WASP—HS

AI, Sustainability and Agenda 2030

Report 2023

Table of Contents

Wallenberg AI, Autonomous Systems and Software Program - Humanities and Society (WASP-HS) would like to thank all chairs and participants of the event AI, Education and Children for contributing to the fruitful discussions which this report is based on.

Introduction

In this WASP-HS community meeting, we navigate the intricate landscape where artificial intelligence (AI) meets sustainability and the overarching objectives of the UN Agenda 2030. The invited talk and the discussions in the three roundtables illuminated the challenges and opportunities that lie ahead. A diverse group of participants, including experts spanning computer science, pedagogy, political science, legal studies, and global health, engaged in a collaborative exploration of the multifaceted challenges and opportunities arising from the integration of AI into education, environmental governance, and healthcare.

Dr. Ricardo Vinuesa's keynote underscored the pivotal role of AI in advancing the Sustainable Development Goals (SDGs) while acknowledging its potential to hinder certain targets. A consensus-based expert elicitation process revealed that AI could facilitate the accomplishment of 134 targets spanning all Sustainable Development Goals but might impede 59 targets. However, a critical observation emerged: the current research focus tends to overlook vital aspects of AI's impact. Dr. Vinuesa emphasized the imperative for regulatory insight and oversight to guide the rapid development of AI-based technologies, ensuring transparency, safety, and adherence to ethical standards. The absence of such regulatory measures could result in detrimental gaps. Crucially, the keynote argued against a pause in AI development, asserting that the path forward involves intensified regulation, private sector engagement, and ongoing research. The development of interpretable AI models, incorporating inductive biases and symbolic regression, was identified as a key element in ensuring AI's alignment with sustainable development.

Shifting the spotlight to accountability beyond AI, the roundtable led by Ericka Johnson and Jonas Ivarsson unveiled complex challenges in assigning responsibility at various governance levels. The conversation navigated through the intricate terrain of regulating AI as it transitions from proprietary to open frameworks. The participants stressed the necessity of engaging multiple voices and perspectives in AI regulation discussions and acknowledged the environmental impact of AI, extending beyond energy consumption.

The round table on "To Monitor Is to Manage – Or Not? Which Data Do We Need to Reach the Environmental SDGs?", led by Sabine Höhler, Adam Wickberg and Erik Ljungberg, delved into the delicate balance of making environmental data politically actionable while being cognizant of its various uses, including commercial and for-profit applications. The dialogue advocated for an integrative approach to understanding how data affects environmental knowledge and governance. The importance of incorporating issues of equity and data justice into environmental governance discussions was emphasized, particularly in the context of the data-rich digital environment. The challenges associated with data quantity and quality, the dichotomy between quantitative and qualitative data, and the complexities of data-driven policymaking were dissected. The conversation cautioned against the assumption that more data, or larger models, equate to better solutions. Instead, it urged a nuanced evaluation of the impact of AI and data on the Sustainable Development Goals, considering historical and critical perspectives.

The round table on the "future of sustainable health in the context of AI, digitized bodies, and care" facilitated by Pedro Sanches, Teresa Almeida, and Eirini (Irene) Kaklopoulou, spotlighted the potential lack of competencies for ensuring that emerging AI technologies benefit historically disadvantaged groups within healthcare systems. The need for ongoing and democratic auditing and certification processes was identified as crucial, particularly in engaging marginalized actors and avoiding exclusions.

The dialogue recognized the risks of exacerbating historical inequalities in healthcare, especially for older adults and persons with disabilities. It underscored the importance of local-level implementation of technology, involving clinicians, and strengthening local capacity. The environmental sustainability of AI in health was also scrutinized, emphasizing the need for a nuanced cost-benefit analysis, and viewing AI not just as technology but as infrastructure.

In essence, the event provided a comprehensive exploration of the multifaceted intersections between AI, sustainability, and the Sustainable Development Goals. It advocated for thoughtful regulation, inclusivity, and ongoing evaluation to harness the potential benefits of AI while mitigating its potential risks in the pursuit of a sustainable and equitable future. The overarching message was that for AI to positively impact sustainable development, it must be implemented inclusively and

with careful consideration of its broader sociotechnical implications. Main conclusions can be summarized as follows:

- Potential and Pitfalls of AI for Sustainable Development: AI presents a promising tool for advancing Sustainable Development Goals (SDGs) but demands vigilant oversight to prevent hindrances to specific targets.

- Urgent Need for Regulation: Robust regulatory frameworks are crucial to steer AI development, ensuring transparency, ethical standards, and addressing governance and environmental impact challenges.

- Inclusive Implementation: Implementing AI inclusively, particularly in healthcare, is vital to prevent the marginalization of specific groups. Sociotechnical considerations must be integral, viewing AI as infrastructure rather than just a technological tool.

Virginia Dignum, CRM organiser

WASP-HS Community Reference Meetings (CRMs) are meeting places for Swedish private and public organizations and WASP-HS researchers. Each meeting has a specially selected theme with the aim of bringing business and research together to expand knowledge and strengthen collaboration.

This report is based on the discussions and conclusions from the CRM on the topic of AI, education and children. The event took place on 4 October, 2023.

WASP—HS

Highlights From Keynote Speech

Keynote Speaker

Ricardo Vinuesa, Associate Professor, Vice Director of KTH Digitalization Platform and Lead Faculty at KTH Climate Action Centre

The emergence of artificial intelligence (AI) and its progressively wider impact on many sectors requires an assessment of its effect on the achievement of the Sustainable Development Goals. Using a consensus-based expert elicitation process, we find that AI can enable the accomplishment of 134 targets across all the goals, but it may also inhibit 59 targets. However, current research foci overlook important aspects. The fast development of AI needs to be supported by the necessary regulatory insight and oversight for AI-based technologies to enable sustainable development. Failure to do so could result in gaps in transparency, safety, and ethical standards.

If AI is used to help achieve the Sustainable Development Goals, it needs to be through the use of interpretable models. There are methods to achieve such interpretability from deep-learning methods, for instance through inductive biases and symbolic regression. Furthermore, we argue that a pause in AI development would be counterproductive to achieve the Sustainable Development Goals, as what is actually needed is more regulation, involvement from the private sector and research development. Finally, we are working on novel methods to optimize policy leveraging natural-language processing.

Ricardo Vinuesa, Associate Professor

Accountability Beyond AI – Thinking of Tools in the Hands of Humans

Authors
Ericka Johnson, Professor of Gender Studies at Linköping University
Jonas Ivarsson, Professor in Informatics at University of Gothenburg

Main Challenges

- The levels of governance at which responsibility and accountability for AI should be assigned are contestable.
- Regulating AI as it (and its data) shifts from proprietary to open frameworks becomes complicated.
- Conversations about regulating AI must engage multiple voices and perspectives.
- AI's environmental impact stems from more than just its energy use.

Our panel brought together academics with backgrounds in computer science, pedagogy, political science, legal studies, and global health along with civil society actors interested in how AI can be used to address the needs of interest groups and how the Sustainable Development Goals can be used to hold governments responsible to the best interests of civil society.

The conversation displayed a shared interest in concepts of equity and sustainability as well as concerns about accountability and auditability. The discussion revolved around the potential implications of AI on the Sustainable Development Goals (SDGs) for 2030. Concerns were raised about AI solutions that may predominantly benefit affluent nations or require extensive resources, thereby making certain goals harder to achieve.

We began by discussing how accountability and responsibility can be assigned to actors at different levels: international bodies, nation states, civil society and commercial interests, to name a few. How these different levels can relate to each other and to AI is tricky, but we do not need to reinvent the wheel; there are various models we could learn from, for example, the way the World Bank works together with national banks; the way the United Nations Guiding Principles on Business and Human Rights (UNGP) are being used by national judicial systems to hold companies responsible for poor behavior or environmental pollution, etc. Our conversation would suggest that there should be regulatory structures and disciplinary bodies at the national level which align with international goals and agreements. However, we also recognized that the ability of states to do this varies widely around the world and is particularly difficult in areas where access to technological infrastructures is very limited.

There are significantly different regulatory challenges when AI (and particularly the development of it and the data it uses) is commercial, proprietary material or when it is the public domain and open source (and the data is open data), the latter being much more difficult to regulate or impose guardrails for. This is also related to the difficulties of regulating an international, unbounded technology as it changes regulatory environment – as it is developed in one country, using data there, and then moves across legal frameworks into others.

We felt strongly that conversations about regulating AI must engage multiple voices and perspectives There can be a tendency for conversations about AI to become siloed, or focus on one aspect of the technology, from a technological perspective. Cross-sectoral conversations are necessary to open space for a variety of concerns.

Another topic we broached was that AI technology has significant environmental impacts, particularly through energy use, in its development and implementation. But the uses of AI also impact existing non-AI aspects of society (like the tourism and housing markets which have been impacted by Airbnb). These second-level impacts are significant but hard to predict, which makes regulating for such impacts difficult.

Finally, we are concerned about the digital dehumanization and the appropriation of human images, words

and voices, which produces data points rather than citizens and subjects. And, as the term AI can address many different types of technology, we suggest it would be useful to be more specific about the various technologies currently referenced by the term.

Next Steps

1. We identified AI familiarity and education for the general public and for those tasked with producing governance – both because we want to be governed by informed policies and laws and because as individual citizens and consumers we need to be able to act upon and monitor AI. This becomes even more important as we see that some states have a tendency to create areas of exception – for example allowing technological development and use beyond the regulatory framework in some spheres like policing and security.

2. We would like to avoid regulations being a box to check. Rather, we would like to see that they have an impact on the core development of AI technologies. To that end, we feel that the development of core values, perhaps even core principles, professional ethics statements or codes of conduct, may be useful. Here we note that some of the older, more established professions like medicine already have such codes of conduct and ethical regulations. Technology should not be an exception.

3. We suggest considering an additional Sustainable Development Goal to specifically address AI; though there is also a benefit to mainstreaming AI given that it will have an impact on so many of the other development goals. However, we do not want to lose sight of the fact that by becoming a signatory to the Sustainable Development Goals, a state then can be held accountable to these by civil society actors – and is thereby also given a larger mandate to regulate private and commercial actors. Thus, perhaps there would be a significant benefit to having AI included explicitly in one or several Sustainable Development Goals.

ORIGINAL ABSTRACT Many of the Sustainable Development Goals for 2030 could be made more difficult to attain by the implementation of AI solutions that benefit rich nations, powerful actors and capital, or which demand extensive resources for technical development, manufacturing, transportation to market, use and data storage, for example:

Goal 8: “Promote sustained, sustainable and inclusive economic growth, full and productive employment, and decent work for all”.

Goal 10: “Reduce inequality in and among countries”.

Goal 13: “Take urgent action to combat climate change and its impacts.”

What responsibility do institutions and individuals in society have to prevent these negative consequences? And which institutions and individuals should we hold accountable for this?

To Monitor Is to Manage – Or Not? Which Data Do We Need to Reach the Environmental Sustainable Development Goals?

Authors

Sabine Höhler, Professor of Science and Technology Studies at KTH Royal Institute of Technology

Erik Ljungberg, PhD student at KTH Royal Institute of Technology

Adam Wickberg, Docent of History of Science, Technology and Environment at KTH Royal Institute of Technology

Main Challenges

- The desire to make environmental data politically actionable in immediate ways is understandable but requires carefully qualifying the aims. Data circulation accelerates various forms of data use, including commercial and for-profit use.
- Data accommodates different stakeholder views and opposing interests. The same data sets can support diverging understandings of "sustainability".

We need a broad and integrative view on how data affects our knowledge and management of the environment. Starting from the observation that the global environment and its local manifestations have become data-rich, digital concepts, issues of equity and data justice need to be integrated with environmental justice in new forms of governance. Sustainability studies, policy, and governance ground their understanding of governable environments increasingly and almost exclusively on data, which forges new power relations in response to the urgency of the green transition. Environmental goals such as the Sustainable Development Goals, indicator systems, and performance reviews build on the ready availability of environmental data, preferably quantified data that can be fed into information processing systems to model global environmental pasts and futures. From autonomous sensor networks to satellite surveillance systems, new AI-based technologies are employed to outsmart anthropogenic environmental changes. Often, the word "smart" features in these efforts (cf. Halpern & Mitchell 2023). Historical and critical perspectives can reveal a problematic paradox of AI's power in global environmental governance, namely that we propose technological solutions to problems that often follow from technological applications, not least digital technologies.

The title provocation, To Monitor Is to Manage, refers to Lord Kelvin, the 19th century English physicist who insisted that you cannot manage what you cannot measure. Today, this expectation seems even more pervasive and persuasive: better technology will lead to better data which will lead to a more sustainable use of natural environments. But does this equation hold? Will better data help us reach the sustainability goals? And what is better data? How can we evaluate the quest for ever larger models and AGI to reach the SDGs? Recent research and critical discussions on AI and datafication suggest that more data and bigger models might not be the solution (Bender et al. 2021; Bakker et al. 2018).

The theme and questions stem from the research project The Mediated Planet at KTH Royal Institute of Technology (Formals Research Program Realising the Global Sustainable Development Goals) which explores the global environment as emerging through environmental data. Forestry, climate science, oceanography, mining, and agriculture provide examples to study how data access and ownership shape environmental perception and politics in relation to Sustainable Development Goals 13 (Climate Action), 14 (Life Below Water), 15 (Life on Land), and 16 (Peace, Justice and Strong Institutions). The project addresses the epistemological dimensions of environmental data ubiquity in environmental science and policy and the political dimensions of how data-making relates to policy making. The following themes require further discussion and research:

• Data quantity and data quality: Vast quantities of data are produced daily, earth observation satellite data being just one example. The image of environmental data "mining" seems inviting but is inaccurate. Big data sets are not easily reusable, nor

retrievable. A "datum" is often stripped from qualifying information. Accessing and interpreting data relies on experts, which creates unequal situations of technical and social data friction.

- Quantitative versus qualitative data: In environmental and sustainability research, qualitative may give way to quantitative studies which lend themselves to automated processes that render fast and vast results, for instance in climate change modeling. Big data modeling using AI seems seductive but is selective, favoring input sources that are measurable or compatible with existing information processing tools and operations. It is a known problem that data gaps and bias increase existing inequalities. Beyond data justice, datafied information will regulate how environments are represented and what remains invisible. The relation of data to comprehensive knowledge remains fuzzy.

- Data-driven policymaking: Translating data into policy seems attractive but is deceptive. Data does not drive environmental policy in linear fashion. Data informs environmental governance more broadly. The global environmental change debates are data-heavy discourses, yet, even collectively held data models do not impact decision-making processes immediately. Unsustainable fisheries respond not with business changes but with shifting baselines of expected yield. The social, cultural, and economic mechanisms at play are as powerful as environmental evidence. Has data become a proxy for the environment and data collection a proxy for environmental action? At the same time, data can serve marginalized stakeholders – often indigenous peoples – to express and legitimate their positions in environmental politics.

To summarize, data-intensive science cannot substitute human knowledge. Data can support stimulating human curiosity and informing policy making. For data to become not only communicable but also actionable, data needs to translate into comprehensive knowledge. The conclusions are, first, that environmental data, even big data, do not effortlessly turn into environmental knowledge, and second, that governance structures can enable but also inhibit data travel between the sciences and other sectors – public and private. Data travel is affected by shifting public and private ownership of data and by shifting data access in open market economies – which in turn may accelerate the development of AI tools.
The challenges are, first, that data travel seems to accelerate not only scientific but also commercial data use. In corporate climate action initiatives, data becomes currency in the green transition, promising green innovation and sustainable growth. Recent policy work such as the EU Corporate Sustainability Reporting Directive (CSRD) and the EU Sustainable Finance Disclosure Regulation (SFDR) makes sustainability quantification and reporting based on environmental data an integral part of business models. Secondly, best practices and possible pathways to sustainability diverge, ranging from more data to more inclusive data to deliberate political positioning and outright activism (Asayama 2023). To narrow the data "usability gap" (Lemos et al. 2012), we need to understand data governance structures in the history of modern states and avoid simplified ideas of mobilizing data for the greater ecological good.

More data literacy education will be needed to manage the increasing ubiquity of data in our lives. The current 17 Sustainable Development Goalsmight be amended by an 18th goal: AI and Digital Literacy.

References

Asayama, Shinichiro et al. (2023): "Three institutional pathways to envision the future of the IPCC", Nature Climate Change, Volume 13: 877–880. https://doi.org/10.1038/s41558-023-01780-8

Bakker, Karen & Max Ritts (2018): "Smart Earth: A meta-review and implications for environmental governance", Global Environmental Change vol. 52: 201-211.

Bender, Emily, Timnit Gebru, Angelina McMillan Major & Shmargaret Shmitchell (2021): "On the Dangers of Stochastic Parrots: Can Large Language Models be Too Big?" FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 610–623. https://doi.org/10.1145/3442188.3445922

Halpern, Orit & Robert Mitchell (2023): The Smartness Mandate (MIT Press).

Lemos, Maria Carmen, Christine J. Kirchhoff, and Vijay Ramprasad (2012): "Narrowing the climate information usability gap", Nature Climate Change, Volume 2: 789–794.

Charting the Future of Sustainable Health: Navigating AI, Digitized Bodies, and Care

Authors

Pedro Sanches, Assistant Professor in Human-Centerered Artificial Intelligence at Umeå University

Teresa Almeida, Associate Professor in Human-Computer Interaction at Umeå University

Eirini Kaklopoulou, PhD student in Human-Computer Interaction at Umeå University

Main Challenges

- There may be a lack of competencies for ensuring that emerging AI technologies will benefit historically disadvantaged groups as they are implemented into healthcare systems. Therefore, there should be more focus on the implementation of technology at a local level.
- Since AI could represent a new paradigm shift for policy making, there is a need to research forms of auditing and certification that are ongoing and democratic. This could require establishing processes for open debate on AI that engage marginalized actors in a way that goes beyond current exclusions.

This roundtable brought together academics and practitioners with experience in public health and welfare administration as well as legal expertise in Sweden and Europe. Our discussions were centered on identifying urgent sociotechnical issues around AI innovation and implementation in the Swedish societal milieu around health and well-being, from the present until 2030.

Digitized Bodies: Who Is Left Behind?

Amid an ongoing digitalization of health services in the Nordic context, now intensified by the emergence of AI technologies, the equal distribution of benefits of technologies for the health and well-being of all individuals is threatened by historical inequalities. With their focus on innovation, researchers often neglect how emerging technologies are implemented in the healthcare system.

Regarding implementation, this roundtable identified two groups that risk being further disenfranchised by the introduction of AI-based technologies: older adults and persons with disabilities (Gulliksen et al., 2021). These are groups who disproportionally depend on health services. Simultaneously, there is a lack of resources at the municipality level specifically to ensure that emerging technologies are used widely by these groups. We extend the recommendations to work with clinicians in the design of AI-based systems, but we add the urgency of strengthening local capacity to work with marginalized groups. Additionally, we recommend evaluating the impact of AI on Sustainable Development Goals by looking not only at the forefront of AI-based technologies but also at how these get implemented and adopted in local contexts. This is especially important as this family of technologies matures. Additionally, evaluations of alignment between AI and Sustainable Development Goals could also be supplemented by related UN frameworks that focus specifically on human rights and inequalities, e.g. Convention on the Rights of Persons with Disabilities (United Nations, 2015).

Health: From Individuals to Communities to Planet

When it comes to considering the environmental sustainability of AI in health, the roundtable discussion understood that a risk/benefit assessment needs to consider how the environmental impact of data centers and the cost of training AI models might affect other groups than those benefiting from AI-based technologies. A nuanced cost-benefit analysis is therefore critical.

AI-based technologies can be beneficial for decision-makers in healthcare and public health. Issues of implementing AI for decision-making are often framed in terms of trust. Participants of the roundtable also understood that there is a pressing need to

understand the impact private AI developers could have on public healthcare. Issues of trust in AI need therefore to be considered at a social and political level. Ultimately, to ensure the long-term sustainability of public health and healthcare services, AI should not only be treated as technology but also as infrastructure, and as such broader sociotechnical issues of healthcare system design and financing need to be considered when evaluating its positive and negative impacts.

AI in Practice: From Care to Certification

AI-based technologies can play an important role in the everyday work practices of clinical personnel. To do so, it is necessary to understand the variability and multitude of care practices by clinical professionals. Locality and transparency of the development of technology can play a role in how trustworthy it is perceived to be, but ultimately the successful implementation of AI-based technologies also depends on reliable standards, certification, and audit processes. It is recommended that the capacity for auditing be developed either at the municipal or national level. The rapid pace of technological change, with frequent updates in data and models also causes challenges in ensuring that these technologies are properly audited. Health professionals require that these processes be done transparently and openly to ensure that AI-based technologies are trusted by all.

The positive impact of AI-based technologies in addressing the Sustainable Development Goals depends on addressing current structural issues with healthcare in Sweden, as technology can only be relevant to social issues when it is imbued with sociotechnical considerations from design to implementation and adoption.

References

Alan Turing Institute. (2022). AI Standards Hub. Retrieved October 20, 2023, from https://aistandardshub.org/ai-standards-search/

Gulliksen, J., Larsdotter, M., & Johansson, S. (2021). Ny teknik och digitala lösningar för ökad inkludering i arbetslivet: En kunskapssammanställning. Nordens välfärdscenter/Nordic Welfare Centre.

Nations, U. (2015). Committee on the rights of persons with disabilities.

Strange, M., & Tucker, J. (2023). Global governance and the normalization of artificial intelligence as 'good' for human health. AI & SOCIETY. https://doi.org/10.1007/s00146-023-01774-2

This Report

This report is made possible by the Wallenberg AI, Autonomous Systems and Software Program - Humanities and Society (WASP-HS), a national research program in Sweden. The vision behind WASP-HS is to promote new interdisciplinary knowledge in the humanities and social sciences on the subject of artificial intelligence and autonomous systems and their impact on human and social development. The research program enables cutting-edge research, expertise and competence building in the humanities and social sciences. In total, the Wallenberg Foundations are investing up to SEK 660 million in the WASP-HS research program.

For more information please visit www.wasp-hs.org.

For questions or inquiries please contact us at contact@wasp-hs.org.

Authors

Adam Wickberg, Docent of History of Science, Technology and Environment at KTH Royal Institute of Technology

Eirini Kaklopoulou, PhD student in Human-Computer Interaction at Umeå University

Erik Ljungberg, PhD student at KTH Royal Institute of Technology

Ericka Johnson, Professor of Gender Studies at Linköping University

Jonas Ivarsson, Professor in Informatics at University of Gothenburg

Pedro Sanches, Assistant Professor in Human-Centerered Artificial Intelligence at Umeå University

Ricardo Vinuesa, Associate Professor, Vice Director of KTH Digitalization Platform, Lead Faculty at KTH Climate Action Centre

Sabine Höhler, Professor of Science and Technology Studies at KTH Royal Institute of Technology

Teresa Almeida, Associate Professor in Human-Computer Interaction at Umeå University

Virginia Dignum, Professor at the Department of Computing Science at Umeå University

How to cite this report

WASP-HS. Community Reference Meeting: AI, Sustainability and Agenda 2030. Report. December 2023.

WASP—HS

Wallenberg AI, Autonomous Systems and Software Program
- Humanities and Society