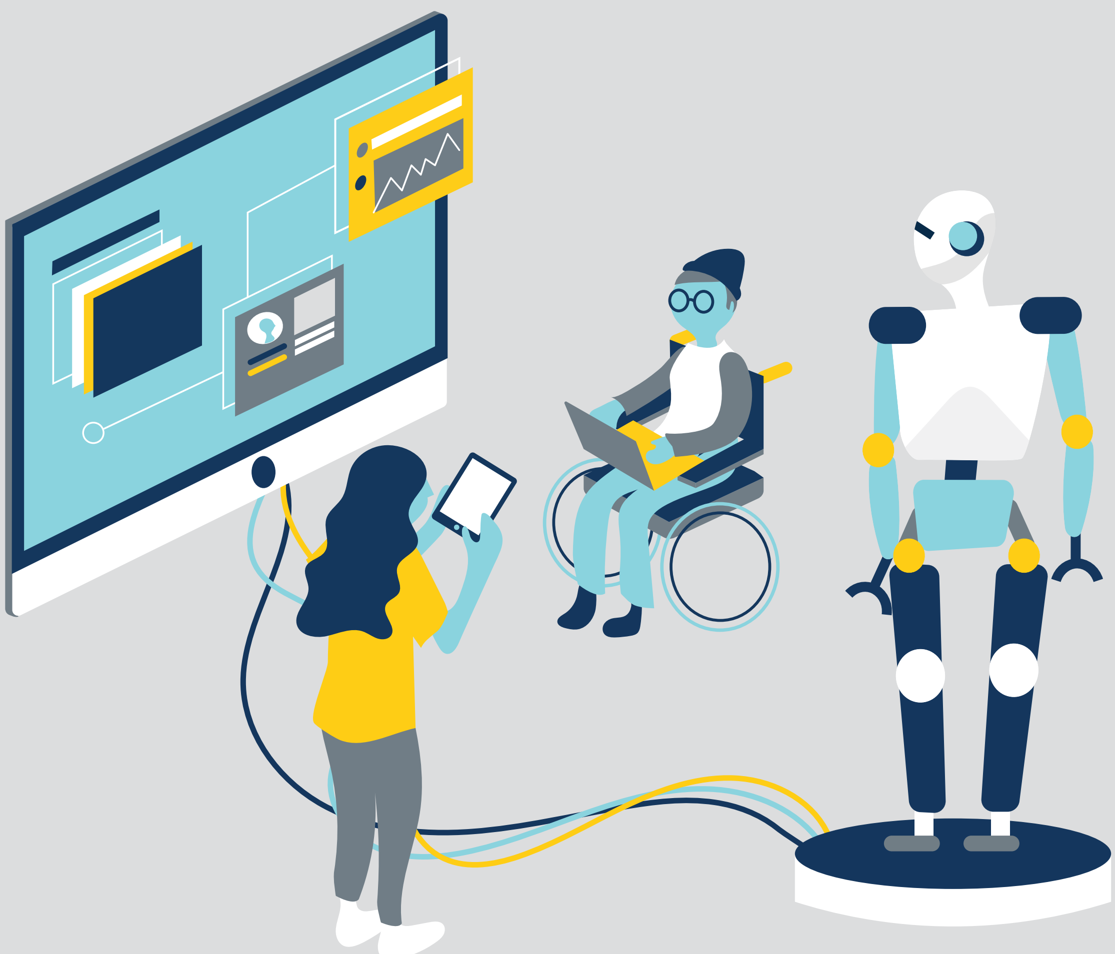
Intro to AI for Policymakers: Understanding the shift

March 2018



brookfield institute
for innovation + entrepreneurship

POLICY INNOVATION HUB

Ontario

TABLE OF CONTENTS

INTRODUCTION: THE STORY OF AI

Artificial Intelligence (AI) has benefited from a number of recent technological advances, from increases in processing power to decreases in battery costs, and the explosion of available data. Beyond technical research and commercial opportunities, this transformative technology has the potential to fundamentally alter our society, with enormous consequences. The current speed of development and adoption of AI poses challenges to policymakers seeking to create regulations, use AI applications to improve government operations and service delivery, and better understand and respond to the socioeconomic impacts. Innovations such as autonomous vehicles are already challenging traditional regulatory regimes and planning in transit, infrastructure, and public safety. Natural language processing and predictive algorithms are already operating inside everyday technology in our homes and workplaces.

To date, the majority of Canadian policy attention and investment has focused on building talent, supporting research, and encouraging the development of commercial AI applications. However, AI is experiencing a renewed prominence in public consciousness and public life, demanding policy responses beyond those focused on research and innovation. Cultural awareness of AI existed centuries prior to its technical feasibility, or the coining of the term in 1956. Following the initial burst of breakthroughs in hardware capabilities in the mid-20th Century, governments around the world commissioned research on the topic from academics and research institutes, recognizing the potential promise of AI to aid military operations and other government initiatives.

Relative decreases in political tensions and coinciding limitations in technical advances resulted in a significant withdrawal of government funding and attention during the 1980s. This shift opened the doors for private sector corporations that restimulated investment and research in recognition of potential commercial applications. Fueled by hype cycles known as "summers" of progress and "winters" of relative inactivity, this series of booms and busts privileged certain technological developments to emerge over others. The development and use of AI remains largely in the realm of academia and industry, with some increasing involvement of civilian government actors. But effective regulation and planning will require a working understanding of the technology and a clear line of sight into the work of researchers, entrepreneurs, and businesses on the frontlines of innovation. Understanding and responding to these shifts will be critical as we move into a world that is increasingly influenced by the adoption of these technologies.

The Brookfield Institute for Innovation + Entrepreneurship, with the Ontario government's Policy Innovation Hub, welcome you to a one-day conference on the public policy implications of AI on March 23rd, 2018, in Toronto, Ont. This event is among the first initiatives in Canada to engage policymakers alongside a cross-section of participants in the AI economy (technical experts, practitioners, academics, and entrepreneurs). The conference aims to:

+ Provide policymakers with a direct line of sight into the AI sector: myths, hype, the evolving state of technological advances, and potential applications.
+ Identify policy areas for further exploration including horizontal policy issues and areas with multi-jurisdictional implications.
+ Strengthen connections between policy and AI experts and between the public, private, and academic sectors.

GOVERNMENT RESPONSES TO AI

Today, Ontario is a recognized hub of AI expertise, both in Canada and internationally. Notable academics in the Toronto-Waterloo corridor and clusters such as the Vector Institute are promoting Canada as a leader in the development of AI, and Ontario institutions are graduating top talent. International companies have stationed their AI R+D efforts in the region (e.g., Uber, Thomson Reuters) and the Creative Destruction Lab at the University of Toronto says it has the largest cohort of AI startups in North America. The Government of Ontario has made strategic investments in research and education: providing $50 million to the Vector Institute; $30 million to boost the number of AI graduates, as part of its effort to attract foreign investments and companies, and to stimulate job creation; and $80 million to establish an Autonomous Vehicle Innovation Network[1] project, including an Autonomous Vehicle Pilot Study[2] which began in January 2016. Ontario has the opportunity to be a national leader in AI, keeping pace with this quickly evolving technology to anticipate and respond to future applications across key policy sectors, including government itself.

The federal government has committed significant financial support, including the $125-million Pan-Canadian Artificial Intelligence Strategy[3] (2017), led by the Canadian Institute for Advanced Research (CIFAR) in partnership with the Alberta Machine Intelligence Institute, the Vector Institute, the Montreal Institute for Learning Algorithms, and the recently announced SCALE.AI[4] Innovation Supercluster. There is also a small but growing number of pilot studies being carried out across Canada, including a suicide detection pilot study[5], led by the Public Health Agency of Canada, which aims to use data mining and machine learning to collect and analyze social media data and identify suicide indicators and risks. The Senate Committee on Social Affairs, Science and Technology has also studied the role of AI and robotics in the healthcare system.[6]

Internationally, some governments have taken proactive approaches to responding to and regulating AI. The United Kingdom has established an All-Party Parliamentary Group on Artificial Intelligence[7] in the House of Commons, as well as a Select Committee on Artificial Intelligence[8] within the House of Lords. The EU General Data Protection Regulation[9] (effective May 2018) will put some restrictions on when decisions that have legal impacts on individuals can be based solely on automated processing without human intervention or oversight. Under the Obama Administration, the United States developed a National Artificial Intelligence Research and Development Strategy[10] that addresses both economic and social considerations. The New York City Council also recently approved the "algorithm accountability bill,"[11] requiring the creation of a task force to audit and provide oversight of public-facing algorithmic decision-making systems.

AI 101

AI is a term used to refer to both the field of research and software capabilities. While the field of AI encompasses a wide range of techniques dating back to the 1950s, the current state of the art uses machine learning, deep learning, and reinforcement learning to identify patterns, produce insights, enhance knowledge-based work, and automate routine tasks.

Recent advances in AI have enabled us to process and analyze the growing amount of data being generated by human actions and behaviours. It has significantly enhanced our ability to spot patterns and generate insights and has automated mundane, rudimentary, and unsafe tasks in both production and domestic life.

TYPES OF AI

Narrow AI

Today, all AI is considered Narrow AI, also called applied or weak AI, because it is able to facilitate individual, repetitive tasks by learning from patterns found in data.

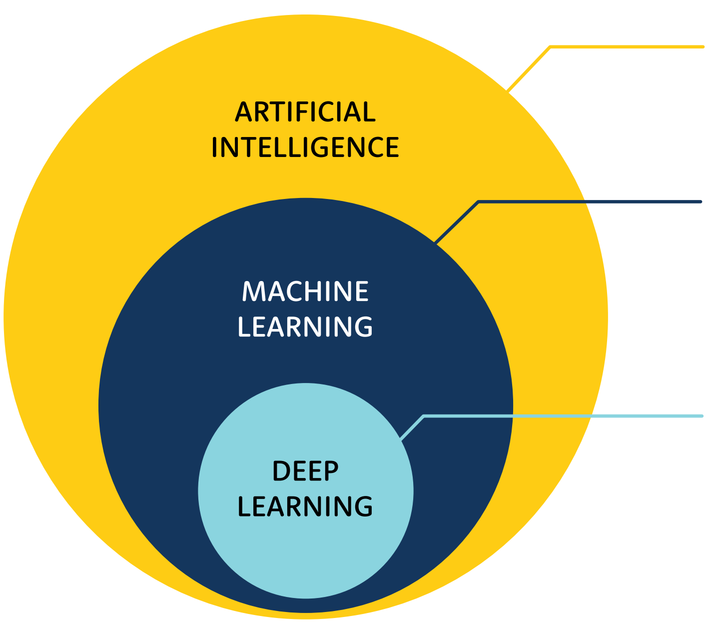
Artificial General Intelligence

Artificial General Intelligence, a notional concept that has not yet been developed in real life, refers to a single system “capable of performing all of the intellectual tasks that a human brain can”.[12] This includes reasoning, learning, and problem solving in complex and changing environments.

Artificial Superintelligence

Artificial Superintelligence is a hypothetical type of artificial intelligence “that surpasses human intellect and abilities in nearly all areas.”[13]

HIERARCHIES



AI relies on a variety of methods and techniques to learn and operate. This includes learning methods like machine learning and deep learning, as well as training approaches ranging from supervised to unsupervised learning.

MACHINE LEARNING is a technique that enables computer systems to learn and make predictions based on historical data. The machine learning process is powered by a machine learning algorithm, a function that is able to improve its performance over time by training itself using methods of data analysis and analytical modelling. Machine learning can be supervised, semi-supervised, or unsupervised.

DEEP LEARNING is a subfield of machine learning that consists of multiple cascading layers, modelled after the human nervous system (a practice referred to as neural coding), known as an artificial neural network. Deep learning architectures enable a computer system to train itself using historical data by recognizing patterns and making probabilistic inferences.

TRAINING DATA

Both machine learning and deep learning models rely on training data to learn relationships, increase the model's efficiency, and improve its ability to achieve the desired output. Training data refers to a data set that has been collected, prepared, and provided to the model for the purpose of teaching prior to active deployment. The quality, quantity, structure, and contents of training data are key determinants of how machine learning and deep learning models will function in a real environment.

TECHNIQUES

Supervised learning

Supervised learning is the task of teaching a machine learning algorithm by providing a labelled training data set, determining what input features will correspond with learned functions, and providing an example of correct outputs.

Reinforcement learning

Reinforcement learning involves the use of "rewards" and "punishments" in the form of functions. Programmers will reward a program when it learns a function or achieves the correct output efficiently. Reinforcement learning differs from supervised and semi-supervised learning in that "correct" inputs and outputs are never specified to the system. Reinforcement learning is often used when a system is operating in a dynamic and changing environment where systems have to take multiple routes to achieve the same output, or in cases where efficiency is valued over structure.

Semi-supervised learning

Semi-supervised learning is a method of training algorithms with a combination of labelled and unlabeled data. Semi-supervised learning contains two methods, *transductive* learning and *inductive learning*. Transductive learning refers to the ability of the system to infer labels upon the unlabeled data by learning from the labelled data it was provided. Inductive learning refers to the ability for the system to achieve the desired output without labelling the data.

Unsupervised learning

Unsupervised learning involves providing unlabelled input data from which a machine learning algorithm must structure data, discover patterns, classify inputs, learn functions, and produce outputs without external validation or support. Unsupervised learning can be employed to discover hidden patterns in data, typically those unrecognizable to, or difficult to discern by, humans.

PREDICTIVE ANALYTICS

What is it?

Predictive analytics is the use of data analytics to predict trends, behaviour patterns and outcomes.

How does it work?

Predictive analytics combines data mining, modeling, and mathematical analysis to produce visualizations of trends from large sets of data. While descriptive analytics analyzes what has already happened, predictive analytics builds on historical analysis and provides insights into likely future scenarios. Predictive analytics has been made possible due to advancements in computing power and data collection, more specifically, the ability to collect, store, and analyze big data. The volume, scale, speed, and accuracy of today's predictive analytics models far exceed those of previous generations of data analytics. Nevertheless, collecting, cleaning, labelling, and standardizing data as well as building more accurate predictive models is still a challenge.

Example

Today, predictive analytics is widely used in the corporate sector to optimize business processes, uncover statistical patterns, identify marketing targets based on predictive markers from previous successful sales, and predict and improve employee performance, among other applications. In the public sector, predictive analytics is being used to varying degrees of success to predict recidivism and risk of hospital readmission, reduce the risk of workplace accidents, pre-approve individuals for benefit programs, reduce the tax gap by identifying potentially fraudulent claims for investigation, and identify children in care who might be at risk of violence.

NATURAL LANGUAGE PROCESSING

What is it?

Natural language processing (NLP) is a functionality that enables machines to process, understand, and/or generate audio and textual speech.

How does it work?

NLP uses deep learning to analyze written and spoken text to generate responses in natural language or in the form of actions. Advances in machine learning and deep learning capabilities have enabled NLP systems to distinguish different voices and learn patterns in large data sets in order to better understand natural language. Current technical challenges include recognizing language variability, vocal ranges and accents.

Example

Today, NLP is deployed as part of commercially available AI assistants (e.g., Google Home, Amazon's Alexa, Apple's Siri) which offer voice-activated interactions with household computers and other technologies. In both the commercial and public sectors, NLP is currently used for service delivery, including automated customer support, language translation, text-based spam filters, and interactive dialogue. NLP has also been used to gauge sentiments on social media in order to trigger stock market trading and more accurately target advertising.

IMAGE RECOGNITION + COMPUTER VISION

What is it?

Image recognition is a system with the ability to identify specific features of digital images and videos. Computer vision is the extraction, analysis, and understanding of useful information from a single image or a sequence of images in order to achieve automated understanding from the visual input(s).

How does it work?

Image recognition employs machine learning and deep learning to identify and classify features within an image. Computer vision uses inputs from image recognition to classify information, make inferences, and take action. The introduction of deep learning (specifically deep convolutional neural nets) to image recognition has enabled the development of advanced learning models using big data and increased classification accuracy.

Example

Image recognition is used for fraud detection, facial recognition, and identifying illegal and/or graphic digital content. Computer vision is currently used for automatic inspection in manufacturing, assisting humans in identifying tasks (e.g., species identification), video-surveillance and motion detection (e.g., 'smart' traffic lights), modelling objections or environments, and as a core component in robot or autonomous vehicle navigation. Image recognition and computer vision capabilities have significantly increased in recent years, enabling computers to identify images better than humans in some cases.[14]

ROBOTICS

What is it?

Robotics is the design, construction, and operation of robots and also machines that embody software and are capable of autonomously performing specific tasks.

How does it work?

Robots include sensory features that collect and analyze information from the surrounding environment that is then used to generate reactions based on pre-programmed parameters encoded in the robots' software. Robotics are increasingly being designed to learn from training data and data collected on their environment in order to improve their responsiveness. Robotics has experienced accelerated growth in recent years due to improvements in computer vision, sensors, predictive analytics, and mechanical structures, as well as reductions in battery costs and sizes. Future developments will likely iterate the shape and design of robots in ways that enable a greater diversity and increasing complexity of tasks.

Example

Robots today are largely divided between those that perform domestic and industrial tasks. Domestic, task-specific robots include functionalities that enable them to perform household duties such as cleaning (e.g., domestic vacuum robots, such as the Roomba). Industrial, task-specific robots include functions that enable them to assist with processes like manufacturing or assembly line production.

PUTTING THE TECHNOLOGY INTO PRACTICE: AUTONOMOUS VEHICLES

An autonomous vehicle, also known as a self-driving car, is a vehicle capable of sensing its environment and successfully navigating with little to no human input. This technology can range from driver assistance in the form of autonomous parallel parking and lane guidance to fully autonomous driving. Autonomous vehicles employ **computer vision**, sensors, and **predictive analytics** in order to generate an understanding of their environment and predict possible paths to inform decisions by drivers/operators. While advanced autonomous vehicles are not yet available for commercial and public use, autonomous vehicles have the potential to be adopted as privately owned family cars, provide public transport (e.g., city buses) and commercial transport (e.g., intercity truck driving and Uber). Consumer acceptance of these technologies will depend on the development, implementation, and safety of applications of autonomous driving technology, as well as advancements in vision systems and technologies viewed to enhance human safety.

CROSS-CUTTING CHALLENGES OF AI

ETHICS

Ethics, in the context of AI, refers to whether, when, and how machines should make decisions, and what values should guide those decisions. The values embedded into AI systems will determine whether, and how, these systems will act in moral situations. Due to the increasing reliance on and trust in automated systems in contexts that may require them to make moral decisions, users should consider whether the values embedded in the code reflect their own. For government, the challenge is to ensure that machine decision making reflects not only public service values and ethics, regulations, and laws, but also broader social and moral norms. Ethical considerations underlie all the cross-cutting implications listed here. Current debates regarding autonomous vehicles revolve around whether, and how, they should have to decide between human lives in the event of an unavoidable collision. If they are designed to have the ability to make the choice, what values should be embedded in their programming to help guide those decisions?

Example

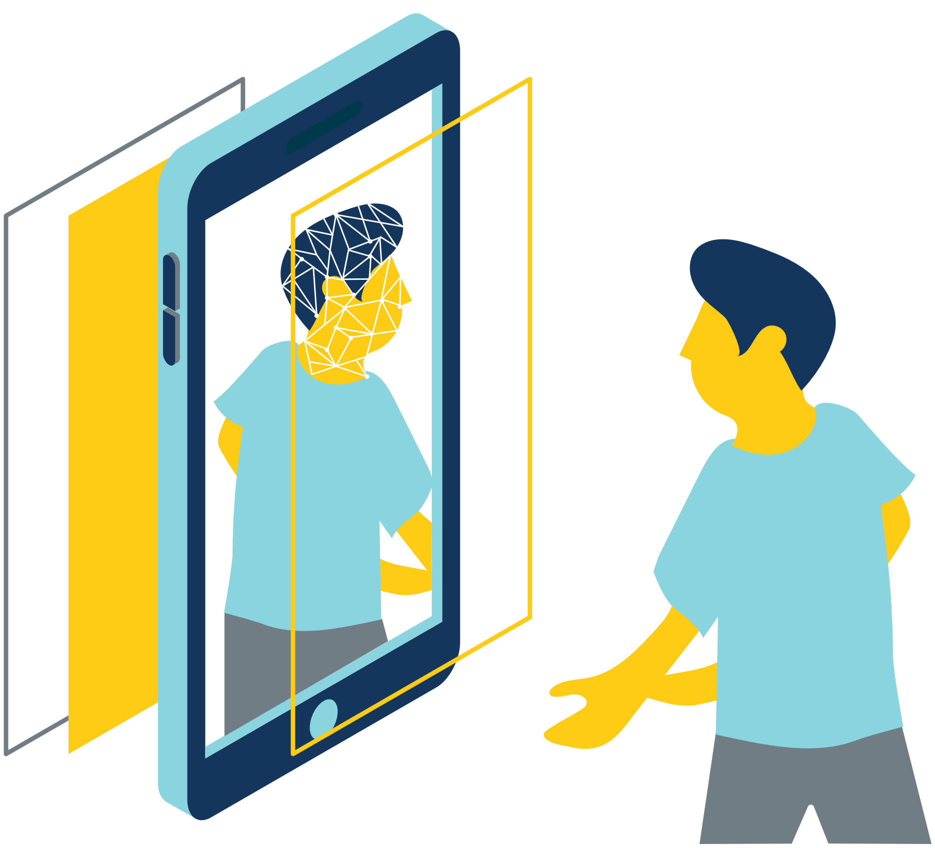
+ The use of drones in warfare is becoming increasingly common for remote, overseas conflicts. These systems employ computer vision, infrared imaging, and sensors to detect targets, assess a range of pre-programmed responses, and strike accordingly. Although the technology exists, autonomous military drones[15] pose significant ethical and legal challenges for the ease with which they can end human lives, and the difficulty of building decision-making systems that can mimic or improve on the human moral compass.

BIAS

Bias occurs when biases from the real-world are encoded and perpetuated within AI systems. Training data used to teach systems patterns, trends, or correct answers can introduce bias if it is incomplete, skewed, draws on non-representative datasets, excludes important information and/or contains existing social prejudices due to how it was collected or labeled. Design assumptions relating to potential user needs and contexts, can also introduce bias by emphasizing certain characteristics over others. Bias, when encoded into algorithms, is a form of what Oscar Gandy describes as "rational discrimination".[16] Distanced from class or racial hatred, rational discrimination[17] merely ignores and fails to correct existing societal biases.[18] As a result, algorithmic bias can introduce and/or reinforce disparities in society by impacting individuals' access to resources and services, the level of surveillance they experience, their treatment by police and government, and even their ability to be seen or heard in a technology-rich environment. When well-designed and implemented, high-tech tools for more precisely monitoring, tracking, sharing information, and predicting user need and eligibility could be used to improve program uptake, access, and impact. When used badly, they can serve as automated gatekeepers, perpetuating bias in the real world.

Example

+ A "crime prognostication system"[19] called PredPol,[20] used by some of the largest police forces in the United States, uses historical policing data to predict and map "crime hotspots," likely perpetrators, and victims. Because of the historic tendency to overpolice low-income and racialized communities and bias in arrests, PredPol is built on unfairly skewed racialized data.[21] When input into the algorithm, this data reinforces existing biases, predicting a disproportionate amount of crime to occur in these historically targeted communities, and further increasing police presence. Advocacy groups have also raised concerns that the use of PredPol could encourage officers to treat people who fit the PredPol profile as criminals without evidence.
+ A University of Virginia researcher found that the photo training data used to train image recognition systems had a heavy gender bias.[22] Two prominent research image collections (including one supported by Microsoft and Facebook) displayed gender bias in their depiction of activities such as cooking and sports, with images of shopping and washing linked to women, and images of coaching and shooting tied to men.[23] Machine learning software trained on these datasets didn't just mirror these biases – it amplified them, strengthening the association between gender and gender-biased activities.
+ Facial recognition software continues to struggle to "see" people with darker skin, due to a lack of diverse training data, and a lack of diversity in the teams developing the software. Two common benchmark data sets used to test facial recognition systems (IJB-A and Adience) are composed of 79.6 percent and 86.2 percent light-skinned faces, which means that they don't test the accuracy of the algorithms for non-light-skinned faces with the same rigour.[24] Recently, a study by MIT and Stanford University found that the facial recognition software produced by three major technology companies was significantly worse at identifying the gender of non-white people.[25]



SAFETY

In the context of AI, safety refers to the ability of artificially intelligent systems to operate without posing a risk or causing harm to humans. As AI systems are increasingly embedded in our lives, both visibly and invisibly, ensuring safety becomes more important. On one hand, safety may become jeopardized due to technical bugs or errors within the system, biased data, neglect of maintenance, lack of privacy, use in unintended contexts, or if the AI learns unsafe behaviour once it is operating. On the other, AI may help create and enforce parameters of safety by codifying a set of behaviours that are known to promote safe practices. This argument assumes these behaviours can be both explained in terms humans can agree on, and the encoded into the functioning of AI program.

Example

+ A 2015 Tesla Model S was involved in a fatal crash when the car's autopilot failed to detect and react to an oncoming transport truck turning left across its path.[26] Tesla's autopilot feature digitally controls steering and speed under the supervision of the driver. This is made possibly by a forward-facing camera that can read speed limit signs and watch lane markings, as well as numerous sensors that provide real-time situational awareness within a 16-foot radius in clear conditions. Neither the driver nor the car detected the oncoming transport fast enough to prevent the accident.
+ Darktrace's Enterprise Immune System,[27] an AI cybersecurity system developed by mathematicians and ex-British spies at the University of Cambridge, automatically detects and responds to cyber attacks using unsupervised machine learning to distinguish between normal and unusual behaviour in real time. It works by observing normal behaviours and identifying and neutralizing abnormalities that do not fit within the evolving pattern. In doing so, it can spot emerging trends that may have otherwise gone unnoticed and adapts quickly to new forms of threats. Darktrace is used by a growing number of Canadian organizations, including Energy+, Pizza Pizza, and DynaLIFE.[28]

PRIVACY

Privacy refers to the condition of being unobserved, and the confidentiality of an individual or group's personal and behavioural data. The collection, analysis, sharing, and use of personal data is becoming an increasingly important feature for AI systems. Personal data is often collected, used, and shared on an opt-out basis, or without the option to consent. The increasing ease with which intelligent systems collect and analyze personal data, as well as the ability for companies to share this information, has been critiqued for "challeng(ing) current understandings of privacy and strain(ing) the laws and regulations we have in place to protect personal information".[29]

Example

+ DeepMind Technologies, an Alphabet subsidiary, received 1.6 million patient records from the UK's National Health Service (NHS) intended for an app used for monitoring and diagnosing acute kidney injuries.[30] This information was shared without patient consent, and both DeepMind and the NHS faced public backlash for sharing personal health data for purposes other than what it was originally collected for.

+ Strava, a popular fitness tracking app that creates a geographical heatmap of a user's activity, accidentally exposed the sensitive locations of several United States military bases.[31] Strava automatically anonymizes and aggregates heat map data to share publicly as means for users to discover new routes to run or find exercise partners. While there is an option to use the app privately, the maps identified what appear to be military bases, and the data was able to be restructured in such as way that identified military personnel by name.

EXPLAINABILITY

Explainability refers to the ability of humans to interpret why a certain decision or action was made by an algorithm or a series of algorithms. Notionally, explainability can be achieved by understanding what and how data is used to produce certain outcomes. Current techniques within the field of AI, namely machine learning and deep learning, apply large volumes of data to non-linear models beyond the realm of human ability and comprehension, which can make the systems opaque and difficult to understand.[32] In cases where autonomous or intelligent systems are used to aid or make decisions that have real-life consequences for individuals or groups, the importance of understanding why and how a decision was made is important for liability, trust, and transparency.

Example

+ Most major credit rating companies are now using machine learning to help determine credit scores, expanding on existing actuarial science and statistical models and applying algorithms to identify segments of historical repayment and client demographic data to build their models on and how to weight it. The challenge for this sector is how to make the models, and the decisions they make, interpretable for regulatory purposes, and to ensure that the data they are based on is not biased.

+ Several American states, including California, New Jersey,[33] and Wisconsin, have integrated risk assessment algorithms in their judicial systems to aid in sentencing criminal suspects and set parole. These algorithms analyze data about the defendant, such as age, gender, criminal record history, in order to determine whether the defendant is likely to recommit a crime or appear for their court date. The algorithms are typically purchased from private businesses, not developed in house by government analysts. In Wisconsin v. Loomis, the defendant, Eric Loomis, was found guilty for his role in a drive-by shooting.[34] The trial judge used COMPAS, an algorithm-based risk assessment tool, to help determine the length of his sentence. Loomis challenged his sentence on the grounds that he was not allowed to assess the algorithm, but the state supreme court ruled that knowledge of the algorithm's output was sufficient transparency for the defendant.

ACCOUNTABILITY

Accountability in AI concerns who, or what, is held accountable when AI systems make decisions that impact human rights, civil liberties, and well-being. Many contemporary legal processes attribute the notion of responsibility to individual and corporate entities. However, these frameworks may lack an understanding of artificial autonomy and the relationships between designers, systems, and users that could prevent individuals seeking recourse for harms that AI has caused. For government, the adoption of AI into service delivery and operations poses questions around ensuring ministerial accountability and fulfillment of commitments on open government and service standards.

Example

+ RADAR, an automated news service, uses natural language processing to write journalistic articles autonomously.[35] This form of “robo-journalism” introduces issues relating to liability; for example, if an article written by this system was found to be defamatory, who (or what) would be held responsible?
+ Health insurance companies are adopting machine learning models to predict insurance premiums more accurately using an individual’s personal data.[36] This raises concerns regarding accountability for the miscalculation of premiums or denial of insurance that may result in medical and financial hardship.
+ Companies (e.g., car manufacturers) currently receive different legal protections than individuals. If a person was struck by an autonomous vehicle, who would be held accountable, and through which legal mechanism (e.g., lawsuit, fines, penalty)? This scenario would currently be treated much differently than if that same person was struck by a car being driven by a human.

GLOSSARY

Agency

The ability of an artificial system to make decisions, typically within a specified context, without human intervention.

Algorithm

A sequence of instructions, rules, and calculations executed by a computer in a particular order to yield a result, typically an answer to a specified problem. Algorithms can be used in combination with other algorithms to solve complex problems.

Applied Artificial Intelligence

"The use of AI to enhance and extend" the capabilities of software applications.[37]

Artificial General Intelligence (AGI), alternatively General Artificial Intelligence (GAI)

An AI system capable of performing beyond problem/task-specific, domain-dependent solutions toward general-purpose systems comparable to human-level intelligence (including but not limited to problem solving, task completion, context-specific knowledge, modes of inquiry, etc). AGI is a notional concept, as these capabilities cannot be supported by current software or combined software and hardware capabilities.

Artificial Intelligence (AI)

AI as technology: Computer programs capable of behaviour commonly thought to require intelligence.[38]

AI as a field/discipline: The study and development of artificially intelligent systems.

Artificial Narrow Intelligence (ANI), also known as Weak AI

All current AI is narrow, meaning it can only do what it is designed to do. Narrow AI is:

a. Domain-dependent and designed for problem/task-specific solutions.
b. Not self-aware

Artificial Neural Network (ANN)

A processing device (algorithms or actual hardware) modeled after the neuronal structure of the cerebral cortex in the mammalian brain, but on significantly smaller scales.[39]

Artificial Superintelligence

A notional AI system containing intellectual capacity beyond "the best human brains in practically every field, including scientific creativity, general wisdom and social skills."[40]

Autonomous

An artificial agent that has the capacity to learn and/or function without external interventions.

Big Data

A dataset with a size beyond the processing capability of a typical database for the purpose of data capture, storage, management, and analysis.[41]

GLOSSARY

Chatbot

An artificial system designed to function as a participant in text-based conversations over the internet.

Computer Vision

Computer vision applies machine learning to automatically extract, analyze and understand high dimensional data from a single image or video or a sequence of images in order to describe and/or generate decisions.

Data Analytics

The analysis of large volume or high-velocity data using advanced analytic techniques.[42, 43]

Deep Learning

"A type of machine learning that trains a computer to perform human-like tasks" by setting up basic parameters about the data. It "trains the computer to learn on its own by recognizing patterns using many layers of processing."[44]

Disembodied AI

Intelligence software existing invisibly, integrated into a variety of platforms.

Embodied AI (also referred to as Cyber-Physical or Robotics)

Intelligent software embedded within physical hardware.

Human-in-the-Loop

The existence of, or requirement for, human approval, assistance, or intervention in order for a system to perform a specific task. Many algorithms still rely on having a human in the loop.

Machine Learning

A technique that enables computer systems to learn and make predictions based on historical data.

Natural Language Processing

"Makes it possible for machines to process and understand audio and text data." Function can include "tasks like translation, interactive dialogue, and sentiment analysis."[45]

Predictive Analytics

The use of data analytics and machine learning to extract information and learn patterns from data in order to uncover past, present, and future events.

Reinforcement Learning

A type of machine learning that "allows machines and software agents to automatically determine the ideal behaviour within a specific context, in order to maximize its performance."[46] Humans supervise and provide reward feedback when the agent performs correctly.

GLOSSARY

Robotics

The use of hardware and software to do a novel task

Sentiment Analysis

The use of use of AI technology to collect, quantify, and analyze online data to determine the affective and emotional states of individuals.

Semi-Structured Data

Data which does not conform to formal data standards or models associated with relational databases, but still contains semantic/lexical tags and/or markers to enforce order.

Structured Data

Data presented and classified in a standardized format, making it easy to organise, search, and analyze.

Supervised Learning

The process of training an algorithm using labelled training data. Labelled data refers to data which has already been categorized, tagged, and/or weighted.

Training Data

Data used to train machine learning and deep learning algorithms. Training data can be structured, semi-structured, or unstructured.

Unstructured Data

Data presented without a pre-defined model or organisational standard. Contains data of various types (text, numbers/ qualitative, quantitative) and is often difficult to search or analyze.

Unsupervised Learning

The process of training an algorithm with unlabelled training data. Unlabelled data refers to data which is raw (uncategorized). This forces the algorithm to learn by creating its own categories for the data it is given.

Use Case

The application of a technology to serve a specific need in a specific context.

Virtual Reality

"...the computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors."[47]

ENDNOTES

1. AVIN. (2017). Autonomous Vehicle Innovation Network (AVIN). Retrieved from: https://www.avinhub.ca/.

2. Ontario Ministry of Transportation. (2015). Automated Vehicles – Driving Innovation in Ontario. Retrieved from: http://www.mto.gov.on.ca/english/vehicles/automated-vehicles.shtml.

3. CIFAR. (2017). Pan-Canadian Artificial Intelligence Strategy Overview. Retrieved from: https://www.cifar.ca/assets/pan-canadian-artificial-intelligence-strategy-overview/.

4. Newswire.ca. (2018). SCALE.AI, the AI-Powered Supply Chain Supercluster, to Receive Funding from Innovation Superclusters Initiative. Retrieved from: https://www.newswire.ca/news-releases/scaleai-the-ai-powered-supply-chain-supercluster-to-receive-funding-from-innovation-superclusters-initiative-674185353.html.

5. Buyandsell.gc.ca. (2017). Artificial Intelligence (AI) pilot project for surveillance of suicide-related behaviours using social media (Advance Contract Award Notice). Retrieved from: https://buyandsell.gc.ca/cds/public/2017/12/22/a6835bc7e03bc761b2d730ebf18fc33e/1000196416_ai_acan_english.pdf.

6. Senate of Canada. (2017). Canadian healthcare system must brace for a technological revolution. Retrieved from: https://sencanada.ca/en/newsroom/soci-challenge-ahead/.

7. APPG. (2018). APPG AI | Party Parliamentary Group on AI. Retrieved from: http://www.appg-ai.org/.

8. www.parliament.uk. (2018). Select Committee on Artificial Intelligence. Retrieved from: http://www.parliament.uk/ai-committee.

9. Privacy-Regulation.eu. (n.d.). General Data Protection Regulation. Article 22. Retrieved from: http://www.privacy-regulation.eu/en/22.htm.

10. Networking and Information Technology Research and Development Program. (2016). The National Artificial Intelligence Research And Development Strategic Plan. Retrieved from: https://www.nitrd.gov/news/national_ai_rd_strategic_plan.aspx.

11. New York City Council. (2017). Automated decision systems used by agencies. Law 2018/049. Retrieved from: http://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0.

12. Centre for Public Impact (2017). Destination unknown: Exploring the impact of Artificial Intelligence on Government. Artificial Intelligence and the future of government. Retrieved from: https://publicimpact.blob.core.windows.net/production/2017/09/Destination-Unknown-AI-and-government.pdf.

13. Centre for Public Impact (2017). Destination unknown: Exploring the impact of Artificial Intelligence on Government. Artificial Intelligence and the future of government. Retrieved from: https://publicimpact.blob.core.windows.net/production/2017/09/Destination-Unknown-AI-and-government.pdf.

14. Johnson, R. (2015). Microsoft, Google Beat Humans at Image Recognition. EE Times. Retrieved from: https://www.eetimes.com/document.asp?doc_id=1325712.

15. North Atlantic Treaty Organization. (2017). Autonomous military drones: no longer science fiction. Retrieved from: https://www.nato.int/docu/review/2017/Also-in-2017/autonomous-military-drones-no-longer-science-fiction/EN/index.htm.

16. Gandy, O. (2009). Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage. 1st ed. Oxon: Routledge.

17. Gandy, O. (2009). Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage. 1st ed. Oxon: Routledge.

18. Eubanks, V. (2018). Automating Inequality: How High Tech Tools Profile, Punish, and Profile the Poor. St. Martin's Press.

19. Tashea, J. (2017). Courts Are Using AI to Sentence Criminals. That Must Stop Now. WIRED. Retrieved from: https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/.

20. PredPol. (2017). Predictive Policing Software. Retrieved from: http://www.predpol.com/.

21. Lum, K., and Isaac, W. (2016). To predict and serve? Significance, 13(5), pp.14-19.

22. Simonite, T. (2017). Machines taught by photos learn a sexist view of women. WIRED. Retrieved from: https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/.

23. Zhao, J., Wang, T., Yatskar, M., Ordonez, V., and Chang, K. (2017). Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints. University of Washington. Retrieved from: https://arxiv.org/abs/1707.09457.

24. Buolamwini, J., and Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Conference on Fairness, Accountability and Transparency. Proceedings of Machine Learning Research, pp.1-15.

25. Tucker, I. (2017). 'A white mask worked better': why algorithms are not colour blind. [online] The Guardian. Retrieved from: https://www.theguardian.com/technology/2017/may/28/joy-buolamwini-when-algorithms-are-racist-facial-recognition-bias.

26. Fung, B. (2016). The technology behind the Tesla crash, explained. Washington Post. Available at: https://www.washingtonpost.com/news/the-switch/wp/2016/07/01/the-technology-behind-the-tesla-crash-explained/?utm_term=.e9616b217e54.

27. Darktrace. (2018). Darktrace. Retrieved from: https://www.darktrace.com/.

28. The IT Nerd. (2018). Darktrace Safeguards Data For Heritage Education Funds. Retrieved from: https://itnerd.blog/2018/02/06/darktrace-safeguards-data-for-heritage-education-funds/.

29. Campolo, A., Sanfilippo, M., Whittaker, M. and Crawford, K. (2017). AI Now 2017 Report. AI Now. Retrieved from: https://assets.contentful.com/8wprhhvnpfc0/1A9c3ZTCZa2KEYM64Wsc2a/8636557c5fb14f2b74b2be64c3ce0c78/_AI_Now_Institute_2017_Report_.pdf

30. McGoogan, C. (2017). NHS illegally handed Google firm 1.6m patient records, UK data watchdog finds. [online] The Telegraph. Available at: http://www.telegraph.co.uk/technology/2017/07/03/googles-deepmind-nhs-misused-patient-data-trial-watchdog-says/.

31. Russell, J. (2018). Fitness app Strava exposes the location of military bases. TechCrunch. Retrieved from: https://techcrunch.com/2018/01/28/strava-exposes-military-bases/.

32. Gunning, D. (2016). Explainable Artificial Intelligence. DARPA. Retrieved from: https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf.

33. Lapowsky, I. (2018). Crime-Predicting Algorithms May Not Beat Untrained Humans. WIRED. Retrieved from: https://www.wired.com/story/crime-predicting-algorithms-may-not-outperform-untrained-humans/.

34. Tashea, J. (2017). Courts Are Using AI to Sentence Criminals. That Must Stop Now. WIRED. Retrieved from: https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/.

35. Ponsford, D. (2017). First robot-written stories from Press Association make it into print in 'world-first' for journalism industry. Press Gazette. Retrieved from: http://www.pressgazette.co.uk/first-robot-written-stories-from-press-association-make-it-into-print-in-world-first-for-journalism-industry/.

36. O'Neil, C. (2017). Big Data Is Coming to Take Your Health Insurance. Bloomberg. Retrieved from: https://www.bloomberg.com/view/articles/2017-08-04/big-data-is-coming-to-take-your-health-insurance.

37. Georgian Partners. (2017). An Overview of Applied Artificial Intelligence. Retrieved from: https://georgianpartners.com/investment-thesis-areas/applied-artificial-intelligence/.

38. National Science and Technology Council (2016). Preparing for the Future of Artificial Intelligence. Committee on Technology. Executive Office of the President. Retrieved from: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

39. University of Wisconsin-Madison. (n.d.). A Basic Introduction To Neural Networks. Retrieved from: http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html.

40. Bostrom, N. (2013). Superintelligence: Paths, Dangers, Strategies. Oxford University Press.

41. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. Retrieved from: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.ashx.

42. Informatica. (2018). What is Data Analytics? Informatica Canada. Retrieved from: https://www.informatica.com/ca/services-and-training/glossary-of-terms/data-analytics-definition.html#fbid=sHMgUo4CDEj.

43. IBM. (2018). Big Data Analytics. [online] Available at: https://www.ibm.com/analytics/hadoop/big-data-analytics.

44. SAS. (n.d.). Deep Learning: What it is and why it matters. Retrieved from: https://www.sas.com/en_us/insights/analytics/deep-learning.html.

45. Centre for Public Impact (2017). Destination unknown: Exploring the impact of Artificial Intelligence on Government. Artificial Intelligence and the future of government. Retrieved from: https://publicimpact.blob.core.windows.net/production/2017/09/Destination-Unknown-AI-and-government.pdf.

46. Reinforcement Learning Warehouse. (n.d.). Reinforcement Learning. Retrieved from: http://reinforcementlearning.ai-depot.com/.

47. Oxford Dictionaries. (2018). Virtual reality. Retrieved from: https://en.oxforddictionaries.com/definition/virtual_reality.

PRESENTED BY

brookfield institute
for innovation + entrepreneurship

POLICY INNOVATION HUB

Ontario

IN PARTNERSHIP WITH

VECTOR INSTITUTE

CIFAR
CANADIAN INSTITUTE FOR ADVANCED RESEARCH

Centre for Public Impact
A BCG FOUNDATION