Artificial Intelligence (AI) for Development Series

Introductory module

July 2018

Work in progress, for discussion purposes
Comments are welcome!
Please send your comments on this paper at: gsr@itu.int by 30 July 2018

The views expressed in this paper are those of the author and do not necessarily reflect the opinions of ITU or its Membership.

This Introductory module was prepared by Malcom Webb, M Webb Ltd, under the direction of the ITU/BDT Regulatory and Market Environment Division, in collaboration with the ITU/BDT Telecommunication Division and under close coordination with the Chief of the ITU/BDT Infrastructure, Enabling Environment, and E-Applications Department. We would like to thank the ITU General Secretariat and the ITU Standardization Bureau for their contributions.

The author would like to thank Callum Webb for his research and valuable insights that substantially contributed to this paper. .

Suggested citation. AI for Development Series. Geneva: International Telecommunication Union, 2018. Licence: CC BY-NC-SA 3.0 IGO.

1. The AI Series

The Telecommunication Development Bureau (**BDT**) of the International Telecommunications Union are promoting an initiative to deepen the understanding, and promote further discussion and collaboration, among policy makers and regulators of the significance of artificial intelligence (**AI**) and the policy and regulatory issues that are beginning to emerge from the development of AI.

The AI Series is a part of this initiative. The AI Series includes:

- This introductory module which introduces some of the key aspects of AI and the important policy and regulatory issues that arise and that are discussed elsewhere in the AI Series;
- A module on AI governance examining governance strategies for AI to limit the risks arising from these innovative applications and helping to unlock their opportunities;
- A module on the ethical and societal issues arising from AI; and
- A module on the relevance of AI in the current and future development of the Internet of Things (IoT) and how security should be addressed, including in relation to data protection and privacy.

2. Introduction to AI

There is no accepted definition of artificial intelligence (**AI**). Professor Nilsson, from Stanford University, describes AI, and intelligence, as follows:

> *"Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment"*[1]

At the ITU's AI for Good Summit 2017, AI was described as:

> *"… a set of associated technologies and techniques that can be used to complement traditional approaches, human intelligence and analytics and/or other techniques".*

AI comprises a broad range of computational technologies, some of which are developments of existing technologies and some brand new.

One of the themes of this AI Series is that policy makers and regulators need to increase their understanding of AI technologies and the policy implications of this technology, exchange experiences and discuss possible governance and regulatory frameworks to capture the benefits of AI and address its challenges. It will be important for policy makers and regulators to develop a cross-sectoral and interdisciplinary approach to facilitate AI.

Although the term AI has only recently come into widespread public consciousness, AI itself is not new. AI traces its roots back over 50 years. However, AI systems and their use in many fields have developed significantly in the last few years, revealing the true potential of this technology. It may still be debated whether AI is a revolutionary technology. Many experts in the field believe so. Whether AI is revolutionary, or an evolutionary technology, it shines light on a wide range of policy

[1] Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge, UK: Cambridge University Press, 2010).

and regulatory issues from a different angle and with a more intense focus than we have seen with other technologies.

As illustrated in the module on AI in society "These two factors – a growing global ubiquity and an emerging set of risks and rewards – is why AI presents such a wide array of increasingly sticky ethical and societal concerns".

AI is popularly used today in a narrow and relatively basic form, including image and voice recognition systems; Siri and Alexa; Amazon and Netflix recommendations; subtitles on over a billion YouTube videos; fraud detection by credit card companies, etc. Today, AI systems recognise images and words as well as, if not better than, most humans.

Robots and AI are different. Robots are automated, but they are not usually autonomous. Automated means being able to do physical or mental work that could have been done by a person . It generally involves repetitive tasks. Autonomous systems are designed to operate in changing circumstances without human control. They look for patterns and learn from their experience, without following a programmed set of instructions. Normally, automated systems do not use AI, but increasingly robots and other automated systems will use AI in performing manual or cognitive tasks.

In this AI Series, when we refer to AI, we do not usually mean "artificial general intelligence", which may be defined as the ability to do any intellectual task that a human is capable of. It is too early in the development of the technology to consider general AI in any substantive way.

Rather, this AI Series generally refers to "narrow AI", which are narrower applications of human-like intelligence[2].

> *"AI products tend to evolve from laughably weak to interesting but feeble, then to artificial but useful, and finally to transcendent and superior to humans"*[3]

3. Applications of AI

There are **three key things** that are propelling momentum for AI today: the availability of far greater quantities of data, increased computer processing power (particularly cloud computing) and algorithmic advances. We may also add to this list the increasing ubiquity of high speed broadband networks.

a) Current and potential applications

Current and potential applications of AI across the digital ecosystem include:

Healthcare	Education
More accurate diagnoses and treatment; personalised medicine; improved medical decision-making; forecasting health risks and improving preventative responses; virtual agents to guide patients; remote patient monitoring and consultations	Automating teacher tasks; virtual teaching assistants; automating assessments; programming assignments; personalised or customised learning; students learning at their own pace; remote teaching and assessments; personalisation at scale

[2] When chapters do discuss "artificial general intelligence", it is highlighted and dealt with specifically

[3] Garry Kasparov, *Deep Thinking: Where machine intelligence ends and human creativity begins* (John Murray, 2017)

Public services	**Utilities**
Better forecasting; more efficient and targeted provision of public services	Optimising management and use of utility infrastructure; better predictions of demand and supply; condition-based maintenance, rather than scheduled maintenance; increasing capital productivity
Meteorology	**Climate change**
Analysis of weather patterns; predicting adverse weather-related events	More accurate climate models
Transport	
More efficient transport systems; as well as autonomous vehicles; making public and private transport safer	

b) Sustainable Development Goals

AI is expected to be a key enabler for countries to achieve the Sustainable Development Goals.

At the ITU's AI for Good Summit 2017, the significance, and implications, of AI for developing countries was discussed:

> *"Developing countries may have the most to gain from AI, but unless we are vigilant, they may also have the most to lose. In order to reap the benefits of AI, vast amounts of data are needed, which are only available through mass digitization – an area where developing countries lag far behind. There can be no mass digitization without universal and affordable access to broadband, which is central to ITU's mission. We need to avoid a deepening of the digital divide, so the benefits of AI can be distributed equitably"*[4].

This highlights the challenges for emerging countries around digitisation, and broadband access, to the successful application of AI technologies, which we discuss further in this module and in this AI Series.

A critical point was also made at the AI for Good Summit 2017 that:

> *"… it is vital that the needs of a diverse range of people, including the most vulnerable, guide the design and development of AI systems. Those who are furthest behind in terms of social and economic development are at the centre of the SDGs and need to be at the centre of design and application of technologies such as AI".*

We discuss the potential of AI, and some of these challenges, further in this module and elsewhere in this AI Series.

4. Status of AI development and availability around the world

AI development is currently mainly concentrated in large wealthy countries or regions (in particular, the United States, China and the European Union). AI development is also concentrated in sectors

[4] ITU, AI for Good Summit 2017 report: https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf

which are early adopters in digital technologies (the high-technology sector, telecommunications, financial services, etc)[5]. A key characteristic of each of these sectors is that industry participants have access to large volumes of structured data.

According to the McKinsey Global Institute:

> *"AI investment is growing fast, dominated by digital giants such as Google and Baidu. Globally, we estimate tech giants spent $20 billion to $30 billion on AI in 2016, with 90 percent of this spent on R&D and deployment, and 10 percent on AI acquisitions. VC and PE financing, grants, and seed investments also grew rapidly, albeit from a small base, to a combined total of $6 billion to $9 billion. Machine learning, as an enabling technology, received the largest share of both internal and external investment".*

Much of the AI investment today relates to machine learning (almost 60% of investment according to McKinsey), which is an enabling technology for other AI developments. Autonomous vehicles, for example, is a relatively small investment class currently, but experts predict it is likely to emerge quickly. Autonomous vehicles is a high public profile technology and will be a bellwether of AI and its acceptance, or resistance, by the public.

Adoption of AI in health and education is growing, from a low base. Both sectors face the challenge of having to build the trust of professionals in that field, the public and regulators.

India case study: In February 2018, India's finance minister Arun Jaitley informed Parliament, during the 2018-2019 budget speech, that Niti Aayog, the premier policy think-tank for the government, will oversee a National Programme on AI, focussing on research and development of AI and its application.

The Digital India programme, the government's initiative for the promotion of AI, machine learning, and other related fields, intends to emphasise promotion of AI in 2018 and has set up four committees to encourage research related to AI. These committees are focused on researching and working on development of AI, creating a data platform, skilling, re-skilling, research and development, legal regulatory, ethical and cyber security issues. Digital India's funding nearly doubled to US$477 million for 2018-2019.

5. The technologies used in AI

a) Machine learning

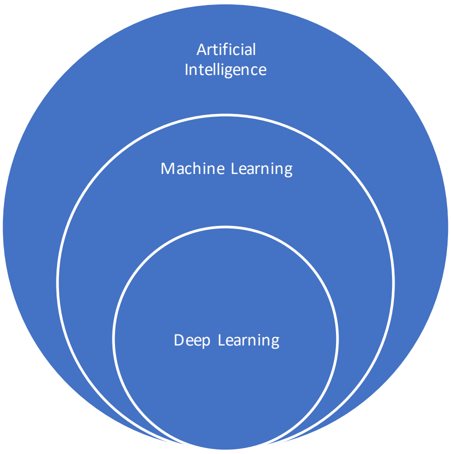
Machine learning is a subset of AI. It's what most people tend to think of when they imagine AI. Machine learning allows systems to learn directly from data, without being explicitly programmed. Structured and unstructured data provide raw material for algorithms, using training data to identify statistical rules and correlating inputs with successful outputs, learning to make predictions and recommendations. Machine learning algorithms tend to emphasise outcomes over processes. They use induction and decision-tree techniques, building context by analysing new data.

As noted in the module on AI and IoT in security aspects, "In a nutshell, machine learning is all about automatically learning a highly accurate predictive or classifier model, or finding unknown patterns in data, by leveraging learning algorithms and optimization techniques".

[5] McKinsey Global Institute

b) Deep learning

Deep learning is in turn a subset of machine learning that uses algorithms to gain more abstract insights from data. There are other forms of machine learning (such as search, symbolic and logical reasoning, etc), but deep learning is the most prominent. There have been some highly successful developments in machine learning and deep learning in recent times. Artificial neural networks are trained on enormous data sets powered by high performance computing.



c) Neural networks

Neural networks, with layers of "neurons", each receiving an input, detect patterns and then provide an input to the next layer of neurons. A neural network generally contains many layers and many neurons in each layer, with intricate webs of connections between the layers. The technology is very loosely inspired by how the human brain and nervous system works.

d) Computer vision and natural language processing

Computer vision and natural language processing are important AI technologies. Computer vision has accelerated with the developments in deep learning. Computer vision takes advantage of powerful graphics processing units (GPUs), which are used in machine learning, which can, e.g., quickly and accurately process images or video to identify objects and position.

Natural language processing enables AI systems to understand what is said or written and its meaning. The technology is now moving beyond responding to simple text enquiries into being able to engage in more complex interactions with people.

Computer vision is an important technology for healthcare and natural language processing for education.

e) Supervised and unsupervised learning

Supervised learning algorithms are trained on datasets that include labels estimated by a data scientist, indicating the importance of features within the problem. Backpropagation applies an algorithm which makes it possible for machines to predict an outcome based on input information provided. The algorithm runs many trials, learning from each trial by analysing the difference between the assigned expected outcome and the outcome reached. The algorithm then adapts its previous guess and attempts the process again. This is repeated until the algorithm has ran all its cycles (or epochs), resulting in the "actual" outcome based on the initial values given.

Unsupervised learning algorithms must determine the importance of features within the problem on its own, by analysing inherent patterns in the data. An amalgamation of these methods are semi-supervised or reinforcement learning algorithms.

The module on AI and IoT in security aspects includes a detailed description of the key technologies underlying the development in AI.

6. Investment and ICT infrastructure requirements

AI is supported by ICT infrastructure. This includes cloud-based computers with high processing powers, but also IoT networks of sensors and devices that can feed vast quantities of real-world, real-time data in to AI systems.

To support AI, ICT infrastructure will need to be flexible, very low latency, reliable, secure and adaptable to different use cases.

a) Communications infrastructure

AI will require "smarter" communications networks, which involve softwarisation, cloud infrastructure, virtualisation and more complex network structures.

a. Mobile telecommunications networks

In mobile telecommunications networks, the foundation communications infrastructure for the foreseeable future will be IMT-2020 or 5G, which is expected to be commercially widely available in the early 2020's. IMT 2020 will facilitate increasing "softwarisation" of the network – greater virtualisation and centralisation of operations (reducing cost, increasing flexibility in meeting customer and network requirements). These technologies benefit network providers by reducing their costs, but also AI users will benefit from the scalability and customisation of these technologies.

Some key IMT-2020 technologies:

- *Software defined networking (SDN)* – allows greater flexibility, agility and control in large networks; the foundation of many emerging network technologies[6] [7]
- *Network function virtualisation (NVF)* - allows operators to use commercial servers for base station hardware; decreases complexity of hardware needs
- *Network slicing* – allows operators to provide isolated sub-networks, each optimised for specific types of traffic characteristics

5G networks, with far greater capacity requirements, will require "densification" of the networks, with more base stations and access points, at both the macro and small cell layers.

[6] See the ITU's SDN portal here: https://www.itu.int/en/ITU-T/sdn/Pages/default.aspx

[7] See: Recommendation ITU-T Y.3150 "High level technical characteristics of network softwarization for IMT-2020".

5G is optimised for Internet of Things (**IoT**) capabilities, where an enormous range of devices will connect to the network. The World Economic Forum estimates there will be as many as 30 billion IoT devices in the next ten years[8].

Smart cities are a use case that goes beyond IoT, but IoT is integral to smart cities.

Although 5G will in time be the foundation communications infrastructure to support AI, networks operating 4G and possibly 3G can still provide a reliable infrastructure for some applications. Sensors, for example, that feed AI applications and are only required to communicate occasionally with small amounts of data may indeed operate over 3G or even 2G.

b. Fibre networks

Increasingly, fibre infrastructure will be necessary to support the more advanced mobile telecommunications networks. Fibre backhaul will be required to connect to the base stations and access points to provide low latency and high capacity.

Although fibre infrastructure is available in main centres of large developed countries, and also in central urban areas of many emerging countries, it remains a huge challenge for governments to facilitate the expansion of the range of fibre networks with the sort of density that will be required for advanced mobile telecommunications networks, but even more so outside of urban areas and into rural areas.

c. Investment and market structures

The investment requirements for new fibre-rich, high-speed mobile broadband networks to support the full realisation of an AI future will be considerable in most countries. In many countries, new duct or pole infrastructure will be required to push fibre deeper into the networks. Environmental and health concerns around the world may create real consenting obstacles for densified high-speed mobile networks.

This pressure may result in calls for single networks, at least at the passive layer, which will be shared by retail service providers and others providing IoT and AI applications. Whether existing market structures built around infrastructure competition will still be fit for purpose in this new era may be a valid question in some countries.

Governments and regulators will need to consider how this new infrastructure will be financed and how access will be provided. Co-investment models may be appropriate in some jurisdictions, or new infrastructure could be owned by non-traditional telecoms investors, such as infrastructure funds. Governments may be investors, supplemented by donor funds in some cases.

There are many issues for governments and regulators to consider to drive the deployment of new broadband infrastructure. AI is a use case for high speed broadband and there are many other applications for broadband networks. However, we emphasise the importance of high quality communications infrastructure as a key enabler of an AI future.

b) Cloud infrastructure

AI relies on robust cloud infrastructure to provide the computing power to run the algorithms and massive data sets that are required.

[8] https://www.accenture.com/t20170411T115809Z__w__/us-en/_acnmedia/Accenture/Conversion-Assets/WEF/PDF/Accenture-Telecommunications-Industry.pdf

Some key cloud technologies:

- *Infrastructure as a Service (IaaS)* – infrastructure elements are hosted by a third party, which may include hardware, software, storage, with associated services

- *Platform as a Service (PaaS)* – a service that allows users to develop, run and manage their own applications, using common infrastructure

- *Software as a Service (SaaS)* – a service that provides access to software over cloud infrastructure and platforms

This also requires physical data centre infrastructure to run these cloud applications. There are hundreds if not thousands of data centres around the world and they are essential for centralised cloud computing. They also consume a large amount of energy and so access to low cost and high quality electricity systems is important for their development. Indeed, as with all ICT technologies, a certain base level of electricity infrastructure will be required to realise AI's full potential.

7. Socio-economic impact of AI

AI will allow certain functions to be performed more accurately and efficiently than humans are capable of. The implications of this are wide ranging and will impact on socio-economic matters such as employment, training and the future of work. The same can be said for automation. Indeed, it is helpful to consider the socio-economic impact of both AI and automation, as they both are emerging as potent technologies and will both have wide ranging effects.

This is not a new issue. Society has been dealing with the impact of technology and mechanisation on jobs for hundreds of years. As with previous generations of technology, the growth of AI and automation is expected to adversely impact on employment in some areas, but also create new employment in other areas (e.g., data science fields).

All of this is uncertain. We may be able to anticipate jobs that are likely to be lost because of AI and automation, but we don't know when this is likely to occur. We expect it won't occur evenly around the world. Some countries, and some sectors, will be affected earlier than others. Some countries may be insulated from its effects for some time.

We also don't know what new jobs will be in demand in an AI future and whether there will be a net gain or net loss of employment.

Nevertheless, we can anticipate that many people around the world will eventually be affected to some degree by the impact of AI and automation. The broad implications of these effects may require considerations of development of social safety nets and ideas such as universal basic income.

Governments need to develop a sense for where and when the benefits and risks of AI will be experienced, how those benefits and risks are likely to be realised (broadly or narrowly) and where the opportunities are for broadly shared benefits.

a) Jobs that will be affected by AI

Over time, we can anticipate that the impact of AI and automation on employment may be profound. In one prominent 2013 study by Frey and Osborne[9], the authors estimate "*… around 47*

[9] https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

percent of total US employment is … at risk – i.e. jobs we expect could be automated relatively soon, perhaps over the next decade or two".

In emerging countries, the impact of AI on employment could be more significant, with greater proportions of low skilled workers performing manual or repetitive tasks. These jobs are potentially most at risk of being replaced by automation.

In a World Bank study[10], the authors found that two-thirds of all jobs are susceptible to automation in the developing world. However, the impact of AI and automation in emerging countries is likely to be cushioned for a period by slower technology adoption than in developed countries and lower wages. Lower wages in emerging countries may attract jobs that cannot be efficiently undertaken in developed countries impacted by AI and automation. On the other hand, the United Nations believes that the inevitable increased usage of robots in developed countries will erode the labour-cost advantage which emerging countries have enjoyed[11].

Jobs will increasingly require people to work collaboratively with AI, just as we do today with new technologies. AI and robotics will tackle manual or repetitive tasks, while humans will undertake more creative or strategic tasks, which complement the respective strengths of machines and humans.

b) Preparing people for the age of AI

It is clear that policy makers should begin the process of adapting their education and training systems to prepare their people for the age of AI.

Throughout formal education, there has been a primary focus on literacy and numeracy, which have been important skills for many jobs in today's workforce. However, recent studies show that current AI techniques are close to performing literacy and numeracy tasks at or above the proficiency of 89% of adults in OECD countries (Elliott, 2017)[12].

This suggests that policy makers should consider preparing students beyond literacy and numeracy to include training and skills in such areas as problem solving, data and statistical literacy, computational thinking and digital technology.

If the future of work is likely to involve humans working in complementary areas alongside AI, then education and training should prepare people in those complementary areas. This training will be required from an early age, through primary and secondary school and on to tertiary education.

Just as importantly, it will also be necessary to consider the needs of those already in the workforce and those of working age, who will require training in new skills. Continuous learning itself will be a core skill going forward. People need to be prepared and adaptable to meet the needs of a changing work environment.

SkillsFuture initiative case study: In January 2016, the Singapore government created the SkillsFuture initiative. This initiative provides guidance on expected areas of employment growth and training subsidies. To enable Singaporeans to take time out of full employment, a credit for workforce

[10] https://openknowledge.worldbank.org/handle/10986/23347

[11] http://unctad.org/en/PublicationsLibrary/presspb2016d6_en.pdf

[12] See http://www.keepeek.com/Digital-Asset-Management/oecd/education/computers-and-the-future-of-skill-demand_9789264284395-en#page90 pg 88-90, 96.

retraining has been given to everyone aged over 25, with further training subsidies available to those over 40 (The Economist).

If countries aspire to be AI hubs, then serious training in AI development is required at graduate and post-graduate level. Considerable investment in tertiary level capabilities and facilities will be necessary.

c) Impact on taxation revenue

On its face, if a robot or AI process displaces a human for the same job, this will not necessarily impact on the income tax revenue from that person, so long as that person is able to find another job for a similar income. However, if another job for a similar income is not available, or not available immediately to replace the displaced job, then income tax revenue for the government will diminish.

This has led policy-makers in some countries to consider how to manage any shortfall in income tax revenue that may arise as a result of robots or AI processes replacing jobs.

Robot tax case study: As part of EU-wide legislative talks on regulation of automation, a robot tax was proposed, and rejected, in February 2017. This tax would have been levied on robot owners, to pay for retraining of workers who lost their job. Robot tax is a colourful description of a tax on automation. EU Commissioner Andus Ansip described such a tax as a "tax on progress", which would result in Europe falling behind others in AI development.

On the other hand, South Korea has begun limiting tax incentives for investments in automated machines. South Korea's "robot tax" involves reducing currently available tax deductions for automation investments. While not directly taxing the employment disruption caused by robots, it is intended to provide comparable results. The reform would reduce the current deductions of three to seven percent for automation investment by up to two percent.

In 2016, the United Nations Conference on Trade and Development (UNCTAD) remarked that:

"Clearly, without the introduction of a major tax on robots as capital equipment, robot-based manufacturing cannot boost the fiscal revenues needed to finance both social transfers, to support workers made redundant by robots, and minimum wages, to stem a decline in the living standards of low-skilled and medium-skilled workers."[13]

d) Safety nets

Universal Basic Income (**UBI**) has been proposed by some experts as a solution to address the social consequences of the expected displacement of jobs by automation (and AI). Under UBI, all citizens would receive a reasonable amount of money to ensure at least a minimum standard of living. Top economists, such as the chief economic advisor to the Government of India, Arvind Subramanian, and economics Nobel prize winner Sir Chris Pissarides, among others, have shown their support for UBI.

In emerging countries such as India, Subramanian argues that a safety net from UBI would support people impacted by poverty due to droughts, declining agricultural opportunities etc. In a similar light, a decline in manual or repetitive tasks due to automation may also require a safety net to

[13] http://unctad.org/en/PublicationsLibrary/presspb2016d6_en.pdf p.3

catch those whose work becomes redundant. Sir Chris Pissarides advocates for UBI as a solution to inequality, which may be expected to rise because of automation.

The idea of a universal basic income has existed since the industrial revolution. In 1849, John Stuart Mill famously proposed that a "certain minimum" should be assigned by the government for the subsistence of every member of the community, whether capable or not of labour.

e) Other policy proposals

The ITU paper on the social and economic impact of digital transformation on the economy for GSR17[14] examined many of these issues in depth. A number of proposals are put forward for policies aimed at promoting innovation in advanced technologies while mitigating workforce disruption in developed economies, including:

- *"Increase public expenditures in education to increase the skills (including digital skills) acquired through formal training;*
- *Implement labor policies focused on workers being able to retain their current jobs or move to new areas of demand (job placement services, special labor market programs, apprenticeship programs);*
- *Put in place subsidies to lessen job disruption of low-skilled workers (tuition-free education, temporary cut in payroll taxes, basic income guarantees);*
- *Implement policies aimed at increasing geographic mobility (reduction of relocation costs, subsidized housing; and*
- *Promote demand for skilled workers by accelerating the rate of innovation in areas likely to be affected by job disruption effects*".

8. Significance of a strong foundation in data

AI requires a strong foundation in data. Access to data is needed to train AI systems, to allow them to identify patterns, which in turn enables those systems to make predictions and recommendations. In comparison to data, computing power has almost become a commodity and so perhaps less important to the development of AI as access to data.

a) Open data and open standards

Open data and open standards for public data are likely to be an important enabler of AI in many countries. Open data improves the quality of public services, through learnings from the data made available and providing new insights, which will also be valuable for AI. Open standards will assist AI systems in making sense of the complexity of data.

Governments can promote open standards to build a robust data ecosystem in their country, particularly for public data, making systems and data interoperable. These include common standards for metadata, which will allow the provenance of data to be traced as data is used and reused for different purposes[15].

[14] https://www.itu.int/en/ITU-D/Conferences/GSR/Documents/GSR2017/Soc_Eco_impact_Digital_transformation_finalGSR.pdf

[15] For further discussion about open data in developing countries, see Verhulst and Young, "Open Data in Developing Economies: towards building an evidence base on what works and how", July 2017

This may be one of the most significant steps that Governments can take to prepare themselves for the age of AI. A quality government data environment, with open standards, will be foundational to maximise the achievement of the Sustainable Development Goals through the use of AI. Conversely, a poor government data environment, with inconsistent and incoherent standards, will impede the potential for AI.

Governments can play an important role in developing and adopting effective anonymisation or de-identification techniques which can be implemented through these open standards, providing an appropriate balance between re-identification risks and the public benefit in using this information. We discuss anonymisation or de-identification techniques further below.

The Open Data Charter is a collaborative effort between data experts and over 70 governments operating with the objective of opening up public data. The Open Data Barometer (ODB) is an initiative of the World Wide Web Foundation[16]. In its most recent survey (2016), they found that 79 of 115 countries studied had operational open data initiatives.

Open data charter case studies: The principles of the Open Data Charter are in summary: open by default, timely and comprehensive, accessible and usable, comparable and interoperable, for improved governance and citizen engagement and for inclusive development and innovation[17].

The G8 have an open data charter[18]. Under the open data charter, all government data is expected to be published openly by default, alongside principles to increase the quality, quantity and re-use of the data that is published.

b) *International data exchange*

Governments can promote the international and regional exchange of data and global collaborative efforts. Medical research is an area where there are expected to be particular benefits from international exchange of appropriately anonymised clinical data.

The European Parliament is currently discussing the final stages of a bill to allow free flow of non-personal data between EU countries. A draft bill allowing nearly unrestricted flow of non-personal data (aside from where there may be concerns for public security) was passed in the Council of the EU on 20 December 2017. The legislative work is expected to be finished by June 2018.

The OECD Privacy Framework encourages transborder flows of personal data between countries where safeguards and effective enforcement exists consistent with the OECD guidelines. Any restrictions which are made to transborder data flows should be proportionate to the risks presented. The OECD framework intends to avoid the creation of unjustified obstacles to economic and social development. They use the example of excessive protection of personal data, exceeding the requirements necessary. The OECD recognises the benefits towards efficiency gains and convenience of increased transborder flows of personal data, however they are concerned with respect to the protection and enforcement of privacy.

c) *Data for public good*

The public good applications of AI are likely to be considerable for all countries. However, the public good applications in emerging countries may be even more significant than in developed countries,

[16] https://opendatabarometer.org

[17] See here for more details: https://opendatacharter.net/principles/

[18] https://opendatacharter.net/resource/g8-open-data-charter/

where commercial applications of AI will likely, at least initially, receive most of the investment and attention of the major players in the field.

However, much of the data that will be valuable from a public good perspective will be either personal data or commercial and proprietary data. This creates a tension between public good on the one hand and personal privacy and commercial strategic value on the other hand.

Personal data, which is provided and may only be used for certain purposes, may have substantial public benefits if it was accessible, for example for research purposes or for providing improved public services on an anonymised or de-identified basis using AI technologies.

Healthcare data case study: More health-related data is being collected than ever before, including by mobile apps, Fitbit, etc. Access to anonymised patient data may be highly beneficial for medical professionals and researchers.

Governments can develop rules for who can access this sensitive data, what it can be used for and how it is stored, protected from cyber risks and how it should be anonymised or de-identified.

Consideration will need to be given to providing incentives and mechanisms to share health data for these public benefit purposes. The same likely applies for education data (information about student performance, etc).

A different set of issues arises with commercial or proprietary data. Some commercial or proprietary data will be derived from personal data provided in return for digital services. Other commercial or proprietary data, with potential public good applications, will be developed for the purposes of providing commercial services (such as mapping data collected by various companies).

There is likely to be strategic value in that data which weighs against use outside of the business concerned. Costs and legal risk will also be a consideration. For example, there may be costs associated with the anonymisation or de-identification of any personal information, which would not have needed to be incurred if that information was not released. There may be legal risks for the business, particularly around re-identification or third-party confidentiality rights.

Nevertheless, public benefit may be realised in accessing appropriately protected, or aggregated, commercial or proprietary data for new AI based public services. Questions of incentives, and safeguards (e.g., protections from liability), for holders of proprietary data to share that data for public services need to be considered.

These are critical issues for governments preparing for the AI age. Where the public benefit from AI technologies and its reliance on high quality data is growing, these tensions between personal privacy and commercial considerations need to be deliberated by governments, and the public and the major holders of personal data. While these issues are present in the current era, they are likely to become more prominent in the AI age.

9. Ethical, legal and regulatory issues

a) Personal data

Data protection laws protect personal data, which is information about an individual. Access to personal data, and its protection, will be critical to the future evolution of AI. AI will not succeed if people lose confidence in the ability of AI to protect their personal data.

The European Union has a comprehensive regulatory framework for the protection of personal data. The approach in the European Union is to treat personal data as information about an identified or identifiable individual. This has been broadly followed in the OECD guidelines and the Privacy Framework of the Asia-Pacific Economic Cooperation[19], among others. By contrast, for example, the United States has pursued more of a sector-specific approach for personal data.

Privacy laws do not apply to non-personal data (information that does not relate to an identified or identifiable person), or to data where the person's identity has been sufficiently anonymised or de-identified[20].

Existing data protection laws were usually established at a time of limited collection and limited usage of personal data. What has changed in the intervening period is that more data is now collected about individuals, in new ways, at far greater scale. Personal data is used (and re-used) for a much wider range of purposes than ever before, often far beyond the original purpose. An increasing number of entities are involved in the collection and in the processing of data, often without explicit knowledge of the individual. There is very limited public awareness of these activities.

Collection case study: Data is being collected through sensors, social networks, vehicles, etc. It is captured as a by-product of interaction with devices, services, etc. Data is collected directly, e.g. through use of device, and indirectly, e.g., through sensors, Wi-Fi hotspots, or just being in places, including in the home. Data may still be private, even if it is captured in public places.

Big data and analytics allows for greater insights to be obtained from collected data, beyond what had been the original purpose of collection. Sometimes those insights may be apparent much later than the time of collection.

The EU approach, and in many other countries, is to protect information that relates to an "identified" individual, but also information that relates to a person that is "identifiable" (that is, they could be identified). In the EU[21], there is a test of reasonable likelihood of identification, but the test is dynamic, in that information may not be "identifiable" at the time of collection, but it may become identifiable as a reasonable likelihood through the progress of technology change.

De-identification case study: The anonymisation or de-identification of data can remove immediate privacy concerns. However, developments in advanced analytics over recent years has meant that

[19] OECD guidelines define personal data as "any information relating to an identified or identifiable individual (data subject)". The Privacy Framework of the Asia-Pacific Economic Cooperation 2004 defines PII as "any information about an identified or identifiable individual."

[20] e.g., in Recital 26 of the GDPR, personal data that is "rendered anonymous in such a manner that the data subject is not or no longer identifiable" is excluded

[21] Recital 29, GDPR

personal data may increasingly be inferred from de-identified data. Professor Paul Ohm has highlighted that re-identification risks, arising out of modern analytics technologies, render data increasingly identifiable[22]:

"Easy reidentification represents a sea change not only in technology but in our understanding of privacy. It undermines decades of assumptions about robust anonymization, assumptions that have charted the course for business relationships, individual choices, and government regulations."

These developments, and these concerns, have in turn propelled the search for advanced new technologies that can substantially reduce re-identification risks. These new technologies include differential privacy and homomorphic encryption.

Differential privacy is a method of data collection which applies random noise to the dataset on collection, where an individual's true information is distorted and will not be recognisable in the dataset.

Homomorphic encryption allows for the computational use of encrypted data, without knowledge of the true (decrypted) data. By never needing to decrypt the data, the privacy of users is uncompromised during the computational process. Although homomorphic encryption has existed as an idea for nearly 40 years, full homomorphic encryption is not expected to be usable for several decades due to the intensive computing power required.

The tensions between data protection and the realisation of IoT will create new grey areas with space to circumvent legislative boundaries. These are addressed in the module on AI and IoT in security aspects. The module also examines how data protection may be threatened in an IoT environment and further discusses AI-based privacy enhancing techniques and mechanisms in greater detail.

Some experts have called into question whether the traditional data protection law models are suitable in the AI and big data age, where information is increasingly identifiable. Some academics argue that the distinction between "identified" and "identifiable" information is becoming meaningless[23]:

Koop argues that[24]:

> *"Current data protection law … might be considerably more productive if, instead of trying fitfully to establish where the border lies between personal and non-personal data, we would allow for categories of data that have certain effects on people when they are processed, regardless of whether or not they relate to identifiable individuals."*[25]

This brings up the issue of context. Attitudes of the public to data protection (the benefits and the risks) depend on the context. For example, people often disclose personal information to receive digital services and feel no strong need to protect themselves. But that same information, used in another context, or when combined with other data, may be very concerning for the individual. Therefore, another possible approach to data protection law is to be more context-specific and

[22] Ohm, P. (2010) "Broken Promises of Privacy," 57 UCLA L. REV. 1701 (2010)

[23] For example, see Schwartz, P. and Solove, D. (2011) "The PII Problem: Privacy and a New Concept of Personally Identifiable Information" 86 N.Y.U. L. Rev. 1814

[24] BJ Koops, 'The trouble with European data protection law' International Data Privacy Law, Volume 4, Issue 4, 1 November 2014

[25] http://www.isaca.org/Groups/Professional-English/privacy-data-protection/GroupDocuments/2014-08-24%20%20The%20Trouble%20with%20European%20Data%20Protection%20Law.pdf

provide different levels of protection of data depending on the sensitivity and proposed use of the data.

These issues extend far beyond AI, but have a relevance in the AI context due to the use of personal data, or anonymised or de-identified data, in AI processes.

b) Notice and consent

While most countries, in most circumstances, require notification of the individual of the purpose of collection, use and disclosure of their personal data, there are different approaches taken in relation to obtaining consent. Some jurisdictions (e.g., the EU countries) adopt a notice and consent approach, while others are notification based, with consent required in limited circumstances.

People have often "consented" to the collection of their data, e.g., to receive the benefit of a digital service. Individuals are presented with detailed terms and conditions, which few are likely to have read or understood before accepting. The purpose of collection is often broadly described and individuals in many cases have little knowledge of information that is collected about them and what it's used for.

And things can change. The use of a service may vary from the time of initial notification and consent. With AI and big data, new insights can be gained from old data, including through combining data, and questions arise whether the original notification or consent was sufficient.

The solution to these difficulties is likely to involve greater transparency and public awareness of the benefits, and the risks, of intensive data usage.

Some experts have argued that data protection law should recognise the trade-offs involved when public good may be derived from personal data. Tene and Polonetsky suggest that: "*Where prospective data uses are highly beneficial and privacy risks minimal, the legitimacy of processing should be assumed even if individuals decline (or are not asked) to consent*".

Case study Singapore data consultation: The Personal Data Protection Commission of Singapore (PDPC) recently conducted a public consultation on approaches to managing personal data in the digital economy[26]. Singapore's data protection legislation primarily provides for consent as the basis for collection, use and disclosure of personal data. The PDPC noted that, in today's analytics-driven world, it may not always be possible to anticipate the purposes for use and disclosure at the outset. Also, it may not be possible always to obtain consent from individuals when their data is collected or attempt to identify the individuals to seek their consent for every new purpose.

The PDPC proposed an approach where notifying individuals of the purpose can be appropriate, where there is no foreseeable adverse impact on the individuals arising out of the collection, etc, of the personal data. The PDPC also proposed that there be a "legitimate interest" in collection, etc, of personal data without consent. This is a more limited ground, intending to apply in situations such as prevention of fraud.

[26] https://www.pdpc.gov.sg/Legislation-and-Guidelines/Public-Consultations#ACTR1

c) Bias or fairness issues

AI systems should be developed to ensure the equal and fair treatment of people that are affected by decisions made by that system.

This issue can arise as a result of bias inherent in the data on which the algorithms are trained (which may be derived from human biases at the time of collection). For example, AI systems may have been trained on limited data with under-representation of certain demographics, or systems which are selectively used for marginalised populations.

As suggested in the module on AI, Ethics and Society, "the risks for bias in AI is probably greater due to the qualities of its datasets than for any "hand coded" biases of its algorithms".

Heat map case study: Jessica Saunders et al.,[27] illustrate the results of bias through police "heat maps", which attempt to predict where best to patrol. Through increased patrolling, more criminals are caught in those areas, leading to the system being trained to increase patrolling further. Saunders et al., discovered that the use of "heat maps" by police has led to disproportionate harassment of African Americans.

The lack of diversity of those involved in AI research (a "sea of dudes" (Mitchell) and a "white guy problem" (Crawford, 2016)) is another issue. This lack of diversity may in turn create certain types of biases in AI systems, created through the lens of white male AI developers.

At a practical and technical level, it is very difficult for developers to ensure data or algorithms are free from bias.

But AI may also be the solution to this problem. AI systems are likely to produce more impartial results than humans as they are not susceptible to conscious or unconscious biases if they are designed properly. They can be used to detect and eliminate biases.

This issue was recognised in the Korea Mid- to Long-Term Master Plan in Preparation for the Intelligent Information Society[28]: "*As the massive quantities of data involved and high complexity of AI algorithms will make it nearly impossible for humans to rid these systems of biases once they begin operating and evolving, policymakers may well need to develop and establish refined methods for applying and testing ethical standards for their development at every stage (e.g., requirements for testing the fairness and reliability of data, enforcing the fiduciary duty of developers, preventing reverse choices, etc.)*"

d) Interpretation and transparency

AI systems today are rarely set up to be transparent and provide reasons for a decision that it makes. As a result, AI systems can be difficult to interpret. However, in certain circumstances (where the outputs are consequential for people, e.g., granting a mortgage, insuring a home, etc.), reasons may need to be provided for an AI decision.

[27] Jessica Saunders et al., Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot, 12 J. EXPERIMENTAL CRIMINOLOGY, 347, 350-51 (2016).

[28] http://www.msip.go.kr/dynamic/file/afieldfile/msse56/1352869/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf

In part, this relates to the bias and fairness issues discussed above. Where decisions are consequential, and where bias or fairness (or simply errors) can be a concern, then issues of interpretation and transparency become increasingly important.

Human decisions, of course, are also not necessarily interpretable or transparent. AI systems are more easily audited than humans.

European law makers have introduced a "right to explanation" in the GDPR, which requires "meaningful information about the logic involved"[29]. Questions arise over what an explanation is and whether disclosure of the program is sufficient. There is also a right for the person concerned "… not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her"[30].

Black box case study: There are inherent difficulties with transparency and interpretation of "black box" deep learning or neural networks. These systems are high performance, but they are also opaque and less transparent. Concerns have been raised by AI engineers over whether transparency requirements may compromise performance.

Specialised tests may be required that look for bad outcomes. Artificial inputs could be used to test for unusual situations that can produce unexpected outcomes. Some details of the system's design may be published, enabling analysis, without revealing proprietary or private information.

Despite the difficulties in "black box" systems, transparency is likely to be essential in building public trust in AI, at least in those circumstances where outputs are consequential for people. Failure to build and maintain public trust will likely lead to underuse of this important technology.

Ethical guidelines will be required. In some consequential areas, people will expect the right to understand the decision-making process, etc. In certain highly sensitive areas, there may be a need to limit development of AI to areas where human explanation is possible.

The AI, Ethics and Society module explains in greater depth the issues involved in interpretability and transparency in AI systems. Dr Best refers to the Statement on Algorithmic Transparency and Accountability released by the ACM US Public Policy Council (2017), that advises transparency of data used to train AI systems, as well as explainability of their decisions. The statement also suggests auditing systems in case of harm, redressing groups adversely affected by algorithms, and holding accountable the entity producing the algorithm.

e) *Accountability and liability*

What happens if AI goes wrong? Accidents and even crimes can happen due to AI decision making.

AI is in many respects no different to other technologies. The designers and manufacturers of AI systems, and the users of those systems, are potentially accountable and liable for how those systems operate, depending on the civil or criminal law of the country concerned.

[29] Articles 13 and 14, GDPR

[30] Article 22, GDPR

In most countries, a victim of harm can sue the wrongdoer under civil law rules for negligence, failure of statutory duty, etc. But who is liable where it is the AI system that is doing the wrong?[31]

While there is discussion in academic circles of the concept of legal personhood for AI systems, it is currently too abstract to be given serious practical consideration. So, the AI system itself, lacking legal personhood, would not be liable for the harm that it causes.

We can imagine the owner or user of the AI system to be potentially liable, notwithstanding that they did not cause the harm. Owners and users of technology commonly have legal responsibility, when the harm is caused by technology that they control (say, industrial machinery). This may be under tort laws or strict liability laws.

What's different with AI is that, being to some degree "intelligent", it operates autonomously and potentially in ways that are not expected by the owner or user of the AI system. To the extent that some fault is required on the part of the owner or user, then it may be difficult to prove this with autonomous systems.

Another option for a victim of harm is legal action against the manufacturer of the AI system, under product liability laws in some countries (usually without having to prove fault by the manufacturer), or under civil law where some fault would have to be demonstrated.

Governments need to consider whether liability for AI systems should be based on fault (like negligence) or strict liability (where no fault needs to be shown). Should AI systems be treated like domestic animals, unpredictable, but where public policy approaches to risk allocation make the owner of the animal liable for its actions? The answer may be different for different types of systems. And if the owner or user of an AI system is strictly liable, then they would need to claim against the manufacturer when the issue arose out of a defect, which may not be straightforward.

This also gives rise to the question of whether compulsory insurance should be acquired by an owner or user of AI, where they are strictly liable, such as occurs with vehicles in a number of countries. Compulsory insurance means that the victim can claim against the insurer.

Compulsory car insurance case study: The UK government has proposed a system where compulsory car insurance will be required to provide cover for motorists when they hand over control to an autonomous vehicle. Motorists, or their insurers, will then rely on existing rules of product liability and negligence to ascertain who's responsible.

In the criminal law domain, there may be a question over whether a person intends to commit a crime when it is committed by an AI-enabled machine that the person owns or uses.

f) Appropriate standards

If regulation is to be applied to AI, what standards should algorithms be required to meet? This is a new and evolving area and policy makers in some countries have begun to consider regulation in the context of autonomous vehicles.

[31] For further discussion on these issues, see European Commission staff working document "Liability for emerging digital technologies" SWD(2018) 137 24 April 2018; also, Petit, "Law and Regulation of Artificial Intelligence and Robots: Conceptual Framework and Normative Implications" 9 March 2017 and the European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))

Autonomous vehicles case study: The Germany Ethics Commission published a report on automated driving guidelines for the programming of automated driving systems in August 2017. The report consisted of 20 proposals, such as that automated driving is an ethical imperative if the systems cause fewer accidents than human drivers, in every driving situation it must be clearly regulated and apparent who is responsible for the driving task – the human or the computer.[32]

During a recent hearing before the United States Congress subcommittee on Digital Commerce & Consumer Protection in the United States[33], the chairman of the committee proposed that AI, such as autonomous vehicles, should be implemented under the condition that they are safer drivers than humans.

Various questions arise when considering the extent of safety required before acceptance of an autonomous system being "safer than humans". For example, is an autonomous vehicle expected to have fewer accidents on average than a human, or is it expected to outperform a human with access to the best safety features currently available? How much safer than humans do we expect an autonomous vehicle to be before they are accepted by policymakers? Further, who undertakes the certification process, testing the safety levels before public implementation, and to what extent?

In many respects, the high-profile area of autonomous vehicles will be the bellwether for these sorts of issues going forward.

g) Verification and validation

For critical systems, AI companies will be expected, or required, to be able to verify whether the technology is operating as intended under actual operating conditions, with no unwanted or unpredictable behaviours. This will require manufacturers to prove, test, measure and evaluate systems before they are deployed.

However, AI machine behaviour can change as algorithms evolve. This creates complications when it comes to verification. How long would a verification be expected to be effective for, before needing to be re-verified? Traditional software verification may not be adequate. In safety critical systems and infrastructure, like planes and bridges, there are robust and accepted processes for addressing verification and validation to ensure safety and reliability. Manufacturers of AI systems will need to address how to manage the risk and building a safety case for the technology.

h) Security threats

AI systems will give rise to cyber-security threats. Hackers will look to access AI machines or datasets used by machines or IoT sensor networks in ways that may negatively impact on AI behaviour.

While AI presents another attack vector for cyber-criminals, AI can also be used to improve cybersecurity by anticipating attacks, identifying vulnerabilities and taking steps to prevent attacks.

The main subject matter of the module on AI and IoT in security aspects is on cyber-risks in IoT environments. It describes in detail the overall features and technical issues that arise in these

[32] See https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile

[33] Self-Driving Vehicle Legislation: Hearing Before the Subcomm. on Digital Commerce & Consumer Prot. of the H. Comm. on Energy & Commerce, 115th Cong. (2017) (opening statement of Representative Greg Walden, Chairman, Subcommittee on Digital Commerce and Consumer Protection).

environments. A framework for the adoption of AI and how it can be used to enforce the security of IoT devices and networks is discussed.

The point is also made that "In security, [available] of big data means AI techniques can be exploited to analyse and recognize patterns of security vulnerabilities to prevent such attacks. Thus, the ability of IoT based platform to learn from data to analyse, identify and mitigate security threats is an important feature that every IoT system should incorporate".

i) *Market structure issues*

A relatively small number of firms in the private sector are currently at the forefront of pioneering AI development and they are deepening their expertise.

Research facilities within private sector organisations are moving towards becoming larger than universities or public facilities in AI, attracting leading practitioners in the AI industry. These firms also possess enormous troves of data, gained as a result of the digital services that they provide. It is difficult for smaller firms to compete in the market, given the concentration of in-depth analysis within these companies and their access to massive data sets.

There may be questions whether this leads to market power, including consideration of issues over barriers to entry in AI-related markets. For example, the data, although deep, may not be unique to that firm, which would reduce barriers to entry. These issues may arise in a mergers and acquisitions context, where one of these firms seeks to acquire another firm with AI capabilities.

More broadly, stresses are likely to emerge from a situation where the key inputs to this important technology (data, algorithms, know-how and IP) are held by the private sector, often outside the jurisdiction concerned, but where the public good benefits for a country are so great.

We expect this is an area where ICT regulators, and competition authorities, will need to examine closely in years to come.

10. Institutional framework and cross-sectoral and interdisciplinary approaches

a) Establishment of an oversight body

In most countries, there is a case for a government body or committee to be responsible for oversight over AI activities. It would not be premature to create such a body now.

This government body or committee may be newly established or it may be an existing body or committee, or indeed an existing regulator or government department, that perhaps has oversight over emerging technologies and their implications for policy-making. It may include people from outside of government, including academics, people from industry, consumer representatives and so on.

However it is constituted, an oversight body would be charged with providing advice to government more broadly. Its tasks may include:

- promoting public knowledge and meaningful public dialogue about AI and its benefits;

- research and analysis of regulatory and policy issues, as well as future technological developments;
- providing support for, and coordination with, sector-specific regulators;
- establishing standards, codes, ethical guidelines reflecting community values; and
- coordinating with other similar bodies internationally.

We see this oversight body having recommendatory powers, rather than enforcement powers.

International case studies: The Advisory Board on Artificial Intelligence and Human Society[34] was established in May 2016 under the Japanese Minister of State for Science and Technology Policy to advance research and development and use of AI technologies.

The French Digital Council[35] was established as an independent advisory commission that issues independent opinions and recommendations on questions relating to the impact of digital technologies on the economy and society and consults on new legislation or draft regulation.

Similarly, the UK Parliament has recommended a standing Commission on Artificial Intelligence.

Governments generally will need to up-skill in AI, to understand its policy implications. AI technical and policy capability should in due course be spread throughout government, providing more diverse perspectives on AI technology within the public sector.

The authors of the module on AI governance, discuss the importance of reducing information asymmetries in government when it comes to AI. As well as building internal capacity within government, they propose various ideas for government to interact with experts in the private sector, including through "tours of duty" and positions that operate outside of traditional bureaucratic structures. They suggest establishing ongoing interfaces with experts, that can supplement or replace the need to hire experts.

b) Sector-specific policy

Sector-specific regulators are likely to lead policy developments in their respective areas. For example, issues around enabling infrastructure would appropriately be dealt with by ICT regulators and ministries, issues around transport by transport regulators, medical applications by health authorities, financial markets by financial regulators, consumer protection by consumer protection authorities, etc.

c) Cross-sectoral policy

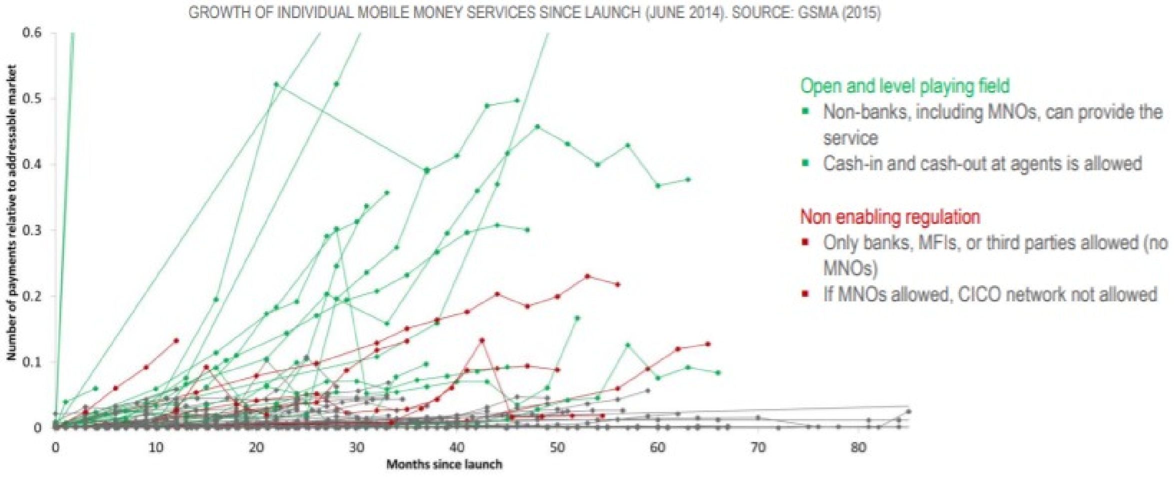
However, we can also envisage that a new technology, such as AI, will require increasing cross-sectoral approaches, which will require collaboration between different sectoral regulators. The ITU has emphasised the benefits of collaborative "G5" regulation, with the need to define the

[34] http://www8.cao.go.jp/cstp/tyousakai/ai/summary/aisociety_en.pdf
[35] https://cnnumerique.fr/en/french-digital-council/

foundation, platforms and mechanisms for working with other sector regulators to help achieve the Sustainable Development Goals[36]. This will also be pertinent in the context of AI policy.

Models for addressing cross-sectoral issues case study: Although not an example directly related to AI, the field of mobile money has required cross-sectoral regulatory approaches, in this context through the financial sector regulators alongside ICT regulators. The financial sector regulators tend to focus on increasing competition and efficiency, while the ICT regulators tend to focus on providing broad policy guidance on data protection, consumer protection etc.

Mobile money is also an example where, by implementing enabling regulation, growth and market penetration increased much faster relative to non-enabling regulation.



In some areas, governments may promote the sharing of incident and safety data related to AI among different sectoral regulators, such as what occurs with civil aviation with incident or near miss data.

d) Multistakeholder governance generally

One of the issues discussed in the module on AI governance is building effective multistakeholder governance groups. They propose a set of principles to guide the establishment of these groups and a range of tools that policy makers and regulators can deploy to engage with diverse stakeholders in advancing AI governance.

e) Data protection regulation

Because of the importance of data, and personal data, to the emergence of AI, the regulator with responsible for maintaining data protection laws will play a prominent role.

This may be an area of responsibility for the ICT regulator, or the privacy or data protection regulator (if a general data protection regulator has been established).

[36] See, for example: https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Documents/Events/2016/Regulatory%20Conference/ITU_RegulatoryTrends%20Sept%202016_J_Ponder.pdf and https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Documents/Events/2017/Regulatory%20Conference/Session%202%20Rosheen%20Collaborative%20Regulation_MontenegroITU.pdf

Whichever authority has responsibility for data protection, it is clear that the privacy implications of AI will be a critical area of focus for that regulator going forward. They will need the appropriate resource and powers to undertake this role.

f) Exploratory regulatory approaches

While we consider it is premature to implement specific AI regulation, we see merit in beginning to put in place structures and methodologies for exploring the potential regulatory implications of AI. These may include regulatory "sandboxes".

Sandboxes case study: Regulatory sandboxes allow for the piloting of new AI technologies in safe environments. The objective would be to promote innovative investment in AI for local application, starting with a contained, low risk, rule that permits something that would otherwise have been limited or prevented by regulation. This allows developers and regulators to observe, experiment, test and adapt further from there.

In the AI governance module, it is suggested that "*policymakers and regulators can create spaces that allow them to experiment in an iterative fashion with policies and regulatory approaches, that still allow for the development of new AI technologies, while still advancing core values of public safety, privacy, consumer protection, and due process*".

11. A roadmap for regulators

There will clearly be a need for policies and regulation that promote and facilitate the use of AI technologies, while at the same time addressing the potential challenges that these technologies present. These challenges may be different between developed and emerging countries.

a) Risk of over-regulation in growth phase

As discussed earlier, there is no clear definition of AI. AI developments are likely to occur gradually and incrementally, but they may experience a rapid acceleration (S-curve model).

There is a risk of over-regulation in the incremental growth phase that we are currently in. Overarching regulation appears to be inappropriate right now. Indeed, existing regulatory frameworks may be fit for purpose or may require relatively minor change. There will be different considerations in different contexts. Policy makers should consider how AI can reduce risks, as well as the risks that it creates.

b) Public awareness and trust

AI has received considerable media attention, but much of it has been superficial. There is little public awareness of the problems that could be solved by AI, with more public attention on extreme situations where AI might go wrong (e.g., accidents caused by autonomous vehicles).

This is an area where governments play an important role in helping society to prepare and adapt to AI. Governments can help to develop public trust and understanding in AI technologies and what the implications of these technologies will be for people.

This stage is critical and will be one of the first things that governments should be doing. If there is a lack of trust and understanding among the public, or if there is excessive fear of the consequences of AI (e.g., in employment), the potential benefits of AI may well not be realised. People may be reluctant to allow their data to be used in the development of these systems and may not be prepared to use them or allow them to be used.

c) Addressing the digital divide

The module on AI governance also highlights supporting local ecosystems of entrepreneurship and start-ups, as well as supporting capacity development at universities. It considers government programmes to facilitate the growth of entrepreneurial ecosystems, technology business incubators and other methods.

d) An AI national plan

Governments should consider developing an AI national plan. This would be a document that outlines the key strategies for preparing the country for AI. It should address the opportunities and risks, many of which are outlined in this AI Series.

For major economies, with research and investment ambitions, the AI national plan will be a comprehensive document.

AI national plan case studies: The United States[37] and China[38] have both produced substantial national AI plans, with India planning to release their own shortly. These plans focus on the research needs, regulatory requirements, data sharing and preparing an AI savvy workforce.

Emerging countries may have different objectives, but planning at a national level is still necessary. While it may be premature to regulate for AI, it is not too early to lay the groundwork for the emergence of AI. As discussed elsewhere in this introductory module, there are some key things that all governments can do now in anticipation of AI, including:

- Beginning a public dialogue to raise awareness of AI as a technology, of its benefits and potential consequences and how the government is preparing for it;
- Developing a quality government data environment, with open standards;
- Engaging with businesses operating in the country that are investing in AI internationally, and that may hold large amounts of personal data, and discussing the development of AI applications with public good purposes and that may fulfil Sustainable Development Goals objectives;
- Considering potential infrastructure roadblocks that could limit the potential for AI (e.g., access to 5G spectrum and deployment of fibre infrastructure); and
- Laying the groundwork for resolving some of the important macro issues that we discuss, including the future of work and how to prepare people for a changing work environment and the fiscal, and safety net, issues that may arise with AI and how to prepare for these impacts.

We also suggest an oversight body, which we discussed in the previous section.

[37] https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf

[38] https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-government-outlines-ai-ambitions-through-2020/

This is an area where potential international or regional collaboration can be highly productive. Many countries are encountering these issues at almost the same time. Also, some of the actors involved will include global international businesses, that will be viewing these issues on an international or regional basis.

We raise many other issues in this introductory module, and in this AI Series, but it is not necessary for individual governments to resolve all of these issues up front, at least in the first iteration of the AI national plan. Some of these issues will take time to emerge and show their true contours and, apart from maintaining awareness of them, will probably not need to be addressed immediately. Other issues are likely to be substantially resolved at a global level and may not require resolution locally.

Bibliography

Artificial Intelligence Index, 2017 annual report, November 2017

British Academy and Royal Society, Data management and use: governance in the 21st century, June 2017

British Parliament Report on AI (https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/14504.htm)

Calo, Artificial Intelligence Policy: A Primer and Roadmap, August 2017

Elliot, 2017. Computers and the Future of Skill Demand. (http://www.oecd-ilibrary.org/education/computers-and-the-future-of-skill-demand_9789264284395-en)

European Parliament, Civil law rules on robotics, February 2017

Government of the Republic of Korea, Mid to long term master plan in preparation for the intelligent information society, [date?]

House of Lords, Select Committee on artificial intelligence, AI in the UK: ready, willing and able?, April 2018

IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html)

Levendowski, How copyright law can fix artificial intelligence's implicit bias problem, July 2017

McKinsey Global Institute, Artificial intelligence the next digital frontier, June 2017

McKinsey Global Institute, Harnessing automation for a future that works, January 2017

Obama White House: Office of Science and Technology Policy, Preparing for the Future of Artificial Intelligence, May 2016

Obama White House: National Science and Technology Council, The national artificial intelligence research and development strategic plan, October 2016

Omer Tene and Jules Polonetsky, Big Data for All: Privacy and User Control in the Age of Analytics, 11 Nw. J. Tech. & Intell. Prop. 239 (2013). http://scholarlycommons.law.northwestern.edu/njtip/vol11/iss5/1

Royal Society; Machine learning: the power and promise of computers that learn by example, April 2017

Stanford University, Artificial intelligence and life in 2030, September 2016

The Economist, 2017. Retraining low-skilled workers. (https://www.economist.com/news/special-report/21714175-systems-continuous-reskilling-threaten-buttress-inequality-retraining-low-skilled)

World Economic Forum (https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/)

World Economic Forum, White Paper, Digital Transformation Initiative, Telecommunications Industry, January 2017 (https://www.accenture.com/t20170411T115809Z__w__/us-en/_acnmedia/Accenture/Conversion-Assets/WEF/PDF/Accenture-Telecommunications-Industry.pdf)