4. AI in the Public Interest

Jianlong Zhou and **Fang Chen**
University of Technology Sydney (UTS), Sydney, Australia
Email: {jianlong.zhou, fang.chen}@uts.edu.au

4.1 Introduction

Like the steam engine in the first industrial revolution, the electricity in the second industrial revolution, and the electronics and information technology in the third industrial revolution, Artificial Intelligence (AI), which has powerful capabilities in prediction, automation, planning, targeting, and personalisation, is claimed to be the driving force of the next industrial revolution (Industry 4.0)[1]. It is transforming our world, our life, and our society and affects virtually every aspect of our modern lives. Specially, AI enables the monitoring of climate change and natural disasters, enhances the management of public health and safety, predicts crops output in agriculture, automates administration of government services, and promotes productivity for economic wellbeing of the country. AI also helps to prevent human bias in criminal justice, enables the efficient fraud detection (e.g. in welfare, tax, trading), enhances the protection of national security (e.g. with face recognition), and others. Therefore, AI can benefit a large range of public interest matters and deliver a revolutionary change in both the efficiency and effectiveness of services in the public interest[2].

This chapter presents how AI serves in the public interest with a focus on AI for enhancing the quality of human life in the area of transport, infrastructural assets, energy, and education. This chapter also demonstrates the capability of AI in supporting better decision making especially decision making in government. Since the ethical implications of AI use are among the top concerns of users, we then introduce privacy related issues in the digital age, and discuss how AI compromises privacy and enhances privacy respectively. Finally, barriers to AI are discussed, which shows that the barriers are not mainly technical but social and cultural, before the summary of the chapter.

4.2 Smart AI applications affect the quality of life

Generally, it is assumed that AI can enable machines to exhibit human-like cognition, and it is more efficient (e.g. higher accuracy, faster, working 24 hours) than humans in various tasks. Claims about the promise of AI are abundant and growing related to different areas of our lives. Some examples are: in human's everyday life, AI can recognise objects in images, it can transcribe speech to text, it can translate between languages, it can recognise emotions in images of faces or speech; in traveling, AI makes self-driving cars possible, AI enables drones to fly autonomously, AI can predict parking difficulty by area in crowded cities; in medicine, AI can discover new uses for existing drugs, it can detect a range of conditions from images, it enables the personalised medicine; in agriculture, AI can detect crop disease, and spray pesticide to crops with pinpoint accuracy; in finance, AI can make stock trades without human intervention, and handle insurance claims automatically; AI can identify potentially threatening weather in meteorology; AI can even conduct various creative work, such as paint a van Gogh painting, write poems and music, write film scripts, design logos,

[1] Sentryo, 2017. The 4 industrial revolutions, https://www.sentryo.net/the-4-industrial-revolutions/
[2] Australian Law Reform Commission. Meaning of public interest, https://www.alrc.gov.au/publications/8-balancing-privacy-other-interests/meaning-public-interest

recommend songs/films/books you like. These diverse and ambitious claims motivate wide adoptions of AI in various sectors including retail, education, healthcare and others. According to surveys by Mckinsey (Bughin et al., 2017), the leading sectors in AI adoption today are mainly high tech and telecommunications, automotive and assembly, financial services, resources and utilities, media and entertainment, consumer packaged goods followed by transportation and logistics as well as others. All these adoptions will ultimately help to deliver a better quality of life with manageable cost of living, better environment, easy access of transport for time saving, etc. This section specifically focuses on the use of AI in transport, infrastructural assets, energy, and education to show how it affects quality of human life for public good.

4.2.1 AI in Transport

With the growing urbanisation and the improvement of people's life, more and more vehicles are on urban roads, which results in significant transport issues such as heavy congestions and serious accidents. These transport issues may cause big economic and social losses. For example, the overall social cost of congestion was $16.3 billion (measured in delay cost by lane kilometre) in 2015 in Australia estimated by Infrastructure Australia. Road users in Sydney and Melbourne need to allow an average 50% more time than free flow to complete their journeys during peak hours, when average speeds can be as low as 29 km/h and 34 km/h respectively[3]. AI can use data from different sources such as road cameras, mobile phones, road networks, and even social media to set up Machine Learning (ML) models, which can revolutionise transportation from different aspects:

- **Traffic management:** AI is already being used in prediction and detection of traffic accidents and congestions. Traffic simulation is also used to simulate what actions can be conducted when traffic accidents or congestions are taken place. These solutions allow to monitor real-time operations of transport networks and identify operational anomalies in a way transport operators and travellers can use for making better decisions for traffic control problems. For example, AI has been used along with collating traffic data to reduce congestion and improve the scheduling of public transport in some cities.
- **Corporate decision making for transport:** AI can predict volumes and speed for the freight transport system. By monitoring the requesting for and providing to volumes with origin-destination pairs, AI can simplify the planning for freight transportation companies for more productive routes and investments.
- **Traffic safety:** AI can predict locations and/or frequencies of traffic accidents in order to take special actions for safety in advance. Pedestrians and cyclists are the most vulnerable road users to serious injuries in a traffic accident. AI can also predict the path of pedestrians or cyclists, which would help in decreasing instances of traffic accidents and injuries. For example, with the predicted path of pedestrians or cyclists, transport signals (e.g. red/green lights at cross-roads) can be automatically adapted to control the movement speed and paths of vehicles on the road to avoid potential traffic accidents with pedestrians or cyclists.
- **Autonomous vehicles:** AI is used to sense the environment of running vehicles and helps vehicles conduct various operations under different circumstances automatically. There are different types of autonomous vehicles. Self-driving cars, driverless buses, and driverless trains could improve the traffic safety by automatically sensing risks around them, resulting in the overall improvement of public safety and economic as well as environmental benefits. For example, computer vision and other AI technologies can be used to recognise objects surrounding running vehicles. If a walking human is recognised in front of a running vehicle, AI can suggest to decrease the running speed of the vehicle or change the lane of the vehicle. Such autonomous vehicles could enhance mobility of people who cannot drive and

[3] Austroads, 2016. Congestion and Reliability Review, https://austroads.com.au/publications/network/ap-r534-16

dramatically change how we get from one place to another. They offer every single person lifestyle benefits on vehicles being able to read a book, watch a movie or talk with others with the optimised path. In such cases, the current traffic rules such as no phone or messaging during driving can be completely removed. Another type of autonomous vehicles is self-driving trucks and remote-controlled cargo ships which can relieve drivers from highly intensive workloads and dangerous work conditions. Autonomous delivery trucks can change the way we receive goods with faster speed and optimised path for significant economic and environmental benefits.

- **Traffic network optimisation and planning:** From the strategical perspective, AI can predict future transport requirements based on historical transportation volume, population increase and other factors. This will help authorities to optimise the traffic network and make reasonable and forward-looking plans for both traffic network and city planning.

Therefore, with the use of AI, traffic accidents and congestions will be well predicted for the improved decisions of public transport systems. The safety for both vehicles and vulnerable road users (e.g. pedestrians and cyclists) will be significantly enhanced with the use of smart predictive technologies. Autonomous vehicles will further enhance public safety and improve the mobility of people with economic and environmental benefits. Beyond these safety, economic, and environmental benefits, AI also benefits the traffic management and network planning, resulting in significant improvement of governmental administration. These demonstrate the dramatic impact of AI in enhancing public interest from the traffic perspective in areas of public safety, administration of government services, and productivity for economic wellbeing of the country[2].

4.2.2 AI in Infrastructural Assets

Infrastructural asset management is one of key strategies conducted by government in sustaining public infrastructural assets such as water supply networks, roads, and bridges. It generally focuses on the maintenance, replacement, and rehabilitation of assets in the later stages of their service life cycles[4]. An effective implementation of infrastructural asset management can not only benefit to the government economically but also help to decrease the adverse effect caused by potential failures of assets. This subsection demonstrates the benefits of AI in infrastructural asset management especially in water supply networks, roads, and bridges.

Water supply networks constitute one of the most crucial and valuable urban assets. The combination of growing populations and aging pipe networks requires water utilities to develop advanced risk management strategies in order to maintain their distribution systems in a financially viable way (Li et al., 2014). Especially for critical water mains, defining risk based on the network size or location (e.g. a single trunk line connecting distribution areas or under a major road) , failure of such pipes typically brings severe consequences due to service interruptions and negative economic and social impacts, such as flooding and traffic disruption (Li et al., 2014; Zhou et al., 2017). For example, Australian Sydney Water had 129.5 ML water leakages per day in 2017-2018, and invested around $1.1 million in 2017-2018 on active leaking detections alone[5]. It took a much more amount of dollars to fix those leakages. From an asset management perspective there are two goals for water pipe management (Whiffin et al., 2013): 1) Minimise unexpected pipe failures by prioritising timely renewals; 2) Avoid

[4] Wikipedia, 2019. Infrastructural asset management, https://en.wikipedia.org/wiki/Infrastructure_asset_management

[5] Sydney Water, 2018. Sydney Water Water Conservation Report 2017-2018, https://www.sydneywater.com.au/web/groups/publicwebcontent/documents/document/zgrf/mdq3/~edisp/dd_047419.pdf

replacing a pipe too early before the end of its economic life. AI can help to predict and identify high-risk pipes before failures. In such way, it is likely that repairs can be completed with minimal service interruption, water loss and negative social and economic impacts. Besides, AI can also be used to prioritise sewer pipes in terms of corrosion risk with a corrosion prediction model built on influential factors that cause sewer corrosion, such as hydrogen sulphide (H2S). Based on the estimation of H2S, reasonable chemicals can be put to corrosion locations to control H2S to reduce the level of sewer corrosions. Therefore, AI can help waste water management systems to more efficiently monitor sewer corrosion and to more effectively optimise sewer water processing by suggesting reasonable chemical dosing at the right location to lessen environmental impacts, resulting in the improvement of environmental sustainability.

Besides water pipes, AI can also be used for the predictive maintenance of other public infrastructure and equipment such as roads and bridges. For example, the road maintenance is one of key responsibilities for government. Thanks for the mobile technologies and various sensors on modern vehicles especially on autonomous vehicles, the condition of the road can be monitored in real-time driven by the data from both passengers (e.g. their mobile phones) and vehicles that are available on each vehicle, instead of using expensive special road inspection vehicles that are based on radar, high-definition cameras and LiDAR technologies. With the use of AI, the information on the location, size and type of damage for road defects identified can be sent to road management authorities automatically in order to take further actions.

Taking the bridge maintenance as another example, the traditional approach to inspect the condition of a bridge is usually conducted by professional people manually at some time intervals, which is time-consuming, less reliable and cannot monitor the condition continuously in real-time. AI technologies can dramatically change the ways for the inspection of bridge conditions. This can be conducted through the installation of sensors for collecting data on physical changes of the bridge and conducting the anomaly detection from collected data using AI technologies. Different sensors (e.g. vibration sensors, displacement sensors) can be installed on key locations of the bridge to record any physical changes (e.g. vibration, displacement) of the bridge. For example, the Sydney Harbour Bridge[6], which was completed in 1932, needs to be inspected regularly to keep the healthy condition of this Sydney Icon. But finding faults along the 1149-metre long deck and 134-metre high steel arch bridge visually by workers is a hard, unreliable and time-consuming process. CSIRO's Data61 has developed an AI system to continuously monitor the structural health of the bridge and provide early warnings of problems before the bridge services are affected based on predictive analytics[7]. Around 2400 sensors have been installed on the bridge to collect information on the condition of the bridge to keep this intelligent monitoring system work continuously.

In a word, it is a key responsibility for the government to keep public infrastructural assets run smoothly for general public good. AI technologies can help to continuously monitor conditions of assets and give warnings before they fail as well as locate failures in assets effectively. These advantages allow government agencies to make active actions before failures of infrastructural assets, which may cause service interruptions and negative economic and social impacts. Therefore, AI improves the management of public infrastructural assets resulting in the significant benefits to the public safety and economic wellbeing of the country, which are major public interest matters.

[6] Wikipedia, 2019. Sydney Harbour Bridge, https://en.wikipedia.org/wiki/Sydney_Harbour_Bridge

[7] Data61, CSIRO. Helping to maintain Sydney Harbour Bridge, https://www.data61.csiro.au/en/Our-Work/Future-Cities/Planning-sustainable-infrastructure/Structural-health-monitoring

4.2.3 AI in Energy

The provision of adequate, reliable and affordable energy to meet future energy consumption needs is one of government's significant missions for people's high quality of life and strong economic growth. AI can help to achieve such ambitious objectives, especially in the generation of renewable energy and efficient use of energy ultimately benefiting our environment.

With the significant increase in the use of renewable energy generation systems, the electricity smart grids are experiencing a shift from the focus of creating intelligent energy distribution and flows within the existing grid structure, to the focus of using AI to restructure the grid by bringing in new, diverse and decentralised energy sources such as solar and wind power[8]. New technologies such as batteries for energy storage and Electrical Vehicles (EVs) propose further requirements for the grid. This kind of future grid can be a complex network with both generation and distribution assets. It can be expected to intelligently match supply and demand and operate automatically or semi-automatically. The future grid can also measure or predict individual customer needs, balance their needs at different time or for different purposes, and then make appropriate actions through the network, aiming to deliver customised energy management solutions. The adoption of AI and smart automation will aid the future grid from different perspectives:

- **AI for matching demand and supply:** The modern electricity ecosystem usually includes a mix of traditional energy (e.g. hydropower or thermal power) and renewables (e.g. solar, wind power) as well as energy storages. However, renewable energy sources are weather dependent and therefore highly unpredictable, making it a challenge to match demand and supply. AI can offer solutions to demand management problems by using predictive analytics to accurately estimate renewables to balance grids.
- **AI for energy efficiency and reliability:** AI can help to improve the economic efficiency of energy. For example, AI can monitor and optimise the turbine parameters of wind power to increase its energy production. The turbine with less performance can be detected by monitoring the generation of different turbines in a wind farm. The parameters (e.g. blade direction relative to the wind direction) of the turbine with less performance can then be optimised based on other turbines in the wind farm. AI can also automatically detect anomalies/faults by monitoring various parameters such as smart meters of the energy infrastructure. Anomalies that are item sets of events that may be anomalous can be identified based on their patterns of appearance in the smart metre data stream.
- **AI and the consumer:** By monitoring the energy consumption behaviour of individuals and businesses, AI can offer solutions to tailor customers' energy consumption and reduce costs by giving suggestions on how and where customers can save energy. Customers can also benefit from AI in the supplier selection, for example, AI learns a customer's energy consumption and generation profile and then matches the most suitable offers from the market to automatically make the switch of suppliers.

Overall, AI can help to provide adequate, reliable and affordable energy by using predictive analytics to accurately match demand and supply, optimising parameters of renewable energy for energy efficiency, detecting anomalies automatically, and providing customised energy solutions for individual consumers. These functions will maximise the use of renewable energy and encourage the efficient use of energy, resulting in the dramatic benefits to our global environment.

[8] Giovanni Polizzi and Ariel Liebman, 2019. AI in Australia's electricity sector, https://www.ecdonline.com.au/content/electrical-distribution/article/ai-in-australia-s-electricity-sector-49048136#axzz5mXPOwyKc

4.2.4 AI in Education

Education is and will always be a foundational part of humanity. Regardless of age, we're constantly developing new skills and understandings. AI technologies are well suited to achieving crucial education objectives, such as enhancing teaching efficiency and effectiveness, providing lifelong education for all, and developing personalised learning.

- **Adaptive learning and personalised teaching:** The conventional classroom teaching delivers one lesson to the entire class without considering individual differences in learning, which not only makes individuals frustrated when they cannot follow the teaching speed, but also wastes the time of students who have already grasped the concepts. AI can improve adaptive learning and personalized teaching by identifying factors or indicators related to learning efficiency. Those factors or indicators are derived from students' behavioural, physiological information or even learning materials. For example, behavioural features such as features from writing activities, mouse movement activities as well as keyboard activities can be derived during the student's learning and link these features with the student's learning performance, confidence, mindset, and cognitive ability. Furthermore, AI can point out places where courses need to improve and adapt teaching materials to student needs. As a result, the learning can be adapted based on student's responses in a loop for a better learning performance. Based on the learning profile derived from student's behavioural, physiological information and learning materials, personalised teaching can be performed for each student by AI. Taking the mathematics learning as an example, the cognitive load level of a student during learning, which is related to the difficulty level of mathematic questions, can be estimated through the examination of the student's behavioural features such as writing speed, orientation of the pen, pressure of the pen tip, etc. (Chen et al., 2016). When the cognitive load level of the student is too high or too low, the difficulty level of mathematic questions can be adaptively adjusted (i.e. decrease or increase the difficulty level) in order to keep the cognitive load to an appropriate level to maximise the learning performance of the student. Therefore, mathematic questions for each student are customised according to their cognitive load levels during learning.
- **Teacher and AI collaboration:** Teacher and AI have complementary strengths and abilities, for example, a teacher has strengths in high level guidance and creativity, while AI has strengths in analysing student's responses to learning materials. It is expected that AI can help fill needs gaps in learning and teaching that schools and teachers cannot provide[9]. For example, AI can provide personalised and streamlined teaching to students by analysing each student's responses and performance during learning, which allows teachers to have the freedom to provide unique human capabilities of high-level guidance, high-order thinking and creativity. Teachers and AIs can form a team in teaching for the best outcome for students.
- **Tutoring and support outside the classroom:** Tutoring outside the classroom is often limited because teachers are not always available. Students can get additional support from AI tutors all the time without the limitation of locations (no matter whether students are in urban areas or in remote rural areas). Most importantly, AI can also provide feedback to students on their success in the tutoring.

[9] Bernard Marr, 2018. How is AI Used in Education -- Real World Examples of Today and A Peek into The Future, https://www.forbes.com/sites/bernardmarr/2018/07/25/how-is-ai-used-in-education-real-world-examples-of-today-and-a-peek-into-the-future/

AI would shift the role of teachers in education. Because AI can help students with adaptive and personalised learning and provide tutoring all the time, teachers will supplement AI based learning, assisting students who are struggling as well as providing mentoring and coaching hands-on experiences for students — value-adding tasks that are uniquely suited to human beings (Bughin et al., 2017).

Overall, with the use of AI, students not only can have efficient learning based on their own capabilities, but also can achieve the lifelong learning which is significant for people in modern society with the rapid advancement all the time. Most importantly, AI can help to achieve the equity of learning for people, no matter how old they are, where they are from, or what professions they are in. All these qualities of AI are interests of public for the more advancement of society.

4.3 AI supports better decision making

AI can help humans by automating tasks that would take much human power or time to deal with. It can also get insights from data to find patterns that are usually difficult to catch by humans. For example, AI has powerful capabilities in coordinating data delivery, extracting data trends, making predictions, quantifying uncertainty, checking data consistency, generating new data, anticipating the user's needs, providing expected information to the user, and suggesting courses of action. Such capabilities can augment human intelligence dramatically in tasks and enable decision making process automatically. As a result, AI powered decision making will not only improve decision quality by reducing human errors and biases that are common with heuristic decisions, but also decrease human's workload involved in critical decision making. Therefore, AI can deliver revolutionary impact on how humans make decisions.

4.3.1 AI and government decision making

McKinsey's analysis of more than 400 use cases representing $6 trillion in values across 19 industries and nine business functions shows the broad use and significant economic potential of AI (Chui et al., 2018). In more than two-thirds of the listed use cases, AI is used to improve the performance beyond that provided by other analytics techniques. For example, AI is already helping financial institutes augment financial planning and investment strategy. AI powered diagnostics systems have proven to be more accurate than human doctors in diagnosing serious disease.

Specifically, applications of AI to the public sector are broad and growing. Today the sources of information accessible to government are much massive ranging from organisation data, program data, public agency data, service data, Medicare data, and data created by different Internet of Things (IoT) devices, as well as many others. With such flood of information, AI can benefit government decisions at least in the following ways:

- **To streamline or automate high frequency, high workload required decisions:** Typical examples of such decisions include making welfare payments and immigration decisions. AI can improve the quality of decisions and reduce the cost of services by automating time-consuming manual bureaucratic processes.
- **To make decisions in complex public sector problems:** AI can identify leading indicators that signal potential problems in public sectors. For example, tax fraud detection is a serious problem that government agencies and departments meet. AI can help to identify such problems, enabling government agencies and departments make informed decisions.
- **To make strategic decisions:** With the powerfulness of using a large amount of data from various sources for analyses and predictions, AI can help government to make high level strategic decisions for different public affairs, such as policy changes or setup of new policies. For example, AI can help government to identify what skill sets required for a

particular program and make workforce planning. AI can also make predictive analytics for the requirement of infrastructural assets in a suburb and help government to make planning decisions.

4.4 AI and privacy

AI usually requires huge volumes of data in order to learn and make decisions. Because of such heavy demand of data especially personal data from AI, the concern of privacy becomes one of important issues in AI. There is no fixed definition of privacy. Broadly speaking, privacy can refer to the right to be let alone, or freedom from publicity by surveillance, interference or intrusion[10], and have control over one's own personal information[11]. From the digital perspective, privacy implies on the ability to control how data especially personal data is being collected, stored, modified, used, and exchanged between different parties[12]. This section firstly presents privacy related issues in the digital age because of data mining and machine learning technologies. We then discuss how AI can compromise privacy and enhance privacy.

4.4.1 Privacy issues in digital age

Britz (Britz, 1996) distinguishes privacy into four categories namely private communications, privacy of the body, personal information, and information about one's possessions. Finn et al. (Finn et al., 2013) further identifies seven types of privacy by considering advances of technologies in digital age:

- **Privacy of the person:** It refers to the right to keep body functions and body characteristics (such as genetic codes and biometrics) private;
- **Privacy of behaviour and action:** It includes sensitive information such as sexual preferences and habits, political activities and religious practices;
- **Privacy of communication:** It refers to avoiding the interception of private communications, examples include mail interception, telephone or wireless communication interception or recording, access to email messages, and interception with devices such as directional microphones;
- **Privacy of data and image:** It includes concerns about making sure that individuals' data and images are not automatically available to other individuals and organisations and that people can control over own data and its uses;
- **Privacy of thoughts and feelings:** It refers to the right not to share individuals' thoughts or feelings or to have those thoughts or feelings revealed. Individuals should have the freedom to think whatever they like;
- **Privacy of location and space:** It means individuals have the right to move about in public or semi-public space without being identified, tracked or monitored;
- **Privacy of association (including group privacy):** It is concerned with people's right to associate with whomever they wish, without being monitored.

A large amount of data for AI are related to these one or more of seven types of privacy issues. Therefore, both AI technology developers and users need to consider in providing proactive protection to individuals in the face of AI technologies. For example, in order to protect data and privacy for all individual citizens of the European Union (EU) and the European Economic Area, the

[10] International Association of Privacy Professionals (IAPP). What does privacy mean? https://iapp.org/about/what-is-privacy/

[11] Office of the Victorian Information Commissioner (OVIC). What is privacy? https://ovic.vic.gov.au/privacy/what-is-privacy/

[12] Michael Deane, 2018. AI and the Future of Privacy, https://towardsdatascience.com/ai-and-the-future-of-privacy-3d5f6552a7c4

EU implemented the General Data Protection Regulation (GDPR) in 2018. The provisions of the GDPR govern the data controller's duties and the rights of the data subject when personal information is processed. It applies when AI technologies are under development with the inclusion of personal data.

4.4.2 AI compromises privacy

It is because of large amounts of data that make AI powerful and function precisely as expected. Furthermore, AI can perform the designed tasks automatically without human's intervention. Although AI can provide efficiencies that we all want, we have to remember that these can come at a cost to affecting our privacy consciously or unconsciously in different ways[12]:

- **Data collection:** For a typical day of a human, he/she gets up in the morning, catches a bus or train to work, surfs webpages during work, goes to a restaurant to have a lunch, goes to supermarket for shopping, etc. All these activities (e.g. travel records, shopping records) even if sleeping states can be recorded by mobile devices or different systems from different sites. Because such a large amount of data is recorded, people are mostly unaware how much data their devices and systems generate, process, or share. With more advanced technologies and devices are introduced, more personal data will be exposed to AI for increased privacy concerns.
- **Identification and tracking:** AI has the capability to identify and track individuals across different data sources, resulting in shadow profiles of individuals. Even if personal data is anonymised, AI can de-anonymise the data based on inferences from different data sources. For example, researchers have developed a way based on Convolutional Neural Network (CNN) (CNN is a type of artificial neural network mostly used to analyse visual imagery) to identify and track individual animals by using animal's movements but without facial recognition[13]. Such results could eventually be applied to public surveillance of humans by using individual's movements such as gaits, which again create privacy concerns.
- **Facial and speech identification:** Face and voice are two typical signatures that we notice or hear to recognise someone. Successful business products based on facial and voice recognition have been developed. For example, many financial service providers are using face biometrics for authenticating transactions (e.g. different banks in China provide ATM services of withdrawing cash by scanning the face). Face biometrics are also used for traffic check-in entrances or border entry/exit self-services in many international airports including Australian Sydney Airport[14]. These identification applications require very strict privacy management. Such identifications help to improve the user experience from a consumer's perspective (e.g. easy authentication and fast check-in services) and the performance of authority management for public good. However, the technology also means that people's biometric data is held by the government or other authorities that may then use those data for other purposes leading to the erosion of privacy.
- **Personal profiling:** Besides individual identification and tracking capabilities as described above, AI also has the ability to use multiple source data as input for the purpose of describing some characteristics of a person, such as credit ranking, trustworthiness profiling.

[13] Chris Burt, 2019. Researchers develop AI method for movement identification and tracking without facial recognition, https://www.biometricupdate.com/201901/researchers-develop-ai-method-for-movement-identification-and-tracking-without-facial-recognition

[14] Kate Schneider, 2018. Big change coming to the way we fly, https://www.news.com.au/travel/travel-advice/flights/big-change-coming-to-the-way-we-fly/news-story/e6db13688970c596a9a278a394c122ca

These personal profiles are also people's privacy and not expected to be revealed to public most of the time.

- **Prediction:** AI can use machine learning to infer or predict sensitive personal information from non-sensitive forms of data. For example, a person's mouse movement behaviour or body movement behaviour can be used to infer their emotional and mental state such as cognitive load and stress (Chen et al., 2016). Even more, AI is capable of predicting a person's political views from collected data such as activity logs and social network messages. These predictions lead to the erosion of privacy.

4.4.3 AI enhances privacy

AI techniques pose a threat to privacy as presented in previous sections. As one coin has two sides, AI can also enhance privacy. By conquering the reasons that result in privacy leakages in AI, innovative methods can be developed to enhance privacy. For example, since AI usually needs a large amount of data for training a model, whether is it possible to reduce the number of training data but still keep a good performance of the model? Or if we do not reduce training data, whether is it possible to encode training data in a way to protect privacy? Whether the explanation of how the data is processed by the model and how the outputs are generated can help to relieve privacy concerns? Therefore, AI can enhance privacy from at least the following perspectives[15]:

- **Reducing the need for training data:** AI needs a large amount of data for the machine learning model training. Fortunately, different techniques have been developed for generating synthetic data which may reduce the requirement of training data and decrease the privacy related risks. Generative Adversarial Networks (GAN) is one of the popular methods for generating the synthetic data. This meets the needs of large amount of data for the training of machine learning models, without the use of data containing real personal information.
- **Upholding data protection without reducing the basic dataset:** Many AI models are trained involving sensitive data such as personal information. Ideally, the parameters of trained machine learning models should encode general patterns rather than facts about specific training examples to protect privacy concerns. AI techniques can help to overcome privacy concerns. For example, Google published a library named "TensorFlow Privacy" [16] for its TensorFlow machine learning framework intended to make it easier for developers to train AI models with strong privacy guarantees. It is based on the principle of differential privacy, a statistical technique that aims to maximize accuracy while balancing the users' privacy, and can prevent the memorisation of rare details (McMahan et al., 2018).
- **Enhancing knowledge share without centralised training data**: Personal data is usually located on isolated islands such as mobile phones, private cloud drivers, private photo albums, etc. While standard AI models are usually trained with centralised training data, which may cause privacy concerns. New machine learning techniques have been developed to relieve this problem. For example, federated machine learning allows to train a model without centralised training data and compromising data privacy (Yang et al., 2019).
- **Avoiding the "black box" issue:** AI is a "black box" for general users. It accepts inputs and generates outputs but does not disclose its internal working. Users do not know how the

[15] Datatilsynet, The Norwegian Data Protection Authority, 2018. Artificial Intelligence and Privacy, https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf

[16] Garey Radebaugh and Ulfar Erlingsson, 2019. Introducing TensorFlow Privacy: Learning with Differential Privacy for Training Data, https://medium.com/tensorflow/introducing-tensorflow-privacy-learning-with-differential-privacy-for-training-data-b143c5e801b6

inputs are processed and how the outputs are generated. This is a challenge for both people who use AI systems and those whose data is used by systems, which results in privacy concerns. Much research has been conducted to make AI systems transparent. Explainable AI (XAI) is one of those approaches which involve people in the process and give explanations for AI outcomes. For example, XAI can explain how the training data is processed by a decision tree algorithm to get output to let users understand that their data is processed in a right way that does not have privacy hackings in the algorithm. As a result, privacy concerns can be relieved because of explanations.

Besides AI techniques, other information technologies can be used to overcome privacy concerns. Such technologies range from ethically-informed design methodologies to using encryption to protect personal information from unauthorised use (van den Hoven et al., 2018). The examples include browsing through anonymous networks which encrypt information in the network, using open-source web browsers such as Firefox which are freely audited for vulnerabilities, and using open-source operating systems such as Linux distributions on which users can have a full check of security backdoors.

4.5 Barriers to AI

While we continuously find ourselves coming across appealing AI-based systems that seem to work (or have worked) surprisingly well in practical scenarios, AI is currently still facing prolonged barriers with user acceptance of delivered AI solutions as well as the availability of data, ethics and legal issues.

4.5.1 Human trust in AI

A key barrier to AI is the human trust in AI technologies and AI-based solutions (Zhou and Chen, 2018). Trust is defined as “the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability” (Lee and See, 2004). This definition shows that uncertainty is tightly coupled to trust. While uncertainty is a common phenomenon in AI technologies and machine learning, which can be found in every stage of learning from input data and its preprocessing, algorithm design, feature selection, to model evaluation and others. Besides uncertainty, the trust issues in AI is mainly caused from the black box nature of machine learning techniques, where users are unaware what is going on inside a machine learning algorithm and how the prediction results are based on input data. Therefore, approaches need to be developed to make AI explainable and transparent for trustworthy decisions driven by AI.

4.5.2 Data issues in AI

An abundance of high quality data is critical for AI training systems. Poor data quality (such as sparse data and missing data) is the number one enemy to the widespread and high performance of AI. The saying of “garbage-in, garbage-out” has plagued analytics and decision making for generations, and this is especially a warning for AI solutions. Data availability is also a big issue in some special areas for training AI with high accuracy. If few data is available for the model training, the overall AI model performance will be decreased with overfitting or low accuracy and other problems. The examples include the cases of some special disease in medicine and the cases of failures of some critical assets, where there are small number of examples with disease or failures. In order to overcome these issues, one of approaches is that government can make public and private sector datasets available for reuse by others in applications of AI that serve the public interest, such as for healthcare or environmental protection. Besides, data issues can also be caused because of privacy concerns as discussed in the previous section.

4.5.3 Government policies/legislation on AI

With the continuous growth of AI uses in various areas, the roles of government are increasingly expected to play in AI regarding both data and the adoption of AI. Privacy is a major concern from the data perspective. Europe's GDPR introduces strict consent requirements for data collection, data processing and the control of the use of data. The essence of GDPR is that personal information shall be utilised in a way that protects the privacy of the data subject in the best possible way, and that each individual has the right to decide how his/her personal data is used[15].

Furthermore, the ethical implications of AI use in government are among the top concerns of people (Carrasco et al., 2019). Developing an ethics framework for AI is becoming of increasing importance, which can regulate what are principles for an ethical AI and how these principles can be implemented. CSIRO's Data61 in Australia has developed a discussion paper which informs Australia's ethics framework for AI (Dawson et al., 2019). Similar frameworks have also been developing in some of other countries. For example, New York in US has reviewed key systems used by government agencies for accountability and fairness; Germany has developed government-led ethical advices on the ethics of automated vehicles.

All these endeavours are currently at their early stages. Progress on these issues will benefit both to realising the potential impact of AI and to safeguarding people from the risks possibly caused by AI, resulting in the significant improvement of public interest matters from the perspectives of privacy and safety when using AI.

4.5.4 Other barriers to AI

Besides issues met by AI as mentioned above, there are also other barriers to AI for its profound impact. Some examples include:

- Currently, the adoption of AI tends to be more concentrated in relatively digitised industries which have access to massive amounts of data collected by their own infrastructures (e.g. Amazon, Alibaba, Facebook, etc.). Broader adoption of AI in different domains and especially in smaller firms could be important to drive improvements in product quality, performance, and markets. The barriers to these smaller firms to adopt AI mainly lie in the lack of data availability as well as the professionals specifically working on AI within the firms. Broader uptake of AI needs to be encouraged.
- Furthermore, the automation driven by AI has an important impact on both jobs and skills potentially required. Employees are afraid of losing their jobs to automation, while employers are also worrying about the difficulty to find people whose skills and capabilities are best matched to the AI-driven work that needs doing. Therefore, concerns on employment and required skills need to be addressed for the wider adoption of AI.

The adoption of AI within government is still relatively low. The possible reasons for this phenomenon could be the ethical concerns of AI such as fairness, transparency and explainability, accountability, and privacy. For example, can AI be prevented from conscious or unconscious bias based on historical data? How to make users accept decisions based on logics from "black box" deep-learning models. Those logics are even difficult to explain or understand by data scientists. While the strong capabilities of AI in prediction, automation, planning, targeting, and personalisation could deliver a revolutionary change in both the efficiency and effectiveness of government services, the adoption of AI within government needs to be strengthened, especially in highly expected areas such as healthcare and education.

4.6 Summary

In summary, AI has powerful capabilities in prediction, automation, planning, targeting, and personalisation. AI has dramatic impact on enhancing public interest matters from the traffic perspective. It can improve the management of public transport systems by accurately predicting traffic accidents and congestions. AI can also enhance the traffic safety and improve the mobility of people with significant economic and environmental benefits. In public infrastructural asset management, AI can benefit the public safety and economic wellbeing by continuously monitoring conditions of assets and giving warnings before they fail for making active actions to avoid any serious social and economic consequences. AI can also help to provide adequate, reliable and affordable energy by maximising the use of renewable energy and encouraging the efficient use of energy, resulting in the dramatic benefits to our global environment. In addition, AI can provide solutions for personalised and lifelong learning, and help to achieve the equity of learning for people which are interests of public for the more advancement of society. Therefore, AI can significantly benefit the public interest in safe and efficient transport, smoothly running infrastructural assets, environmental good energy, and personalised and equal learning, resulting in dramatically improved quality of human life.

Moreover, AI can deliver revolutionary impact on how humans make decisions by not only improving decision quality, but also decreasing human's workload in making critical decisions. Specifically, AI can help government to streamline and automate high frequency, high workload required decisions, to make decisions in complex government service problems such as tax and welfare fraud detections, and to make high level strategic decisions for different public affairs.

However, since AI usually requires large amounts of data to learn and make decisions, the concern of privacy becomes one of important issues in AI. For example, more personal data exposed to AI, identification and tracking capabilities of AI, facial and speech identification technologies, and personal profiling may affect individual privacy consciously or unconsciously. Fortunately, AI has the capability to enhancing privacy. For example, a generative approach of GAN can be used to generate synthetic data to reduce the need for real training data, training data examples can be encoded with AI techniques to hide facts to protect privacy, and well-explained AI models can relieve concerns of privacy.

Despite the rapid advancement of AI in both technology development and applications, AI is currently still facing prolonged barriers with user acceptance of delivered AI solutions as well as data, ethics and legal issues. Furthermore, the adoption of AI tends to be more concentrated in relatively digitised industries which have access to massive amounts of data. Smaller firms have the barriers to broader adoption of AI because of the lack of data availability as well as the professionals specifically working on AI. There are also various social concerns that limit the use of AI in wider applications related to public interest.

In the future, besides the development of more advanced AI techniques to better solve human interest matters, more attention needs to be paid to ethics and legal issues as well as social concerns related to AI. Government will play an irreplaceable role in setting up ethical frameworks and policies to relieve these concerns and promote AI benefits to public interest.

References

Britz, J.J., 1996. Technology as a Threat to Privacy: Ethical Challenges and Guidelines for the Information Professionals. Microcomput. Inf. Manag. 13, 175–93.

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Henke, N., Trench, M., 2017. Artificial Intelligence: The Next Digital Frontier? McKinsey Global Institute.

Carrasco, M., Mills, S., Whybrew, A., Jura, A., 2019. The Citizen's Perspective on the Use of AI in Government. Boston Consulting Group.

Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S.Z., Khawaji, A., Conway, D., 2016. Robust Multimodal Cognitive Load Measurement. Springer.

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P., Malhotra, S., 2018. Notes from the AI frontier: Applications and value of deep learning. McKinsey Global Institute.

Dawson, D., Schleiger, E., Horton, J., McLaughlin, J., Robinson, C., Quezada, G., Scowcroft, J., Hajkowicz, S., 2019. Artificial Intelligence - Australia's Ethics Framework. Data61 CSIRO, Australia.

Finn, R.L., Wright, D., Friedewald, M., 2013. Seven Types of Privacy, in: Gutwirth, S., Leenes, R., de Hert, P., Poullet, Y. (Eds.), European Data Protection: Coming of Age. Springer Netherlands, Dordrecht, pp. 3–32.

Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. Hum. Factors 46, 50–80.

Li, Z., Zhang, B., Wang, Yang, Chen, F., Taib, R., Whiffin, V., Wang, Yi, 2014. Water Pipe Condition Assessment: A Hierarchical Beta Process Approach for Sparse Incident Data. Mach. Learn. 95, 11–26.

McMahan, H.B., Andrew, G., Erlingsson, U., Chien, S., Mironov, I., Papernot, N., Kairouz, P., 2018. A General Approach to Adding Differential Privacy to Iterative Training Procedures. ArXiv181206210 Cs Stat.

van den Hoven, J., Blaauw, M., Pieters, W., Warnier, M., 2018. Privacy and Information Technology, in: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University.

Whiffin, V.S., Crawley, C., Wang, Y., Li, Z., Chen, F., 2013. Evaluation of machine learning for predicting critical main failure. Water Asset Manag. Int. 9, 17–20.

Yang, Q., Liu, Y., Chen, T., Tong, Y., 2019. Federated Machine Learning: Concept and Applications. ACM Trans. Intell. Syst. Technol. 10, 12:1–12:19.

Zhou, J., Chen, F. (Eds.), 2018. Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent, Human–Computer Interaction Series. Springer International Publishing.

Zhou, J., Sun, J., Wang, Y., Chen, F., 2017. Wrapping practical problems into a machine learning framework: using water pipe failure prediction as a case study. Int. J. Intell. Syst. Technol. Appl. 16, 191–207.