inclusiveness. It relates to all countries and regions of the world. It addresses issues of economic development, but attributes the same importance to social and environmental issues. It highlights human rights, but is also sensitive for concerns at the level of groups and countries. It addresses public and private actors and includes civil society. What is even more important that sustainable development works on the micro, the meso and the macro level. Sustainability can be an issue for specific products and applications. Yet, it can be also a standard measuring overall impacts on the environment or social justice. In contrast, the focus of the discourse on AI tends to be on the consequences of specific applications, particularly in their performance concerning fairness, accountability and transparency. Save for the future of work, the impacts of AI on society as a whole are only rarely discussed. Sustainable AI development as a frame could change that.

Governance

One of the remarkable features of sustainable development is its focus on governance and implementation. The discourse on sustainable development has continuously been improved. The 2030 Agenda for Sustainable Development has established a multi-stakeholder process in order to support achievement and development of specific goals and related indicators. A sustainable perspective on AI does not only highlight goals and limits, but necessarily implies the question how they are to be realized. From this perspective, it is not enough to talk about ethics and human rights but to take specific actions in order to achieve certain goals or prohibit certain impacts of technologies. While the current state of governance mechanisms is far from perfect, they are steps in the right direction. AI technologies might be further tools to strengthen the governance of sustainable development.

Fit

Generally, the framework of sustainable development has been used to deal with similar issues and, therefore, is a good fit for AI's current challenges. First, it sets limits and goals for technological development. Hence, it is neither exclusively directed to the potential nor to the risks of AI. Sustainable development highlights both sides of the coin. What's more, the concept of sustainable development is not shy to address conflicts. From the initial discussions, there was an inherent conflict between sustainability and development. While sustainable AI development does not necessarily deal with the need for progress and environmental impacts, there are often trade-offs that need to be addressed in a clear manner. These trade-offs can concern the functionality of the systems and the many things that can be achieved on the one side and notions of privacy, transparency and individual autonomy on the other.

III. LEARNINGS FROM SAID

Sustainable AI development is an inclusive framework. Yet, it is different than other frameworks and has the potential to highlight specific aspects of artificial intelligence and digitization. This change of perspectives offers specific learnings that can be implemented when discussing the future of artificial intelligence.

1. Understand change and share knowledge

It is important to understand change caused by AI in a comprehensive manner and to map its impacts. This applies to beneficial and harmful impacts. It is one of the major features of the discourse surrounding sustainable development that changes are tracked, goals are defined and monitored. Best practices are shared and communicated across communities. Such a model would have great benefits for the governance of artificial intelligence. This is particularly important in the field of AI for discussions are sometimes guided by specific very visible applications like automated cars or narratives from Hollywood movies. In order to assess the change, it would be necessary to have a complete picture of where AI is already in use and what the repercussions of that use are. Instead of focussing on very specific issues like the trolley dilemma,[1] it would be interesting to see how AI becomes a part of the everyday infrastructure. Such a systematic approach might uncover that there are many more applications that are not automated decision-making systems but nevertheless play an ever-increasing role.

2. Develop a strategy of strategies

There is currently a proliferation of national AI strategies. Canada, China, Finland, France, Germany, and Japan are among the countries that have published such an AI-strategy. While it is certainly a good thing that more and more countries position themselves regarding such a set of transformational technologies, the question remains how countries coordinate. While every country wants to become a leader in AI, this is not a realistic option,

save for a few countries. And even the countries that can be leaders in AI technology contemplate whether they are strong enough to outweigh the rest of the world. Despite certain quotes from different politicians in 2016, it becomes more and more clear that AI research cannot be compared to the race for the nuclear weapon. It is about different technologies, different areas of application and different purposes. While machine learning based on neural networks works well today, it is far from clear what the next breakthrough will be based on. In such a situation, we are in need of a strategy of strategies allowing for competition and cooperation at the same time. Such a strategy of strategies should provide incentives for inventions but help to avoid that the same inventions and mistakes are made in different countries again and again. It should also coordinate efforts where countries actually have the same goals and can work together. The sustainable development agenda is such a strategy of strategies.

3. Justice

One of the most important features of sustainable development is the constant quest for justice equity and fairness. Sustainable development is concerned with current affairs but also takes into account future generations. As a framework, SAID makes such questions of justice visible on the micro and the macro level. There are currently issues that are not really resolved. One issue that is projected to be a great concern in the future is the distribution of wealth gained from trained neural networks. This is particularly problematic when the training data is based on the efforts of people that do not profit from the AI applications. This has been discussed in the context of music and health. It is even more significant when this problem is combined with the gig-economy. There are concerns that certain platforms might take advantage of missing labor regulation in developing countries and exploit the workforce there for simple labelling tasks. To train an algorithm in such a setting means effectively to make one's own ability redundant. However, gig-workers without legal protection have little leverage to influence such a setting. If these and comparable situations are not addressed from the perspective of sustainability and justice, we might be soon faced with another aspect of what has been coined "digital slavery".[ii]

4. Build an infrastructure and an ecosystem for sustainability

From the perspective of sustainable AI development, an infrastructure for sustainable AI development is key. This infrastructure has to transgress the goals of building fast and functional AI applications. It will also have to deal with questions of making applications sustainable by taking into account individual, group and societal needs. Take for example the need to assess and test AI applications. This is an interest shared by developers, companies, regulatory authorities and policy makers alike. Testing and evaluation are much easier when there is an infrastructure allowing to do this. A focus group of the ITU currently undertakes such an effort by building evaluation and test sets for specific AI driven eHealth applications.[iii] They are in the process of creating a test environment that would allow the evaluation and benchmarking of eHealth applications.

What has been stated for infrastructure is also true for the ecosystem. While it is certainly necessary to have hubs that spark innovation and help to make innovations applicable to real world problems, those ecosystems should also include reflections on the ethical, legal and social impacts of AI. Many AI strategies consider so-called "AI-observatories". In certain domains, there are also agencies that have to test and certify AI applications.

»If it is true that ›responsible design eats oversight for breakfast‹, it will be important to think about agents of good design. «

If it is true that "responsible design eats oversight for breakfast", it will be important to think about agents of good design.

Those agents have to be close to the processes of building AI and able to include ethical, legal and social aspects right from the beginning. In certain jurisdictions, data protection officers have become the agents of data protection in companies, organizations and public administration. In the case of using AI in public, the ecosystem should also include citizens. There are different participation tools that could be used to include citizens in the design process. These methods include sortition or random sample voting. When there are important decisions to make for instance in the smart city context, it would be possible to consult such a group that would be selected to be trained and work as a citizens' data ethics commission. The ecosystem for sustainable AI development should comprise various actors and organizations, from the UN High Level Panel on Digital Cooperation to local citizens' data ethics commissions.

IV. THE FUTURES OF AI

In conclusion, the preliminary learnings from sustainable AI development show that there is a lot of potential to think about AI in terms of sustainable development. As a set of general-purpose technologies, AI has not one single future. It has many futures, and these futures depend on more than just technological inventions. They depend on the decisions taken to fund research and innovation, they depend on decisions to implement technologies in society, and they depend on choices about how to design and how to implement technologies. AI applications can play an important role in society but it is us who decide which role they play. To employ sustainable AI development as a framework makes us aware of the choices.

[1] In the context of AI, this is often related to situations in which a car can do nothing than choose between harming and potentially killing different groups of people. The problem is that this choice can be programmed.

[i] U. Gasser and V. A.F. Almeida, 'A Layered Model for AI Governance', IEEE Internet Computing 21 (2017), 58–62.

[ii] D. Snower, 'Emancipation in the age of digital slavery', Global Solutions Journal 1 (2018), 18–20.

[iii] M. Salathé, T. Wiegand, M. Wenzel and R. Kishnamurthy, Focus Group on Artificial Intelligence for Health. https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents