A MANIFESTO CONCERNING ARTIFICIAL INTELLIGENCE FOR MONITORING SUSTAINABLE DEVELOPMENT: THE MISSING LINK BETWEEN SDGS, INVESTMENT AND TRUST

JOHN SHAWE-TAYLOR

Professor of Computational Statistics and Machine Learning and UNESCO Chair of Artificial Intelligence at University College London and Director of the International Research Centre on Artificial Intelligence (IRCAI) under the auspices of UNESCO at the Jožef Stefan Institute in Slovenia.

DANIEL MIODOVNIK

Director at Social Finance and co-founder of their Digital Labs, Advisor to the International Research Centre on Artificial Intelligence (IRCAI) under the auspices of UNESCO.

DAVOR ORLIC

Honorary Research Assistant at UCL Centre for Artificial Intelligence and Chief Operations Officer at the International Research Centre on Artificial Intelligence (IRCAI) under the auspices of UNESCO at the Jožef Stefan Institute in Slovenia.

SDG6 - Clean Water and Sanitation
SDG7 - Affordable and Clean Energy
SDG9 - Industry, Innovation and Infrastructure
SDG10 - Reduced Inequalities
SDG11 - Sustainable Cities and Communities
SDG12 - Responsible Consumption and Production
SDG13 - Climate Action
SDG15 - Life on Land
SDG16 - Peace, Justice and Strong Institutions
SDG17 - Partnerships for the Goals

A MANIFESTO CONCERNING ARTIFICIAL INTELLIGENCE FOR MONITORING SUSTAINABLE DEVELOPMENT: THE MISSING LINK BETWEEN SDGS, INVESTMENT AND TRUST

ABSTRACT

We have seen an immense surge of interest in tackling the United Nations Sustainable Development Goals (SDGs). Countries, companies and investors around the world are committed to addressing the global economic, social and environmental crises. Investors have already committed US$89 trillion in assets to investments targeting SDG outcomes as part of the Principles for Responsible Investment (PRI) program.

However, there is the danger that without objective and reliable ways of assessing progress the momentum will be lost. We've seen an erosion of trust between citizens and governments, tech companies and industry alike. The lack of a consistent framework and the current subjectivity of data and ratings are holding us back.

We believe that artificial intelligence (AI) and data are fundamental to building the trustworthiness and evidence of measurable progress against the SDGs. We are already seeing examples of how clearly defined and measurable outcomes can unlock investment to solve the SDGs. For example, clear outcome metrics and data collection underpinned a $10 million outcomes contract to address rural sanitation in Cambodia (SDG Goal 6: Ensure availability and sustainable management of water and sanitation for all).

Therefore, we have designed a manifesto that calls on NGOs, the UN, companies, investors and countries to collaboratively build a robust, accessible and transparent system for measuring and certifying attainment of the SDGs. Together, we can build the AI and data ecosystem to create trust and enable investors, companies and governments to demonstrate progress, secure investment, and ultimately, change the world.

INTRODUCTION

At the moment, there is a lack of objective understanding of where exactly we stand in terms of progress towards reaching the SDGs. We believe that a universally acknowledged innovative technical mechanism, as well as a mechanism for finance and investment, would create trust between all stakeholders. The missing link is the convergence of both the usage of AI for measurement of the progress of the SDGs and social impact bonds that together can be used by governments to finance such technical endeavors. In this chapter, we propose a narrative that could potentially solve this missing link.

SOCIAL IMPACT BONDS

Poor sanitation, especially in places where open defecation routinely occurs, is linked to poor health outcomes, from spreading diseases to contaminating drinking water. To help the Royal Government of Cambodia bring safe sanitation to some of the poorest and most vulnerable households in Cambodia, Social Finance partnered with Stone Family Foundation, International Development Enterprises (iDE) and the United States Agency for International Development (USAID) to design the world's first impact bond for sanitation.

The goal of the impact bond was to reach 85% rural sanitation coverage in target areas by 2023, with 1,600 villages achieving open-defecation-free (ODF) status. Reaching this milestone would accelerate Cambodia's efforts to reach universal sanitation ahead of the 2030 SDG targets (Social Finance, 2021).

After an impressive decade of growth in sanitation coverage in rural Cambodia, remaining households tended to be in the poorest and hardest-to-reach areas. To help the Cambodian government realize its ambitious target by 2023, iDE, a leading rural sanitation provider, needed to access funding to innovate. The impact bond provided funding to innovate. Stone Family Foundation contributed the upfront funding to iDE, which gave iDE the resources to develop and deliver a rural sanitation program to reach the poorest and most vulnerable households. USAID agreed to deliver up to £10 million in outcome funding to Stone Family Foundation if iDE's program enabled these villages to achieve ODF status.

The impact bond was launched in November 2019. USAID last reported that 500 villages had achieved ODF status, with 88,738 households now having confirmed access to sanitation, in line with the Cambodian government's ODF guidelines. USAID has paid $3,125,000 in outcomes to date (USAID, 2019). This is an example of how data can enable innovative financing to drive progress towards the SDGs.

Challenges and importance of verification

The previous example confirms our belief that there is a significant appetite among investors to commit their money to companies that are able to contribute to sustainable development. In other words, such investors are willing to potentially settle for a lower or longer financial return on their investment if they can be reassured that the money will be used to further specific or general SDGs. This should be no surprise, given the interest in ethical investment that has over many years seen investors remove their support for companies whose actions are seen as unethical, such as promoting smoking, using cheap labor in sweatshops, and so on. The key difference between the constraint to investing in unethical businesses and investing in sustainable development is that the former is derived from evidence that the company has performed unethically in a specific way that is relatively easy to verify. On the other hand, proving a company is consistently contributing to sustainable development requires a very different level of evidence.

An example of this difficulty is illustrated by a recent article analyzing green bonds in the Brazilian forestry sector (Ferrando et al., 2021). We quote from the abstract:

> Through the study of recent green bond issuances realized by private companies active in the forestry sector in Brazil, we discuss how green bonds as a "new" form of "green" debt put nature at work and transform the territories and natural elements in the Global South into "temporal and spatial fixes" for the needs of global financial capital.

This is just one example of how demonstrating one's green credentials is difficult to underpin with objective and verifiable data. A recent scientific study has investigated the extent to which carbon offsets are generating the promised effects and found evidence of overestimation: "Results suggest that the accepted methodologies for quantifying carbon credits overstate impacts on avoided deforestation and climate-change mitigation" (West et al., 2020).

At the core of the difficulties with verification is the question of whom we can trust to provide objective and accurate information. Indeed, the whole ESG Initiative (Environment, Social and Governance) has been called into question by Tariq Fancy, BlackRock's first global chief investment officer for sustainable investing: "But there are other issues with ESG investing, including its subjectivity and the unreliability of data and ratings" (Amaro, 2021).

The key issue here is that the people generating the ratings and data are those that will potentially profit from a positive assessment, creating a conflict of interest and consequent erosion of trust at the heart of the initiative.

There are certainly very encouraging reports, such as the work in Costa Rica that received Prince William's environmental Earthshot Prize or a recent submission to the IRCAI Global Top 100 Outstanding Project list[46] based on using computer vision to detect carbon emission in Zambia's forests:

> Our project is based on detecting and reducing carbon emissions in forests using computer vision. We intended to collect data using satellite and also data science, machine learning and artificial intelligence. After collecting the data, we are going to pre-process it, and it will be ready for training and metrics and performance evaluation using Keras software for analysis. The impact of this project is on about 300 people within and near the national parks near the forest that will benefit from this project (Zamculture, 2019).

The surge of support and interest in tackling the UN SDGs is currently at record high levels. While this is an extremely positive development, there is a real danger of disillusionment setting in if companies and countries are found to short-change on the truth, and as we have seen, there is already evidence that this is happening. There is also the danger that without objective and reliable ways of assessing progress, for example, social media could be used to stain a company's image by spreading unfounded rumors that the credentials they claim are not true. Such developments could significantly undermine the interest in and support for SDG investment.

An example of the scale of support for SDG investment is given by the Principles for Responsible Investments (PRI) program of investing with SDG outcomes (UNPRI, 2022) to which investors have committed a combined US$89 trillion in assets under management. Their framework is summarized in their diagram (UNPRI, 2020), which includes the following steps:

46. IRCAI is a center under the auspices of UNESCO. Website at ircai.org

1. Identify outcomes
2. Set policies and targets
3. Investors shape outcomes
4. Financial system shapes collective outcomes
5. Global stakeholders collaborate to achieve outcomes in line with the SDGs

The framework is well-constructed and identifies the aims of the program to direct investment to address the UN SDGs. At the heart of this approach is the need for "investors [to] individually seek to increase positive outcomes, decrease negative outcomes and measure progress toward established targets."

While the question of measurement is highlighted, the broader question of trust is also important to capture. Again, in the words of this report:

> With more objective assessment of SDG Key Performance Indicators (KPI) there is greater opportunity for stakeholders to support initiatives that are making verifiable impact: these could be individual investors, governments, other companies making informed choices about collaboration, etc.

However, we are living in a time of widespread mistrust of institutions and leaders, with most people believing government and business leaders are seeking to mislead them (UNESCO, 2020). Set against this backdrop of the erosion of trust, we believe that this missing piece of the jigsaw is crucial for the role of AI in sustainable development. Therefore, we propose the following manifesto:

There is an urgent need to create a robust system for measuring and certifying the attainment of SDG KPIs, where possible giving evidence for the interventions that were responsible for any changes (positive or negative). The system and its operation need to earn the trust of all stakeholders: citizens, governments, tech companies and industry.

REALIZING THE MANIFESTO

We now turn our attention to the question of how this manifesto can be brought to life. Here we will argue that trust can be created if the conclusions are based on collected and verifiable data and that there is an even-handed presentation of the strengths and weaknesses of the inferences that are drawn from the data.

The role of data

All types of datasets can form the basis for assessing several aspects of the realization of different KPIs of the SDGs. Data has the potential:

- to measure whether an outcome has occurred;
- to record that outcome in a manner that is trusted by all;
- to ensure verifiability and attributability of the outcome to that service or product;
- to use that data to make a payment and to analyze how to improve services, as we shouldn't be satisfied until the SDGs have been fully delivered.

Data are being collected at an unprecedented rate using local and remote sensors. There is also a well-established movement that is arguing for such data collections and science more generally to be made open. For example, UNESCO has established a Recommendation on Open Science:

> The idea behind Open Science is to allow scientific information, data and outputs to be more widely accessible (open access) and more reliably harnessed (open data) with the active engagement of all the stakeholders (open to society) (Masakhane, 2022).

Open Science captures perfectly the potential role and approach that can engender trust in data, but also encourage broader participation in scientific exploration. This is an important part of building trust, namely that all groups should feel that they can participate, in terms of collecting data but also in verifying and contributing to its analysis. By groups here we could be referring to different regions of the world, different sections of society, different scientific disciplines, different governments, NGOs or corporations. The model of developed nations bringing ready-made solutions to bear on remote problems can very easily result in solving the wrong problem or overlooking critical local conditions, resulting in a poor solution or, even worse, no solution at all, with the consequent erosion of trust in both the collaboration and science in general.

An important part of open science and open data is a recognition that local challenges need local participation, in defining the challenge, collecting the data, and collaborating in developing solutions. The Masakhane initiative is an excellent example of an organization trying to do this for African languages with considerable success:

> Masakhane is a grassroots organization whose mission is to strengthen and spur NLP (natural language processing) research in African languages, for Africans, by Africans. Despite the fact that 2,000 of the world's languages are African, African languages are barely represented in technology. The tragic past of colonialism has been devastating for African languages in terms of their support, preservation and integration. This has resulted in technological space that does not understand our names, our cultures, our places, our history (Fairtrade Foundation, n.a.).

The technologies required to certify validity of data are well studied and are being increasingly deployed. In some cases, this can be relatively straightforward, for example for data collected remotely by satellite. The Fairtrade brand has a more challenging problem of tracking its products and producers to ensure that their standards are maintained, but this is an example of a trusted brand that has succeeded in managing this complex task:

> FLOCERT, an independent organization, checks that the Fairtrade standards have been met by the farmers, workers and companies that are part of the product supply chains. In order to reassure consumers that this has happened, we license the use of the FAIRTRADE Mark on products and packaging to signal the standards have been met (VideoLectures.NET, 2020).

Hence, while we do not want to underestimate the challenge, we believe that there is reason for optimism that the Open Science initiative can provide a framework within which the task of collecting and certifying relevant data can be developed and realized. However, collecting and certifying data in itself is not sufficient to attest to the achievement of the KPIs, let alone attribute responsibility. For this, we need to extract insights and knowledge from the data, and it is here that AI can play a vital role.

The role of AI

AI and machine learning are technologies that can be used to extract useful information from data in a verifiable and transparent way: hence they have an increasingly key role to play. As an example, Aidan O'Sullivan has used AI to analyze multispectral satellite imagery to assess water quality in lakes anywhere in the world (Schölkopf, 2019). While this might at first sight only appear to require access

to satellite imagery, there is a vital role of some "ground-truth" data concerning the quality of the water taken from different lakes in order to provide the training data that enables the AI to correctly identify the quality from the multispectral measurements and generalize from a small number of ground-truth measurements. This is an example of the need for local data collection requiring appropriate validation and certification, while there may also be a need for further refinement of the AI methods in order to quantify the accuracy of the predictions in specific cases.

This example again illustrates the variety of contributions that are needed and how a collaboration of the willing can potentially create an ecosystem that will inspire trust through transparency, openness and connectivity. We return to this theme below, but first we should discuss a critical technological component that is required, but which has yet to reach the necessary level of maturity: AI digital twins and mathematical modeling that allow for complex models to track KPIs and provide causal evidence between actions and their outcomes.

The challenge is the need to assign credit or responsibility for changes in the KPIs to the various actors involved. This could be evidence of continued exploitation of a resource such as in deforestation or evidence of interventions that address the issues causing the negative trend, such as for example interventions to improve water quality. The analysis of causality in machine learning is well-established (Schölkopf, 2019) but needs to be scaled to what is often now referred to as digital twins. These are computer models of a particular phenomenon or ecosystem that can be used to test how various interventions have influenced, or could influence, the different KPIs. Hence, through building a complex model of a particular environment we are able to answer "what if" questions and apportion responsibility for the observed and documented changes. As indicated above, a complex model will require advances in AI and mathematical modeling, in particular building on recent advances such as the data-centric engineering program at the Alan Turing Institute (ATI, 2021).

WHAT ARE THE BARRIERS TO REALIZING THE MANIFESTO?

There are a number of issues that may hinder implementation of the manifesto and it is sensible to assess the risks they might pose to its realization. Here we list them briefly.

The first is a lack of common definitions of outcomes and ways to measure them that are trusted by the public, companies and NGOs. The KPIs of the SDGs developed by the United Nations provide a starting point, but this issue will require careful attention, coupled with technical and public engagement, in order to build the necessary level of agreement and trust.

This naturally leads to the second concern that there is a collective action problem around who is, and should be, responsible for developing the definitions of outcomes and the technology solutions that capture and record them. This topic of building solutions that measure and verify outcomes does not represent an obviously attractive focus for funding, because given its nature we are not sure what would be the ideal funding body or the timeline for the return on investment for such a type of initiative. Our manifesto is designed to make the case for this funding by arguing that it makes sense to invest in such an initiative, but leaves open the question of the potential sources of that funding.

A third area of concern is the issue of data ethics and privacy and what is appropriate and ethical to collect and store. This concern needs to be addressed in collaboration with the people and communities affected in order to build trust in how data is being used, following the guidelines of the UNESCO Recommendation on the Ethics of Artificial Intelligence.

Governance is the final issue that we want to highlight: the question of who is responsible for "approving" an outcome definition or the AI for measuring SDGs. It is, of course, critical that the governance be made accountable and transparent in order to engender the necessary trust. This last component builds on the previous ones and is essentially the linchpin for making the manifesto credible and effective.

We need to overcome all these barriers if we are to unlock the potential of data and AI to measure progress against the SDGs, create accountability, and enable investment in companies focused on them.

HOW MIGHT WE OVERCOME THE BARRIERS?

Perhaps the one most important guiding principle for addressing all of the barriers and risk factors is to work collectively: the public, investors, companies, governments, international organizations and NGOs need to come together to define standards around outcome definition, collection, verification, attributability, etc. It is only by ensuring a consensus that the agreed methodology will not become discredited by the criticisms of one or more stakeholders.

The second guiding principle is to start small to build trustworthiness: building trust does not happen overnight. Instead of trying to tackle all 17 SDGs from the beginning, we should rather start with a small number of SDGs in order to demonstrate the potential for data and AI in overcoming trust barriers and building credibility in the approach. This will help test key assumptions around building trust, perception of risks, and whether it results in the unlocking of more investment into tackling the SDGs.

The third guiding principle is to leverage existing tools and applications that scale. There are so many emerging AI solutions that could support this ambition. We should understand what exists and what can be used and scaled without reinventing the wheel.

We have already stressed the need to create transparent governance. We believe this can be achieved by establishing agreed methodologies for determining what outcome definitions, approaches to recording data, and other mechanisms are acceptable to all stakeholders, and how this approval process happens. We need an accounting framework for SDG outcomes that enables organizations to audit what has been achieved. The accounting framework needs to be developed in partnership with the public, investors, companies, governments and NGOs in order to build trust and utility. Furthermore, engagement and partnership efforts need to include the affected people and communities so that the efforts reflect their experiences and expectations. We cannot let companies or NGOs detached from the day-to-day experiences of people determine what is an outcome for them and how it should be measured.

Overall, it will be essential that companies become involved but equally we need to ensure that the methodology is defined by a broader group of stakeholders with the interests of all societies being represented at the international level. It is natural to assume that international organizations such as the United Nations and UNESCO should take a leading role in this process, with UNESCO Category 2 centers such as IRCAI providing technical assistance.

TOWARDS A GLOBAL PARTNERSHIP

The range of expertise and geographies involved make the challenge of measuring SDGs a truly global one that requires the engagement of local teams of researchers in every region that can respond to the call to action. In this sense, we believe that bottom-up funding will be the most effective. This means distributed funding not coming from one funding source or body but diverse sources and scoping, including the size and amounts of contributions, ranging from open calls for technical solutions to micro projects at AI research institutions. The key to success is building trust in the approach and this cannot be imposed, but rather can be achieved only by creating a broad coalition including NGOs, companies, governments and international organizations. It is only by ensuring citizens everywhere feel represented that the support and trust of all societies can be commanded.

For this to be successful, a vital feature will be transparency in terms of what any given technology or solution can deliver. In other words, the description of the pluses and potential minuses, so that criticisms cannot create a narrative of "You are being misled." It is also vital that we create the common language of data to facilitate cross-partisan discussion and agreement on appropriate strategies – in other words, depoliticization of the discussion. It may be easier to achieve all of these desiderata if there is an initial focus on a single or small subset of SDGs, where perhaps the views are less polarized. By building trust in this setting, the opportunity would be created to extend to other more challenging SDGs.

This global partnership could initially be piloted by building a research community in sustainability and AI, via a network that strengthens AI research excellence centers across the world and facilitates collaboration and networking. The objective of this vibrant global network of AI excellence centers in sustainable development would be to boost the research capacity in this domain, and make it attractive for scientists and investors – both social impact and venture capital – and policymakers. This initiative is also expected to contribute to the development of ethical and trustworthy AI, as described in the UNESCO recommendations.

CONCLUSION

> We have argued for the need for the manifesto earlier in this chapter, but it is worth exploring what additional benefits might accrue from its successful realization. One useful analogy is the view that financial markets offer a very sophisticated machinery for ensuring that invested resources deliver the biggest financial return. The sustainable development agenda challenges the belief that this should be the only way in which investments should be measured, and we have argued that there is growing support for this view. However, there is no corresponding mechanism for measuring performance of companies against these new criteria. If we are to literally "put our money where our mouth is," we urgently need to create such mechanisms as our manifesto has urged. Only through the more effective use of data and AI can we avoid the "greenwashing" effect, where companies, via marketing and PR, spin claims to the public and their customers that they are delivering against the SDGs, when in reality they are not. More importantly, this will open up a robust and verifiable route for investors to support sustainable development and for companies to make the case of their products' value for society. It will also allow for companies to showcase their products' added value and potential savings they can bring to governments in terms of quantifiable improvements to SDGs, hence informing social impact bonds.

REFERENCES

Amaro, S. 2021. Blackrock's former sustainable investing chief says ESG is a dangerous placebo. August 24. CNBC. https://www.cnbc.com/2021/08/24/blackrocks-former-sustainable-investing-chief-says-esg-is-a-dangerous-placebo.html

ATI, 2021. Data-centric engineering. The Alan Turing Institute. https://www.turing.ac.uk/research/research-programmes/data-centric-engineering

Edelman. 2020. *21st Annual Edelman Trust Barometer*. https://www.edelman.com/sites/g/files/aatuss191/files/2021-01/2021-edelman-trust-barometer.pdf

Fairtrade Foundation. n.d. What Fairtrade does. https://www.fairtrade.org.uk/what-is-fairtrade/what-fairtrade-does/

Ferrando, T., Miola, I., Junqueira, G. O., Prol, G. M., Vecchione-Goncalves, M. and Herrera, H. 2021. Capitalizing on green debt. *Journal of World-Systems Research*, Vol. 27, No. 2, pp. 410–437.

Schölkopf, B. 2019. Causality for machine learning. *arXiv:1911.10500.* https://arxiv.org/abs/1911.10500

Social Finance. 2021. Cambodia rural sanitation: Bringing safe sanitation to rural communities in Cambodia. https://www.socialfinance.org.uk/projects/cambodia-rural-sanitation

Masakhane. 2022. Masakhane home page. https://www.masakhane.io/

UNESCO. 2020. Open Science home page https://www.unesco.org/en/natural-sciences/open-science

UNPRI. 2020. *Investing with SDG Outcomes: A Five-Part Framework*. https://www.unpri.org/sustainable-development-goals/investing-with-sdg-outcomes-a-five-part-framework/5895.article

UNPRI. 2022. About the PRI. https://www.unpri.org/about-us/about-the-pri

USAID. 2019. *The Cambodia Rural Sanitation DIB: Lessons Learnt from the First Year.* https://www.thesff.com/wp-content/uploads/2022/02/Development-Impact-Bond-lessons-learnt-March-2021-1.pdf

VideoLectures.NET. 2020. AI and Climate: Water Quality Measurement System. http://videolectures.net/IRCAILaunch2021_sullivan_ai_climate/

West, T. A. P., Börner, J., Sills, E. O. and Kontoleon, A. 2020. Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. *PNAS*, Vol. 117, No. 39, pp. 24188–24194. https://doi.org/10.1073/pnas.2004334117

Zamculture. 2019. Zambia school mapping project. https://github.com/Zamculture/Zambia-School-Mapping-Project-