Introduction

Artificial Intelligence for Social Good

Yoshiaki Fukami and Jiro Kokuryo
Keio University

1. Harnessing AI to Achieve the United Nations Sustainable Development Goals

We live in a complex world in which various factors affecting human wellness are interconnected and cannot be analyzed by simple models. For example, solutions to the challenges of pandemics require understanding of not just biology and/or medicine but of social activities, as well as the psychology of people who spread groundless or even malicious rumors on social media.

Expectations are high that artificial intelligence (AI) can help develop solutions to many issues facing the world by identifying patterns in the vast body of data that is now available through today's sensor networks. By enabling machines to identify and analyze patterns in data, we will be able to detect issues and causal relations in complex systems that were previously unknown. Such knowledge is essential in our efforts to overcome complex issues.

We should also be mindful that both wellness and these complex issues are embedded in local contexts that are diverse and depend on geographic and social backgrounds. While recognizing such diversity, it would be useful to have a meta-level understanding of how AI could be applied to accomplish our goals. An integrated and comprehensive vision, as well as its related policies, are needed to realize effective approaches for more people to enjoy the benefits of AI.

With this in mind, the United Nations (UN) has already begun to take a higher-level approach to solving social issues with AI. Set at the General Assembly (2015) and to be accomplished by 2030, the UN Sustainable Development Goals (SDGs) look to harness AI in support of inclusive and sustainable development while mitigating its risks. For example, SDGs look to:

- Provide people with access to data and information
- Support informed evidence-based decisions
- Eliminate inefficiencies in economic systems, as well as create new products and services to meet formerly unmet needs
- Provide data-driven diagnoses and prevent harmful events such as formerly unpredictable accidents
- Support city planning and development

This report understands AI for social good as being the use of AI to support SDG achievement by providing institutions and individuals with relevant data and analysis.

Table 1 is a non-exhaustive list of initiatives by the UN and other institutions to use AI in support of achieving SDGs. Supplemented with additional examples, the table mainly presents initiatives included in the UN Activities on Artificial Intelligence report by International Telecommunications Union (ITU, 2019). While the table presents projects that use AI for social good, it does not include initiatives that attempt to mitigate the risks of AI, such as to address bias or other ethical concerns.[1]

SDG		Use of AI
1	No Poverty	• Implementation of AI on the Global Risk Assessment Framework (GRAF) to understand future risk conditions to manage uncertainties and make data-driven decisions (ITU, 2019, p.54)
2	Zero Hunger	• FAMEWS global platform: Real-time situational overview with maps and analytics of Fall Armyworm infestations (ITU, 2019, p.3) • Sudden-onset Emergency Aerial Reconnaissance for Coordination of Humanitarian Intervention (SEARCH), and Rapid On-demand Analysis (RUDA) using drones and AI to greatly reduce the time required to understand the impact of a disaster (ITU, 2018, p.54)
3	Good Health and Well-being	• Ask Marlo: An AI chatbot designed to provide sources for HIV-related queries in Indonesia (ITU, 2019, p.22) • Timbre: a pulmonary tuberculosis screening by the sound of the cough (ITU, 2019, p.22)
4	Quality Education	• AI to ensure equitable access to education globally: Provide hyper-personal education for students and access to learning content (UNESCO, 2019, p.12) • Using AI and gamification to bridge language barriers for refugees: Machine learnt translation for lesser-resourced languages (UNESCO, 2019, p.11)
5	Gender Equality	• Sis bot chat: 24/7 information online services to women facing domestic violence (United Nations Women, 2019)

Table 1: Notable initiatives using AI in support of achieving SDGs
(Created by Daum Kim)

1. It should be noted that most projects supporting *Goal 5: Achieve gender equality and empower all women and girls* focus on removing gender bias. We only found one initiative using AI to empower women – a project that uses AI to fight against domestic violence.

6	Clean Water and Sanitation	• Water-related ecosystem monitoring through the Google Earth Engine and the European Commission's Joint Research Centre to use computer vision and machine learning to identify water bodies in satellite image data and map reservoirs (ITU, 2019, p.32) • Funding analysis and prediction platform using Microsoft's Azure Machine Learning Studio to capture global funding trends in the areas of environmental protection by donors and member states (ITU, 2019, p.32)
7	Affordable and Clean Energy	• Mitsubishi Hitachi Power Systems (MHPS) in the development of autonomous power plants: A real-time data monitoring action to reduce supply or increase generation and automated capability to manage power plants (Wood, 2019) • Intelligent grid system to increase energy efficiency through AI (Microsoft & PwC, 2019, p.17)
8	Decent Work and Economic Growth	• Analysis of the impact on jobs and employment by investigating the rise and effect of reprogrammable industrial robots in developing countries, along with exploration of patent data in robotics and AI to understand the future impact of AI robots on work (ITU, 2019, p.9)
9	Industry, Innovation, and Infrastructure	• E-navigation: Exchange and analysis of marine information on board and ashore by electronic means for safety and security at sea (ITU, 2019, p. 13) • Maritime Autonomous Surface Ships (MASS): Attempts to apply automated ships (ITU, 2019, p.13)
10	Reduced Inequalities	• Implementation of AI in a Displacement Tracking Matrix (DTM) to detect and contextualize data such as migration, urban and rural land classification, and drone imagery in displacement camps (ITU, 2019, p.16)
11	Sustainable Cities and Communities	• Risk Talk: An online community to exchange climate risk transfer solutions. AI builds a neural network by mapping the expertise of the users through interactions on the platform (ITU, 2019, p.37) • United for Smart Sustainable Cities initiatives (U4SSC): A global platform for smart cities stakeholders which advocates public policies to encourage the use of ICT to facilitate smart sustainable cities transition (ITU, 2019, p.29)
12	Responsible Consumption and Production	• AI-driven system and robotics to reduce food waste by predicting customer demand (Fearn, 2019) • iSharkFin: Identification of shark species from shark fin shapes to help users without formal taxonomic training (ITU, 2019, p.3)
13	Climate Action	• Shipping digitalization and electronic interchange with ports (ITU, 2019, p.12) • Cyber-consistent Adversarial Networks (CyberGans) to simulate what houses will look like after extreme weather events to allow individuals to make informed choices for their climate future (Snow, 2019; Schmidt et al., 2019)

(Cont.) Table 1: Notable initiatives using AI in support of achieving SDGs
(Created by Daum Kim)

14	Life Below Water	• Maritime Single Window (MSW) to electronically exchange maritime information via a single portal without duplication (ITU, 2019, p.12)
15	Life on Land	• DigitalGlobe's Geospatial Big Data platform (GBDX) using machine learning to analyze satellite imagery to predict human characteristics of a city and respond to health crises (ITU, 2018, p.50) • Land governance and road detection through satellite "computer vision" (ITU, 2018, p.60)
16	Peace, Justice, and Strong Institutions	• International Monitoring System of Comprehensive Nuclear-Test-Ban Treaty Organization (ITU, 2019, p.1) • Toolkit on digital technologies and mediation in armed conflict (ITU, 2019, p.27)
17	Partnerships	• The International Telecommunication Union (ITU) Focus Group on AI for Health (FG AI2H) (ITU, 2019, p.19) • The AI for Good Global Summit: Identifying practical applications of AI towards SDGs (ITU, 2019, p.19) • Social Media Data Scraper: AI on natural language processing helps to understand the thoughts of users (ITU, 2019, p.38)

(Cont.) Table 1: Notable initiatives using AI in support of achieving SDGs
(Created by Daum Kim)

2. Report Objectives: Research-based Policy Suggestions

Having reviewed how AI can be applied to promote social good, we now turn to policies that adequately promote and control AI, so that they can be used for the good of society. This is important, as we believe our goals cannot be accomplished through a laissez-faire approach. An adequate governance system for the development, management, and use of AI is crucial in ensuring that the benefits of integrating and analyzing large quantities of data are maximized, while the potential risks are mitigated.

Following an agreement between APRU, UN ESCAP, and Google to share best practices and identify solutions to promote AI for social good in Asia-Pacific, the project AI for Social Good was launched in December 2018 at the Asia-Pacific AI for Social Good Summit in Bangkok. Each chapter of this report presents a unique research project (Table 2), as well as key conclusions and policy suggestions based on the findings. The projects were selected following a competitive process that sought research inputs to inform policy discussions in two broad areas:

1. Governance frameworks that can help address risks/challenges associated with AI, while maximizing the potential of the technology to be developed and used for good.

2. Enabling environment in which policymakers can promote the growth of an AI for Social Good ecosystem in their respective countries in terms of AI inputs (e.g., data, computing power, and AI expertise) and ensuring that the benefits of AI are shared widely across society.

Focusing on specific local contexts and with the objective of informing international policy debates on AI, the research reports offer a range of unique perspectives from across the Asia-Pacific region.

Chapter	Title	Resaerch Member(s)	Affiliation
1	AI for Social Good: Buddhist Compassion as a Solution	Soraj Hongladarom	Chulalongkorn University, Thailand
2	Moralizing and Regulating Artificial Intelligence: Does Technology Uncertainty and Social Risk Tolerance Matter in Shaping Ethical Guidelines and Regulatory Frameworks	M. Jae Moon Iljoo Park	Yonsei University, Republic of Korea
3	Definition and Recognition of AI and its Influence on the Policy: Critical Review, Document Analysis and Learning from History	Kyoung Jun Lee	Kyung Hee University, Republic of Korea
4	Regulatory Interventions for Emerging Economies Governing the Use of Artificial Intelligence in Public Functions	Arindrajit Basu (Team leader) Elonnai Hickok Amber Sinha	Centre for Internet & Society, India
5	AI Technologies, Information Capacity, and Sustainable South World Trading	Mark Findlay	Singapore Management University
6	Governing Data-driven Innovation for Sustainability: Opportunities and Challenges of Regulatory Sandboxes for Smart Cities	Masaru Yarime	The Hong Kong University of Science and Technology
7	Including Women in AI-Enabled Smart Cities: Developing Gender-inclusive AI Policy and Practice in the Asia-Pacific Region	Caitlin Bentley	University of Sheffield, Australian National University
8	AI and the Future of Work: A Policy Framework for Transforming Job Disruption into Social Good for All	Wilson Wong	The Chinese University of Hong Kong

Table 2: List of project titles and their authors

The AI for Social Good Project believes that objective, evidence-based, and logical academic analyses which are free from political and/or economic interests can play critical roles in the formation of sensible policies. At the same time, we are aware of the tendency of academics to stop at simply understanding the phenomena and not take a position in prescribing policies. Hence, we specifically asked the participants of this report to come up with short summaries of their findings, as well as suggested policy implications (see Appendix 1).

We also firmly believe in the effectiveness of a multi-disciplinary research approach for policy formation. To that end, the project organizers were careful to include both the technical and social sciences/humanities. We are extremely happy to report that all of the diverse teams, who shared a similar passion for taking a multi-disciplinary approach, were able to conduct fruitful discussions which led to even stronger projects.

3. Overview of the Recommendations

Based on discussions with the project members, this section presents the editors' own overview of the policy agenda, giving readers a general idea of the issues that need to be addressed.

3.1. Developing a governance framework

3.1.1. Ensuring equality and equity

In Chapter 1, Hongladarom makes an important suggestion in that policymakers should start by agreeing on the basic principles for the governance of data. That is, he discusses how altruism, as opposed to individualism, should be seen as the guiding principle to realize the benefits of data sharing. He also emphasizes its usefulness in correcting existing social and economic inequalities, which may expand with advances in technology. While this assertion may be controversial, it nevertheless addresses the fundamental question of whether data should belong to the individual or society, since we know that the value of data increases as they accumulate. This line of thought is also significant in that it reflects the communal traditions of Asian societies.

In Wong's discussion of AI's impact on employment (Chapter 8), he also calls for social security policies and a fair re-allocation of resources in the governance of AI. The editors' interpretation of such calls for social equity surrounding AI is that there may be strong scale advantages in AI (or data) economy that give unfair advantages to already powerful entities; and that policy intervention is necessary for fairness and to ensure the productive power of AI is able to materialize. Bentley's call (Chapter 7) for the inclusion of women as beneficiaries of AI is also along the same lines.

3.1.2. Managing risk to allow experimentation

All of the researchers recognize the potential for AI to both benefit and cause harm to society. The problem is, we will not know for sure what the positive and the negative impacts might be until we test them. It is therefore necessary to formulate a bold strategy to realize full potential of AI and manage the risks involved at the same time.

In Chapter 6, Yarime looks at the possibility of taking a "sandbox" approach to testing. In this way, experimental use of technology can be undertaken for proof of concept in a controlled environment, and the results can then be used to take the technologies outside the "box" to be implemented in societies at large. He also discusses the importance of preparing mechanisms for compensation, such as insurance, to mitigate damage done to individuals or institutions despite all necessary preventative measures having been taken. This function is crucial, not just to protect citizens but also to promote innovation.

Uncertainty and unpredictability are inherent characteristics of emerging technologies and cannot be eliminated completely. It is worth remembering that we should not sacrifice innovation through excessive safety precautions. If we want to benefit from technological advancements, we must be willing to take certain risks. As such, we should be thinking about "managing" risk rather than "avoiding" risk.

3.1.3. Multi-stakeholder governance and co-regulation

In Chapter 2, Moon and Park call upon the participation of different stakeholders representing industries, researchers, consumers, NGOs, international organizations, and policymakers in setting guidelines for the ethical use of AI. Most AI applications require cooperation of multiple organizations, particularly in the preparation of integrated datasets. For example, automobile driving data from a car manufacturer are only useful when combined with other data sources. The value of such data is further enhanced when combined with data from local and national governments that control infrastructure, such as traffic lights. Each of these actors have different objectives and, in the absence of adequate incentives, tend to tailor their systems to maximize the effectiveness of their own services without regard for the needs of others. Thus, not only do we need mechanisms to promote collaboration, governments should play a role in preparing them.

Although a natural temptation under such circumstances is to centralize control, we must also be aware of the dangers of a centralized approach both technically and societally. On the technical side, centralized databases are vulnerable to attacks and can result in large-scale data leaks once the system is breached. On the societal side, a monopoly over data gives excessive power to the institution that controls it, raising fears of a breach of human rights. A multi-stakeholder governance structure involving government, non-profit organizations, industry groups, and specialist groups should be established to provide oversight of the major players controlling the data. It is important that young policymakers and engineers participate in the discussion (Chapter 5). Given the rapid advances in technology, we must also develop and establish governance mechanisms that can evolve in a timely manner.

3.1.4. Providing accountability

Basu, Hickok, and Sinha (Chapter 4) identify accountability as one of five major areas where states should play a role. This is an extremely important point in light of the fact that AI can easily become a "black box" both technically and institutionally.

Accountability is a fundamental issue across various aspects of AI utilization, from the collection of data to the determination of evaluation functions in AI algorithms. As such, it is vital that we review and evaluate the process by which AI functions, as well as identify appropriate entities to manage the technology.

Accountability must be realized not only through legal systems, but also in the technical specifications of systems that ensure transparency of data management. Due to the pace of technological advancement, this is a challenge. Hence, governments need to assist in the development of a coordination mechanism that can cope with the progress in a timely manner.

3.2. Developing an enabling environment

3.2.1 Correctly understanding the technology

In Chapter 3, Lee cautions that, before discussing policies concerning AI, we should first have a proper understanding of the definition of AI. He points out the dangers of perceiving AI as simply machines that imitate and replace humans. Instead, he favors the perspective of the Organization for Economic Co-operation and Development (2019) that defines AI as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments" to form adequate expectations for the benefits of the technology.

An adequate definition of AI is therefore important, as it greatly influences the design of the governance

structure around the technology. Whether or not we recognize "intelligence" and "personality" (or at least legal personality as we recognize corporations as pseudo-personalities) in machines that seemingly have an intelligence of their own is becoming a serious topic of debate. If we are to adopt Lee's argument, then perhaps we should not.

3.2.2. Ensuring universal access to data

In Chapter 5, Findlay looks at how information asymmetries can create inequities for disadvantaged economies, and calls for systems to guarantee them access to data which enables them to negotiate fairly in international trade. This reminds us that AI cannot work on its own. In the application of AI, datasets, computing power, and expert analysts are all necessary to meet society's needs.

Naturally, the opportunities which computer networks create should not be underestimated. Recent advances in the reduction of communication costs, improvement of computing capabilities, and diffusion of sensing technology have facilitated the generation of big data that can then be analyzed by data scientists. Findlay's concern over inequity is especially important as there still remain many areas where access to essential data are limited and necessary data analyses are not possible. No matter how sophisticated the AI algorithm, it can only work effectively in an environment in which the dataset is properly generated and stored for analyses, there is the necessary computing power, and there is reliable and affordable access to expertise and the Internet.

It is worth remembering that network ubiquity does not exist yet either. There are still many people in Asia-Pacific that do not have access to reliable, affordable, and high-speed Internet. As such, governments should continue their efforts to provide everyone with Internet connectivity so that they have access to the data that empowers them.

3.2.3. Standardizing data models

Standardization of data formats is important in order to ensure universal access to data for a more equitable use of the technology. Not only does the differences in data models (formats) hinder data integration, a lack of standardization nullifies the power of ubiquitous Internet connectivity that enables us to gather data quickly and cheaply. In other words, aggregated data does not automatically mean big data suitable for AI analysis. Data must still be standardized to be collectively meaningful. In addition, data specifications (e.g., syntax and vocabulary) facilitate interoperability among distributed data resources and enable the generation of relevant big data. Furthermore, quality criteria enable data consumers to appropriately handle diversified data resources.

However, standardization is a complex issue, not because it is technically difficult but because it is a political process involving many different stakeholders, pursuing different goals. Therefore, a top down approach to forcefully impose a single set of standards will not work. That said, governments should still play a facilitator role, together with many non-governmental standardization initiatives, to prevent excessive proliferation of standards across every sector of society. Governments should also ensure interoperability among systems that of different standards.

3.2.4. Universal access to human resources for utilization of AI

Findlay also stresses the need for adequate assistance (e.g., technology, training, and domestic policy advice) to fully realize the benefits of AI. This is a reminder that AI systems require people to function. In other words, effective use of AI requires people to fine tune the algorithm and prepare the dataset to be fed into the system. It is also necessary for people to interpret the outcome and give it practical meaning. As the use of AI grows, so too does the demand for data scientists who can use the technology for social good.

However, as data scientists are fast becoming an expensive human resource only available to more developed economies and large corporations, the fewer number of them in less fortunate communities is limiting the opportunities to make use of AI.

When talking about human resources, it is important to recognize that not just software engineers and expert statisticians need to be trained. Senior executives and ordinary people also need to be aware of the benefits, risks, and mitigation measures surrounding AI, so that they are better informed and able to take advantage of the technology.

Another aspect is the need to educate engineers about the ethical, legal, and social implications (ELSI) of AI. As the power of AI grows, so too does its impact on ELSI. For the technology to be developed and used properly, governments need to ensure that technical experts are educated to be sensitive to the concerns of ordinary people concerning AI.

3.2.5. Removing the fear of using personal data

Another policy goal that the editors would like to propose is the removal of (perceived) risk associated with personal data disclosure. We believe that it is important to make available as much data as possible for the use of AI for social good. Of course, this is only achievable when people feel safe about disclosing their information.

There are two main reasons why citizens and consumers are currently holding back from offering their data for social good. First, they fear that data disclosure can lead to discrimination. This is especially true in socially sensitive areas. For example, when disclosure of infection to a disease leads to exposure to social stigma and criticism for non-compliance to social norms, people will be reluctant to cooperate with contact tracing. Second, certain consumers dislike the idea of having their data commercially exploited without their consent.[2] For example, the emergence of target marketing as the key revenue generator for online businesses has led to significant hostility towards the use of personal data.

To address this issue there are technical and institutional solutions available. On the technology side, various forms of anonymization, encryption, and distributed approaches in managing data have been proposed. Institutionally, various forms of regulations are in place to protect individuals from breach of privacy. For both types of solutions, government involvement seems essential in light of the incentives that exist, particularly in the private sector, to keep data secret for financial reasons. Not only should incentives be offered to make data public, but enforcement power must be used in the protection of privacy.

2. We should also be aware of people who are willing to give their information away for free, because they feel compelled or see a benefit in doing so.

References

International Telecommunication Union. (2018). United Nations Activities on Artificial Intelligence (AI) 2018. https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2018-1-PDF-E.pdf

International Telecommunication Union. (2019). United Nations Activities on Artificial Intelligence (AI) 2019. https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf

Microsoft & PwC. (2019). How AI can enable a sustainable future. https://www.pwc.co.uk/sustainability-climate-change/assets/pdf/how-ai-can-enable-a-sustainable-future.pdf

OECD 2019, "Recommendation of the Council on Artificial Intelligence", https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

Schmidt, V., Luccioni, A., Mukkavilli, S.K., Sankaran, K., Bengio, Y. (2019). Visualising the Consequences of Climate Change Using Cycle-Consistent Adversarial Networks. https://arxiv.org/pdf/1905.03709.pdf

Snow, J. (2019). How artificial intelligence can tackle climate change. https://www.nationalgeographic.com/environment/2019/07/artificial-intelligence-climate-change/

United Nations Educational, Scientific, and Cultural Organization (2019). Artificial intelligence in education, compendium of promising initiatives: Mobile Learning Week 2019 https://unesdoc.unesco.org/ark:/48223/pf0000370307

United Nations General Assembly. (2015). Transforming our world: the 2030 Agenda for Sustainable Development. https://doi.org/10.1163/157180910X12665776638740

United Nations Women. (2019). Using AI in accessing justice for survivors of violence. https://www.unwomen.org/en/news/stories/2019/5/feature-using-ai-in-accessing-justice-for-survivors-of-violence

Wood, J. (2019). This is how AI is changing energy. https://spectra.mhi.com/this-is-how-ai-is-changing-energy

Appendix 1

Summaries of Papers and Policy Suggestions

AI for Social Good: A Buddhist Compassion as a Solution

Soraj Hongladarom, Department of Philosophy, Faculty of Arts, Chulalongkorn University

Abstract

In this paper, I argue that in order for AI to deliver social good, it must be ethical first. I employ the Buddhist notion of compassion (*karunā*) and argue that for anything to be ethical, it must exhibit the qualities that characterize compassion, namely the realization that everything is interdependent and the commitment to alleviating suffering in others. The seemingly incoherent notion that a thing (e.g., an AI machine or algorithm) can be compassionate is solved by the view—at this current stage of development—that algorithm programmers need to be compassionate. This does not mean that a machine cannot itself become compassionate in another sense. For instance, it can become compassionate if it exhibits the qualities of a compassionate being. Ultimately, it does not matter whether or not a machine is conscious in the normal sense. As long as the machine exhibits the outward characterization of interdependence and altruism, it can be said to be compassionate. I also argue that the ethics of AI must be integral to the coding of its program. In other words, the ethics—how we would like the AI to behave based on our own ethical beliefs—needs to be programmed into the AI software from the very beginning. I also reply to several objections against this idea. In essence, coding ethics into a machine does not imply that such ethics belongs solely to the programmer, nor does it mean that the machine is thereby completely estranged from its socio-cultural context.

Policy Recommendations

1. **Programmers and software companies need to implement compassionate AI programs.** This is the key message from this article. No matter what kind of "social good" the AI is supposed to bring about, the software needs to be compassionate and ethical in the Buddhist sense.
2. **The public sector needs to ensure that rules and regulations are in place in order to create an environment that facilitates the development of ethical AI for social good.** Such rules and regulations will ensure that private companies have a clear set of directives to follow, and will create public trust in the works of the private sector.

Moralizing and Regulating Artificial Intelligence: Does Technology Uncertainty and Social Risk Tolerance Matter in Shaping Ethical Guidelines and Regulatory Frameworks?

M. Jae Moon and Iljoo Park, Institute for Future Government, Yonsei University

Examining technology uncertainty and social risk in the context of disruptive technologies, this study reviews the development of ethical guidelines for AI developed by different actors as a loosely institutional effort to moralize AI technologies. Next, we specifically examine the different regulatory positions of four selected countries on autonomous vehicles (AVs). Based on the status of moralizing and regulating AI, several policy implications are presented as follows:

1. Moralizing disruptive technologies should precede, and should be fully discussed and shared among different stakeholder prior to regulating them. Before a society adopts and enacts specific regulatory frameworks for disruptive technologies, ethical guidelines (i.e., AI principles or AI ethical guidelines) must be jointly formulated based upon a thorough deliberation of particular disruptive technologies by different stakeholders representing industries, researchers, consumers, NGOs, international organizations, and policymakers.
2. AI ethical guidelines should support sustainable and human-centric societies by minimizing the negative socio-economic and international consequences of disruptive technologies (i.e., inequality, unemployment, psychological problems, etc.), while maximizing their potential benefits for environmental sustainability, quality of life among others.
3. Once a general consensus is made on general ethical guidelines, they should be elaborated and specified in details targeting individual stakeholder groups representing different actors and sectors. Specific AI ethical guidelines should be developed and customized for AI designers, developers, adopters, users, etc. based on the AI lifecycle. In addition, industry and sector specific ethical guidelines should be developed and applied to each sector (care industry, manufacturing industry, service industry, etc.).
4. In regulating AI and other disruptive technologies, governments should align regulations with key values and goals embedded in various AI ethical guidelines (transparency, trustworthiness, lawfulness, fairness, security, accountability, robustness, etc.) and aim to minimize the potential social risks and negative consequences of AI by preventing and restricting possible data abuses or misuses, ensuring fair and transparent algorithms, in addition to establishing institutional and financial mechanisms through which the negative consequences of AI are systematically corrected.
5. Governments should ensure the quality of AI ecosystems by increasing government and non-government investment in R&D and human resources for AI by maintaining fair market competition among AI-related private companies, and by promoting AI utilities for social and economic benefits.
6. Governments should carefully design and introduce regulatory sandbox approaches to prevent unnecessarily strict and obstructive regulations that may impede AI industries but also facilitate developing AI and exploring AI-related innovative business models.

Definition and Recognition of AI and its Influence on the Policy: Critical Review, Document Analysis and Learning from History

Kyoung Jun Lee, School of Management, Kyung Hee University
Yujeong Hwangbo, Department of Social Network Science, Kyung Hee University

Abstract

Opacity of definitions hinders policy consensus; and while legal and policy measures require agreed definitions, to what artificial intelligence (AI) refers has not been made clear, especially in policy discussions. Incorrect or unscientific recognition of AI is still pervasive and misleads policymakers. Based on a critical review of AI definitions in research and business, this paper suggests a scientific definition of AI. AI is a discipline devoted to making entities (i.e., agents and principals) and infrastructures intelligent. That intelligence is the quality which enables entities and infrastructures to function (not think) appropriately (not humanlike) as an agent, principal, or infrastructure. We report that the Organisation for Economic Co-operation and Development (OECD) changed its definition of AI in 2017, and how it has since improved it from "humanlike" to "rational" and from "thinking" to "action". We perform document analysis of numerous AI-related policy materials, especially dealing with the job impacts of AI, and find that many documents which view AI as a system that "mimics humans" are likely to over-emphasize the job loss incurred by AI. Most job loss reports have either a "humanlike" definition, "human-comparable" definition, or "no definition". We do not find "job loss" reports that rationally define AI, except for Russell (2019). Furthermore, by learning from history, we show that automation technology such as photography, automobiles, ATMs, and Internet intermediation did not reduce human jobs. Instead, we confirm that automation technologies, as well as AI, creates numerous jobs and industries, on which our future AI policies should focus. Similar to how machine learning systems learn from valid data, AI policymakers should learn from history to gain a scientific understanding of AI and an exact understanding of the effects of automation technologies. Ultimately, good AI policy comes from a good understanding of AI.

Policy Recommendations

1. Policy experts should be well educated about what AI is and what is really going on in AI research and business. Specifically, AI should be considered a discipline that allows entities and infrastructures to become intelligent. This intelligence is the quality that enables agents, principals, and infrastructures to function appropriately. AI should not be considered a humanlike or super-human system. As such, previous AI policies based on the old paradigm should be rewritten.
2. Governments should create programs to educate administrative officials, policy experts in public-owned research institutes, and lawmakers in national assemblies.
3. Similar to how machine learning systems learn from valid data, policymakers should learn from history, as well as recognize the positive impacts of automation technology. New AI policies should then be established based on this new recognition.
4. When adopting AI, governments and society should recognize its characteristics as an optimization system in order to create more public benefit, faster business outcomes, and less risk.

Regulatory Interventions for Guiding and Governing the Use of Artificial Intelligence by Public Authorities

Arindrajit Basu, Elonnai Hickok and Amber Sinha, Centre for Internet & Society, India

Summary

The use of artificial intelligence (AI)-driven decision-making in public functions has been touted around the world as a means of augmenting human capacities, removing bureaucratic fetters, and benefiting society. This certainly holds true for emerging economies. Due to a lack of government capacity to implement these projects in their entirety, many private sector organizations are involved in traditionally public functions, such as policing, education, and banking. AI-driven solutions are never "one-size-fits-all" and exist in symbiosis with the socio-economic context in which they are devised and implemented. As such, it is difficult to create a single overarching regulatory framework for the development and use of AI in any country, especially those with diverse socio-economic demographics like India. Configuring the appropriate regulatory framework for AI correctly is important. Heavy-handed regulation or regulatory uncertainty might act as a disincentive for innovation due to compliance fatigue or fear of liability. Similarly, regulatory laxity or forbearance might result in the dilution of safeguards, resulting in a violation of constitutional rights and human dignity. By identifying core constitutional values that should be protected, this paper develops guiding questions to devise a strategy that can adequately chart out a regulatory framework before an AI solution is deployed in a use case. This paper then goes on to test the regulatory framework against three Indian use cases studied in detail – predictive policing, credit rating, and agriculture.

Key Recommendations

1. To adequately regulate AI in public functions, regulation cannot be entirely "responsive" as the negative fall out of the use case may be debilitating and greatly harm constitutional values. We therefore advocate for "smart regulation" – a notion of regulatory pluralism that fosters flexible and innovative regulatory frameworks by using multiple policy instruments, strategies, techniques, and opportunities to complement each other.
2. The five key values that must be protected by the state across emerging economies are: (1) agency; (2) equality, dignity, and non-discrimination; (3) safety, security and human impact; (4) accountability, oversight, and redress; and (5) privacy and data protection.
3. The scope, nature, and extent of regulatory interventions should be determined by a set of guiding questions, each of which has implications for one or more of constitutional values.
4. Whenever the private sector is involved in a "public function", either through a public–private partnership or in a consultation capacity, clear modes, frameworks, and channels of liability must be fixed through uniform contracts. The government may choose to absorb some of the liability from the private actor. However, if that is the case, this must be clearly specified in the contract and clear models of grievance redressal should be highlighted.
5. The case studies point to a need for constant empirical assessment of socio-economic and demographic conditions before implementing AI-based solutions.

6. Instead of replacing existing processes in their entirety, decision-making concerning AI should always look to identify a specific gap in an existing process and add AI to augment efficiency.
7. The government must be open to feedback and scrutiny from private sector and civil society organizations, as that will foster the requisite amount of transparency, trust, and awareness regarding the solution – all of which are challenges in emerging economies.
8. In situations where the likelihood or severity of harm cannot be reasonably ascertained, we recommend adopting the precautionary principle from environmental law and suggest that the solution not be implemented until scientific knowledge reaches a stage where it can reasonably be ascertained.

VALUE	QUESTIONS
AGENCY	Is the adoption of the solution mandatory?
	Does the solution allow for end-user control?
	Is there a vast disparity between primary user and impacted party?
EQUALITY, DIGNITY, AND NON-DISCRIMINATION	Is the AI solution modelling or predicting human behavior?
	Is the AI solution likely to impact minority, protected, or at-risk groups?
SAFETY, SECURITY, AND HUMAN IMPACT	Is there a high likelihood or high severity of potential adverse human impact as a result of the AI solution?
	Can the likelihood or severity of adverse impact be reasonably ascertained with existing scientific knowledge?
ACCOUNTABILITY, OVERSIGHT, AND REDRESS	To what extent is the AI solution built with "human-in-the-loop" supervision prospects?
	Are there reliable means for retrospective adequation?
	Is the private sector partner involved with either the design of the AI solution, its deployment, or both?
PRIVACY AND DATA PROTECTION	Does the AI solution use personalized data, even in anonymized form?

AI Technologies, Information Capacity, and Sustainable South World Trading

Mark Findlay, Singapore Management University, School of Law – Centre for AI and Data Governance

This research is supported by the National Research Foundation, Singapore under its Emerging Areas Research Projects (EARP) Funding Initiative. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the National Research Foundation, Singapore.

Abstract

This paper represents a unique research methodology for testing the assumption that AI-assisted information technologies can empower vulnerable economies in trading negotiations. Its social good outcome is enhanced through additionally enabling these economies to employ the technology for evaluating more sustainable domestic market protections. The paper is in two parts; the first presents the argument and its underpinning assumption that information asymmetries jeopardize vulnerable economies in trade negotiations and decisions about domestic sustainability. We seek to use AI-assisted information technologies to upend situations where power is the discriminator in trade negotiations because of structural information deficits, and where the outcome of such deficits is the economic disadvantage of vulnerable stakeholders. The second section is a summary of the empirical work piloting a more expansive engagement with trade negotiators and AI developers. The empirical project provides a roadmap for policymakers to adopt model reflections from focus groups and translate these into a real-world research experience. The research method has three phases, designed to include a diverse set of stakeholders – a scoping exercise, a solution exercise, and a strategic policy exercise. The empirical achievement of this paper is validating the proposed action-oriented methodology through a "shadowing" pilot device, where representative groups engaged their role-plays and represented essential understandings. General findings from the two focus groups are provided.

Principal Policy Projections

- At the initiation of the project, **an intensive needs analysis should be initiated**, grounded in developing local skills around what questions to ask regarding information deficit, then translating into learning about what format to store and order data, and what data can accomplish in trading negotiations and domestic market sustainability. This exercise will empower domestic counterparts and achieve ownership. This exercise should be a collaboration between ESCAP, sponsor companies, and agencies;
- **Trading information asymmetries should be addressed** by sponsor companies, donors, and associated international agencies, through AI-assisted technologies for domestically empowering information access capacity building. **UN ESCAP should promote the use of AI-assisted technologies** to flatten information asymmetries that exist among trading partners in the region;
- While AI has the potential for empowering presently disadvantaged economies to negotiate in equal terms to raise the well-being of all people, such empowerment will not materialize without **adequate assistance**, in the form of technology, training, and domestic policy advice;
- **Product sustainability** is essential for the success of the project ongoing. Sponsor companies, and ESCAP in oversight, should ensure certain crucially sustainable deliverables covering: *data sources, data integrity and validation, accountability, and the technical sustainability of technical products.* These issues require allied services from sponsors, providers, advisers, and locally trained experts.

Governing Data-driven Innovation for Sustainability: Opportunities and Challenges of Regulatory Sandboxes for Smart Cities

Masaru Yarime, Division of Public Policy, The Hong Kong University of Science and Technology

Abstract

Data-driven innovation plays a crucial role in tackling sustainability issues. Governing data-driven innovation is a critical challenge in the context of accelerating technological progress and deepening interconnection and interdependence. AI-based innovation becomes robust by involving the stakeholders who will interact with the technology early in development, obtaining a deep understanding of their needs, expectations, values, and preferences, and testing ideas and prototypes with them throughout the entire process. The approach of regulatory sandboxes plays an essential role in governing data-driven innovation in smart cities, which faces a difficult challenge of collecting, sharing, and using various kinds of data for innovation while addressing societal concerns about privacy and security. How regulatory sandboxes are designed and implemented can be locally adjusted, based on the specificities of the economic and social conditions, to maximize the effect of learning through trial and error. Regulatory sandboxes need to be both flexible to accommodate the uncertainties of innovation, and precise enough to impose society's preferences on emerging innovation, functioning as a nexus of top-down strategic planning and bottom-up entrepreneurial initiatives. Data governance is critical to maximizing the potential of data-driven innovation while minimizing risks to individuals and communities. With data trusts, the organizations that collect and hold data permit an independent institution to make decisions about who has access to data under what conditions, how that data is used and shared and for what purposes, and who can benefit from it. A data linkage platform can facilitate close coordination between the various services provided and the data stored in a distributed manner, without maintaining an extensive central database. As the provision of personal data would require the consent of people, it needs to be clear and transparent to relevant stakeholders how decisions can be made in procedures concerning the use of personal data for public purposes. The process of building a consensus among residents needs to be well-integrated into the planning of smart cities, with the methodologies and procedures for consensus-building specified and institutionalized in an open and inclusive manner. As application programming interfaces (APIs) play a crucial role in facilitating interoperability and data flow in smart cities, open APIs will facilitate the efficient connection of various kinds of data and services.

Policy Recommendations

1. Data governance of smart cities should be open, transparent, and inclusive to facilitate data sharing and integration for data-driven innovation while addressing societal concerns about security and privacy.
2. The procedures for obtaining consent on the collection and management of personal data should be clear and transparent to relevant stakeholders with specific conditions for the use of data for public purposes.
3. The process of building a consensus among residents should be well-integrated into the planning of smart cities, with the methodologies and procedures for consensus-building specified and institutionalized in an open and inclusive manner.
4. APIs should be open to facilitate interoperability and data flow for efficient connection of various kinds of data and sophisticated services in smart cities.

Including Women in AI-enabled Smart Cities: Developing Gender-inclusive AI Policy and Practice in the Asia-Pacific Region

Caitlin Bentley, Katrina Ashton, Brenda Martin, Elizabeth Williams, Ellen O'Brien, Alex Zafiroglu, and Katherine Daniell, 3A Institute, Australian National University

Smart city initiatives are widespread across the Asia-Pacific region. AI is increasingly being used to augment and scale smart city applications in ways that can potentially support social good. We critically reviewed the literature on two key AI applications for social good: increasing safety and security in public spaces through the use of facial recognition technology, and improving mobility through AI-enabled transportation systems including smart traffic lights and public transportation route optimization. We find that there is an urgent need to consider how best to include women in the design, development, management, and regulation of AI-enabled smart cities. After all, poorly designed or delivered AI-enabled smart city technology could potentially negatively and differentially impact women's safety, security, and mobility. To address these pitfalls, we conducted interviews with a range of female and feminist scholars, activists, and practitioners – many of whom are working in the technology space. We carried out an analysis using the 3A Framework. This Framework focuses on investigating smart city initiatives through the themes of agency, autonomy, assurance, interfaces, indicators, and intent. We suggest the following actions be required: (1) commit to gender inclusive policymaking and praxis in national smart city policy; (2) institute formal consultation and participatory processes involving diverse women and community representatives through all stages of a smart city initiative; and (3) devise clearer roles and responsibilities surrounding the protection and empowerment of women in AI-enabled smart city initiatives.

1. **Commit to gender inclusive policymaking and praxis in national smart city policy:** High-level national smart city documentation frequently makes reference to social inclusion goals, but little is mentioned on how social inclusion is practiced. AI-enabled smart cities involve an interlaced network of actors, such as government ministries, private sector actors, and community groups. Governments can play a key coordination role, whilst guiding the establishment of common goals and practices. Moreover, countries across Asia-Pacific should review national policy to take into account the interconnected nature of smart city initiatives, and how they connect to multiple targets across the Sustainable Development Goals (SDGs). National governments should institute a process to develop indicators that map smart city progress in the pursuit of achieving SDGs, namely SDG 5 and 11.
2. **Institute formal consultation and participatory processes involving diverse women and community representatives through all stages of a smart city initiative:** Our research identifies new models of design, community ownership, and public debate supported by AI. Municipal actors, industry partners, and women's community groups should invest greater resources into experimenting with innovative engagement and representation models, as well as building into project plans the time needed for engagement. The 3A Framework can be used to guide discussions with communities, women, and their representatives. Our research highlights how the Framework sheds lights on multiple and interrelated systemic factors that need to be taken into consideration, rather than focusing only on the perspectives of individuals.
3. **Devise clearer roles and responsibilities surrounding the protection and empowerment of women in AI-enabled smart city initiatives:** There is an urgent need for policymakers to establish greater transparency and clearer rules around the handling, ownership, and protection of data with, for, and about women. Better understanding of the impacts, not only the performance of these systems, should guide this discussion. Consequences for mistreatment, harm, and mismanagement across all levels of smart city initiatives should be carefully and clearly outlined. More opportunities for women to be consulted and involved in the design, management, evaluation, and regulation of AI-enabled smart city initiatives are warranted.

AI and the Future of Work: A Policy Framework for Transforming Job Disruption into Social Good for All

Wilson Wong, Chinese University of Hong Kongy

Abstract

This paper examines the impact of artificial intelligence (AI) on the future of work to develop a policy framework for transforming job disruption caused by AI into social good for all. While there is a considerable amount of research and discussion on the impact of AI on employment, there is relatively less research on what governments should do to turn the risk and threat of AI into job opportunities and social good for all. This paper consists of two major parts. It first builds on the typology of job replacement and AI to establish a policy framework on the role of the government, as well as the policy responses it should make to address various concerns and challenges. On the principle of "rise with AI, not race with it", the government must play an active or even aggressive role not only for retraining knowledge, skill-building, and job re-creation, but also for social security and a fair re-allocation of resources in the job disruption process. Second, the paper conducts a survey of national AI strategies to assess the extent to which AI policy of job disruption is addressed by other countries. It concludes that many countries, especially developing ones, are not well-prepared for AI, and most countries seem to be overlooking fairness and equity issues under job disruption in the arrival of the AI era.

Policy Summary: Major Recommendations

1. Theory and Practice: Governments should have more alignment and integration between theory and policy in formatting their AI strategies. For example, they should discuss how enabling technologies as well as social and creative intelligence are included in their retraining, reskilling, and education programs.
2. International Organization and Developing World: AI impacts on both developed and developing worlds. Many developing countries are ill-prepared due to limitations in resources and other factors. International organizations such as the United Nations (UN) should offer more support to these nations to help set up their own AI strategies to evaluate threats and opportunities and formulate solutions.
3. AI for All (No One Left Behind): Equity, social security, and fair re-distribution, such as introducing Universal Basic Income (UBI) to protect vulnerable populations, are the missing pieces in the AI strategies of most countries. Governments should confront these important issues head on and incorporate them explicitly in their national AI strategies.

Appendix 2

Project History

The AI for Social Good Project is the heir to two series of policy advocacy initiatives on the digital economy by the Association of Pacific Rim Universities (APRU). The first series is the Digital Economy initiative and its successor, the AI for Everyone project, hosted by Keio University. The second series, led by The Hong Kong University of Science and Technology, is "Transformation of Work in Asia Pacific in the 21st Century: Key Policy Implications".

The project also stems from the partnership UN ESCAP has been building with ARTNET on STI Policy – a regional research and training network supporting policy research to leverage science, technology, and innovation as powerful engines for sustainable development in Asia Pacific.

In addition to the authors represented in this project, the following advisory board members, to whom we are extremely grateful for their valuable input, were chosen to provide feedback about the projects.

Name	Affiliation
Hideaki Shiroyama	The University of Tokyo
Pascale Fung	The Hong Kong University of Science and Technology
Toni Erskine	Australia National University
Yudho Giri Sucahyo	University of Indonesia
P. Anandan	Wadhwani Institute of AI, Mumbai
Hoyoung Lee	Korea Information Society Development Institute
Punit Shukla	World Economic Forum
Yongyuth Yuthavong	National Science and Technology Development Agency

Table 1: List of advisory board members

To kick-off this collaborative project, the first face-to-face meeting was held on June 5, 2019 at Keio University's Mita campus. A virtual policy fora for the dissemination and discussion of project findings is planned to be held later in the year.

One last face-to-face meeting before final submission of the output, together with an open-to-public forum, was originally scheduled for February 20 – 21, 2020. However, due to the COVID-19 pandemic, it was replaced by an online meeting of just the project members. The project outputs were submitted in May 2020 for editing and subsequent publication in August 2020. When it is safe to do so, an open-to-public forum will be held.

The project was organized by the following members:

Name	Affiliation
Jiro Kokuryo, Project Coordinator	Keio University
Yoshiaki Fukumi	Keio University
Cherry Wong	Keio University
Daum Kim	Keio University
Minkyoung Cho	Keio University
Christina Schönleber	APRU
Tina Lin	APRU
Sanghyun Lee	Google
Jake Lucchi	Google
Marta Perez Cuso	UN ESCAP

Table 2: Organizing members

We are grateful for all the efforts of those involved and sincerely hope that this document will help policymakers in the region accomplish their goals.