Section IV
Ethics of AI

Table of Contents

Introduction

The increasing development and deployment of AI technologies is already impacting all aspects of our economy and society. Simply put, AI is changing everything around us and the speed this change is unravelling is exponential.

Consequently, policy-makers worldwide have a new challenge to urgently address: how do we seize the benefits these AI technologies offer while also protecting the society and individuals from socio-ethical risks?

The economic opportunities - promising to increase productivity and drive down inefficiencies - are revolutionary; however, at the same time, there are a good number of ethical issues we must confront. Will the new technologies be fair and transparent? Will the benefits be distributed to all? Will they reinforce existing inequalities? These are only but a few complicated problems we are now left to resolve.

Ethics has always been an interesting area for humanity, with individuals from different backgrounds contemplating and debating what moral principles can be established to govern an individual's behaviour. With the introduction of AI technologies in society, we are seeing many conversations around ethics now resurfacing while new ones are also now starting to form.

Ethics can be defined as well-founded standards of right and wrong. They are used to prescribe what humans ought to do in terms of rights, obligations, benefits to society, fairness or specific virtues.

In regards to AI, the most important ethical question UK is trying to address is: **even if technologically and/or legally an AI system can be developed and deployed, *should it*?**

The answer to this question relies largely on a country's value and culture.

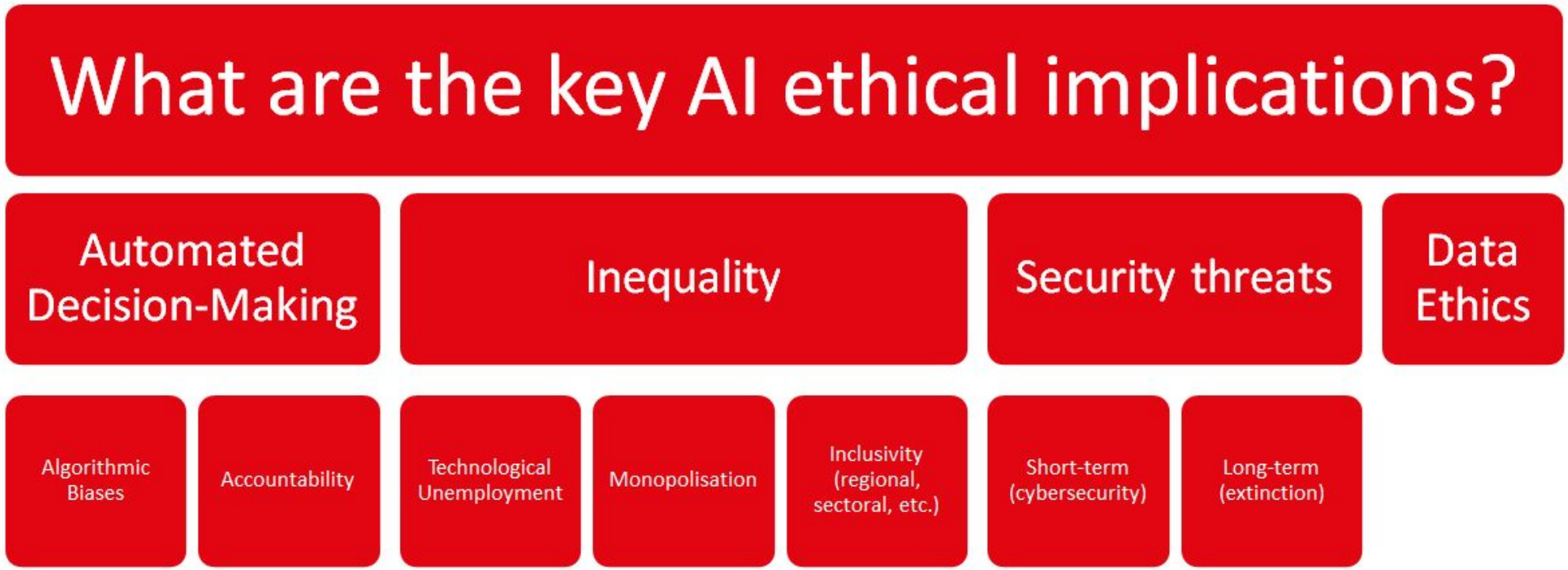
AI ethics, along with AI governance and AI regulation, will shape the 'rules of the game' around how these AI technologies will be developed, used, sold, and managed.

Source: https://www.ibe.org.uk/userassets/briefings/ibe_briefing_58_business_ethics_and_artificial_intelligence.pdf

Main problems

There are many ethical implications that are connected to AI technologies. According to the All Party Parliamentary Group on AI, these implications can be grouped in four key categories:

1. Automated decision-making: When a machine is wholly or partly responsible for a decision, this establishes a new set of ethical concerns for mankind. This is particularly the case when the decisions are around important matters that have to do with one's education, security, and/or wellbeing. The issue around automated decision making can further be broken down into topics. The first has to do with algorithmic biases. As datasets are often reflections of ourselves, this implies that the implicit biases embedded within humans are also reflected within the



Source: APPG AI PowerPoint Presentation

Main problems

data. As a consequence, opaque and potentially biased mathematical models are remaking our lives and making life-critical decisions about us. Algorithms that may conceal hidden biases are already routinely used to make vital financial and legal decisions. Proprietary algorithms are used to decide, for instance, who gets a job interview, who gets granted parole, and who gets a loan.

The second issue around automated decision-making is linked closely to accountability. Who is responsible for a decision if it goes bad? If a decision is based on the output of an algorithm than who is responsible if it ends up ultimately being a harmful or wrong decision? We need new accountability structures when AI is involved in decision-making to allocate responsibility accordingly.

2. Inequality: The second ethical challenge has to do with inequality. There are fears that AI technologies are increasing inequality gaps worldwide rather than shrinking them. Of course, the impact of AI on the labour market might mean higher levels of unemployment. Modern-era automation means even professional 'white-collar' jobs are in risk. Furthermore, because the benefits of AI are so huge, this means that those with the advantage of having the right data now will win in the long run. The risk of monopolisation is therefore becoming increasingly likely. Lastly, as the benefits of AI might not be distributed across various social groups fairly. We need policy and regulation to make sure all parts of society - regardless of geography, age, race, etc. - receive the gains.

3. Security threats: The concerns around security risks are also high in the agendas of policymakers worldwide. Short-term concerns include what cyber-threats and the increasing vulnerability of systems to internal and external malicious use. There can also be security consequences that are unintended but ultimately end up damaging humanity and/or the wellbeing of individuals.

4. Data ethics: Lastly, as it is hard to separate AI from data, there are several ethical implications related to how data is collected, stored, managed, and used. Data ownership, the sense of personal data, data monopolisation, and data privacy are now some of the most common worries for legislators and regulators.

Source: APPG AI PowerPoint Presentation

AI in the UK: ready, willing and able?

The UK is in a strong position to be a world leader in the development of artificial intelligence (AI). A report by the House of Lords Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able?*, published on 16th April 2018 concludes that the UK has a unique opportunity to shape AI positively for the public's benefit and to lead the international community in AI's ethical development, rather than passively accept its consequences.


HOUSE OF LORDS

Select Committee on Artificial Intelligence

Report of Session 2017–19

AI in the UK: ready, willing and able?

Ordered to be printed 13 March 2018 and published 16 April 2018

Published by the Authority of the House of Lords

HL Paper 100

The Chairman of the Committee, Lord Clement-Jones, said:

"*The UK contains leading AI companies, a dynamic academic research culture, and a vigorous start-up ecosystem as well as a host of legal, ethical, financial and linguistic strengths. We should make the most of this environment, but it is essential that ethics take centre stage in AI's development and use.*"

One of the recommendations of the report is for a cross-sector AI Code to be established, which can be adopted nationally, and internationally. The Committee's suggested five principles for such a code are:

1. Artificial intelligence should be developed for the common good and benefit of humanity.
2. Artificial intelligence should operate on principles of intelligibility and fairness.
3. Artificial intelligence should not be used to diminish the data rights or privacy of individuals, families or communities.
4. All citizens should have the right to be educated to enable them to flourish mentally, emotionally and economically alongside artificial intelligence.
5. The autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence.

At earlier stages of education, children need to be adequately prepared for working with, and using, AI. The ethical design and use of AI should become an integral part of the curriculum.

Source: https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf
https://www.parliament.uk/business/committees/committees-a-z/lords-select/ai-committee/news-parliament-2017/ai-report-published/

Chapter I:

AI and Data Security and Protection

Highlights

- Data security and protection is, on the one hand, a promising use-case for AI, and on the other hand a pressing issue relating to AI safety, ethics and governance (an area which the UK is taking a strong international position on)

- The application of AI to data security can be broken down into two broad categories: (1) built-in machine learning, where firms build machine-learning based security solutions (e.g. machine learning algorithms to automatically detect and prioritize security alerts and behavioural anomalies) directly into their security protocols, and (2) machine learning toolkits, consisting of customized machine learning solutions built by security experts that are used supplementrily to a firm's core security solutions

- The UK House of Lords Report on Artificial Intelligence presents an "AI Code" consisting of five core guiding principles in order to ensure that AI is used in such a way as to prevent the monopolization of data by large technology companies operating in the UK.

- The European Union has also released guidelines for data protection through their General Data Protection Regulation (GDPR) framework, which sets guidelines on the amount of data companies can collect and keep, the applications that they can use data for, and requires that companies can alter or delete data on request, inform people on the use of their data, and explain the login behind decision making processes that use data to make automated decisions about people

Two approaches of AI application in Security

APPROACH 1 - BUILT-IN MACHINE LEARNING

The first is to procure solutions to solve predefined cyber problems with machine learning capabilities built in. Those solutions are typically straightforward to deploy as a given set of pre-defined input is required to make the machine learning models work. An example of where this is used today via a built-in solution without needing data scientists is User Behavior Analytics. These solutions free up a security team by analysing machine data, correlating user and system activity with different algorithms and machine learning models, prioritising security alerts and creating anomalies based on all of them. It is also possible with Machine Learning to 'stitch' an attack kill chain of different anomalies together to present a security analysts the full picture of a potential incident. Doing this manually can be very time consuming or expensive as highly skilled incident investigators need to be employed who already know what to look for.

APPROACH 2 - MACHINE LEARNING TOOLKITS

The other approach is for an organisation to hire and employ data scientists with a security background. These skills are rare, but security centric Data Scientists might be able to focus on specialist security use cases such as fraud or create their own customised machine learning solutions. Data Scientists need to innovate quickly. They need to capture data quickly to evaluate new developed models based on features they might want to validate. Most of their time (60%) is spent on data validation rather than testing and working on algorithms or new use cases to solve. By centralising all machine data in a machine data platform and utilising Machine Learning Toolkits data scientists can focus on delivering insights rather than less beneficial tasks such as collecting data, transforming it and finding out it's outdated or incomplete.

Source: https://www.techuk.org/insights/opinions/item/12926-cyber-security-and-ai-opportunities-and-challenges

Data Protection principles and rights

Everyone responsible for using personal data has to follow strict rules called ‘data protection principles’. They must make sure the information is:

- used fairly, lawfully and transparently
- used for specified, explicit purposes
- used in a way that is adequate, relevant and limited to only what is necessary
- accurate and, where necessary, kept up to date
- kept for no longer than is necessary
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage

Under the Data Protection Act 2018, you have the right to find out what information the government and other organisations store about you. These include the right to:

- be informed about how your data is being used
- access personal data
- have incorrect data updated
- have data erased
- stop or restrict the processing of your data
- data portability (allowing you to get and reuse your data for different services)
- object to how your data is processed in certain circumstances

You also have rights when an organisation is using your personal data for:

- automated decision-making processes (without human involvement)
- profiling, for example to predict your behaviour or interests

Source: https://www.gov.uk/data-protection

UK approach to Data Protection and AI

One of the recommendations of the report is for an AI Code to be established based on five guiding principles:

- *Artificial intelligence should be developed for the common good and benefit of humanity.*
- *Artificial intelligence should operate on principles of intelligibility and fairness.*
- *Artificial intelligence should not be used to diminish the data rights or privacy of individuals, families or communities.*
- *All citizens have the right to be educated to enable them to flourish mentally, emotionally and economically alongside artificial intelligence.*
- *The autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence.*

Source: https://dataethics.eu/en/new-uk-ai-report-warns-data-monopolies/

UK approach to Data Protection and AI

Organisations should carefully consider whether the big data analytics to be undertaken actually requires the processing of personal data. Often, this will not be the case; in such circumstances organisations should use appropriate techniques to anonymise the personal data in their dataset(s) before analysis;

Organisations should be transparent about their processing of personal data by using a combination of innovative approaches in order to provide meaningful privacy notices at appropriate stages throughout a big data project. This may include the use of icons, just-in-time notifications and layered privacy notices;

Organisations should embed a privacy impact assessment framework into their big data processing activities to help identify privacy risks and assess the necessity and proportionality of a given project. The privacy impact assessment should involve input from all relevant parties including data analysts, compliance officers, board members and the public;

Organisations should adopt a privacy by design approach in the development and application of their big data analytics. This should include implementing technical and organisational measures to address matters including data security, data minimisation and data segregation;

Organisations should develop ethical principles to help reinforce key data protection principles. Employees in smaller organisations should use these principles as a reference point when working on big data projects. Larger organisations should create ethics boards to help scrutinise projects and assess complex issues arising from big data analytics;

Organisations should implement innovative techniques to develop auditable machine learning algorithms. Internal and external audits should be undertaken with a view to explaining the rationale behind algorithmic decisions and checking for bias, discrimination and errors.

Source: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf

Europe's General Data Protection Regulation (GDPR)

The privacy law, which came into effect across the EU in May 2018, has several elements that will make life very difficult for companies building machine learning systems, according to a leading Internet law academic.

Machine learning—the basis of AI—involves algorithms that progressively improve themselves. They do this by feasting on data. The more they consume, the better they get at spotting patterns: speech patterns that make it easier for a bot to sound like a human; visual patterns that help an autonomous car system recognize objects on the road; customer behavior patterns that train a bank's AI systems to better spot fraud.

Europe's new General Data Protection Regulation (GDPR) says:

- When companies collect personal data, they have to say what it will be used for, and not use it for anything else.
- Companies are supposed to minimize the amount of data they collect and keep, limiting it to what is strictly necessary for those purposes—they're supposed to put limits on how long they hold that data, too.
- Companies have to be able to tell people what data they hold on them, and what's being done with it.
- Companies should be able to alter or get rid of people's personal data if requested.
- If personal data is used to make automated decisions about people, companies must be able to explain the logic behind the decision-making process.

Algorithmic transparency means you can see how the decision is reached, but you can't with [machine-learning] systems because it's not rule-based software. The issue becomes even more fraught where companies use people's data to infer things about them—sensitive personal data, which includes things like sexuality and political and religious beliefs, gets even stronger protections under the GDPR.

The GDPR gives companies other legal justifications that they can use to process people's data, such as the need to use that data in order to provide core services. But where this kind of sensitive data is concerned, people have to give their explicit consent to its processing.

Source: http://fortune.com/2018/05/25/ai-machine-learning-privacy-gdpr/

The Impact of the GDPR on AI

1. Requiring companies to manually review significant algorithmic decisions raises the overall cost of AI.
2. The right to explanation could reduce AI accuracy.
3. The right to erasure could damage AI systems.
4. The prohibition on repurposing data will constrain AI innovation.
5. Vague rules could deter companies from using de-identified data.
6. The GDPR's complexity will raise the cost of using AI.
7. The GDPR increases regulatory risks for firms using AI.
8. Data-localization requirements raise AI costs.
9. Data portability will stimulate AI competition, albeit at a cost.

Source: http://www2.datainnovation.org/2018-impact-gdpr-ai.pdf

Chapter II:

Transforming Labour Market and Growing Skills Gaps

Highlights

- Neutralizing skills gaps in AI and IT technologies is a core ambition of the UK Government in order to promote the accelerated and informed adoption of AI solutions in both the public and private sectors and to combat the “brain drain” facing the nation (i.e. the lack of a sufficient number of AI specialists to meet the demands of their new national AI industrial strategy).

- Part of the funds injected into the AI Industry through the Uk AI Sector Deal is earmerked for the support of 1000 new UK AI PhDs and 8,000 new computer science teachers by the year 2025, as well Government-funded training on how to use AI and Robotics by NHS staff In order to help deter skills gaps due to advances in AI.

- A Deloitte survey found that less than half of tech company executives feel that they have the skills to lead their organisation in the digital economy, that only 16% believe their talent tool has the required expertise to deliver on their company's digital strategy, and that only 12% believe that UK graduates have the necessary skills to succeed in the digital economy.

- The issue being faced by UK leaders is not whether AI will cause changes in the labour market, but what changes need to be made in order to ensure that automation creates more jobs than it displaces, and that the workforce is equipped with the skills necessary to weather such changes.

- A PwC report argues that AI will create more jobs (7.2m) than it will displace (7.0m), that 20% of jobs will be automated over the next 20 years, and that AI will leave no single sector unaffected.

Myths about AI

From health care to transportation to national security, AI has the potential to improve lives. But it comes with fears about economic disruption and a brewing “AI arms race”.

According to By Bill LaPlante and Katharyn White from The Washington Post there are the following myths:

- **You can differentiate between a machine and a human.**
- Garbled sentences and ridiculous responses of Alexa or Siri or Cortana often make clear just how poorly machines mimic human capabilities — or even, sometimes, how they process information. Garry Kasparov told TechCrunch in 2017 that “*Machines don’t have understanding, they don’t recognize strategical patterns. Machines don’t have purpose.*”
- **AI will automate the economy and put people out of work.**
  In transforming work AI will create new jobs. Historically, technological change has initially diminished, but then later boosted, employment and living standards by enabling new industries and sectors to emerge. A report from PricewaterhouseCoopers argued that AI would create slightly more jobs (7.2m) than it displaced (7m) by boosting economic growth.
- **AI can remove human bias from decision-making.**
  In one example that shows AI’s vulnerability to bias, ProPublica found that a program intended to play a key role in criminal justice decisions from bail to sentencing was almost twice as likely to rate black defendants as probable repeat offenders than white defendants. The program also incorrectly rated white defendants as low-risk more often than blacks. In another example, a 2015 Carnegie Mellon University experiment found that far fewer women were being shown online ads for jobs paying more than $200,000 than were men.
- **Artificial intelligence is a threat to mankind.**
  The truth is we simply don’t know where AI will lead us. The more pressing concern might not be that AI is a risk to us, but that we’re a risk to ourselves if we don’t exercise caution in how we push ahead with our AI experiments. A 2017 Rand Corp. report, for example, concludes that introducing autonomous automobiles to the streets sooner could prevent hundreds of thousands of deaths.

Source:
https://www.washingtonpost.com/outlook/five-myths/five-myths-about-artificial-intelligence/2018/04/27/76c35408-4959-11e8-827e-190efaf1f1ee_story.html?noredirect=on&utm_term=.ddc594d6add8

Less than half of executives are confident in their own digital skills

The rise of big data and analytics talent will be necessary to drive change. However, such growth could be restricted by a lack of skilled people in the market.

According to a Deloitte survey less than half (45 per cent) of executives are confident in their own digital skills and ability to lead their organisation in the digital economy, while just 16 per cent believe their talent pool has enough knowledge and expertise to deliver their digital strategy.

Confidence in digital skills is currently low, almost half (49 per cent) of executives plan to invest more than £10 million in digital technologies and ways of working by 2020. 35 per cent plan to invest more than £10 million in the 2018 alone. 38 per cent of executives who say their organisation will invest in three or more emerging technologies over the next two years say that they do not have a coherent strategy in place.

The lack of confidence in digital leadership has not stopped organisations from embracing new technologies. Two in five (41 per cent) businesses have invested in AI technology, up from one in five (22 per cent) who said they had in 2017. Overall, 10 per cent have already invested more than £5 million in AI technology, with 15 per cent planning to invest more than £5 million in the coming year. Despite significant investments having already been made in AI, less than one in four (23 per cent) say that their leadership team has a clear understanding of the technology and how it will impact their business. Overall by 2020, 82 per cent of executives plan to invest in AI, while 70 per cent plan to invest in robotic and cognitive automation and 57 per cent in blockchain.

Only 12 per cent of leaders believe UK school leavers and graduates have the right digital skills, down from 20 per cent who said the same in 2017. Over three-quarters are experiencing challenges in recruiting employees with the relevant digital skills. Data scientists and analysts remain the most difficult roles to recruit and retain. While executives continue to worry that not enough school-leavers and graduates have the right mix of digital skills, only 17 per cent believe that UK companies lead the way with digital.

Source: https://www2.deloitte.com/uk/en/pages/press-releases/articles/less-than-half-of-executives-believe-they-have-digital-skills.html

The reality is that, the economy will be changed, and new workplaces will appear

Robots can take over communications, computing, and thinking, but there will be limits even here. There will still be the need for highly qualified professionals, such as engineers, architects, or judges. Also, tasks, where the dexterity remains beyond that of robot fingers, will remain for the foreseeable future. Machines cannot replace the emotional intelligence of a person. Occupations that require social skills or creativity or represent a high-quality personal service cannot be substituted by AI.

Tech companies such as Apple and Microsoft want to automate as many working processes as possible with learning machines. The employees of tomorrow must be more flexible because in the future man will have to adapt to the computer and not vice versa. Employees will have to adjust to more flexible working practices: a fixation on permanent locations and times is often no longer required. This implies more flexibility and freedom on the one hand, but on the other hand, work and personal time will intermingle.

In addition to the technical expertise, specialist and managerial staff must bring a more in-depth process knowledge and have a higher willingness to undergo independent and ongoing training in the appropriate technologies. Furthermore, a good understanding of all security-relevant questions relating to IT technology and legal security will be a basic requirement.

According to Bernhard Rohleder, Managing Director of BITKOM, new, exciting, and challenging jobs will be created. The number of workers with low-grade qualifications will fall, but staff with correspondingly high, mainly digital skills are now already increasingly in demand. The number of employees in the IT industry has been increasing continuously for years, as do the number of vacancies in this sector that are difficult to fill.

Artificial intelligence will bring about a tremendous shift in the labor market. The important thing is to remain flexible and open to new ideas.

Source: https://www.hrtechnologist.com/articles/digital-transformation/artificial-intelligence-and-the-future-of-human-labor/

AI creating jobs

A report from PricewaterhouseCoopers argued that AI would create slightly more jobs (7.2m) than it displaced (7m) by boosting economic growth. The firm estimated about 20% of jobs would be automated over the next 20 years and no sector would be unaffected.

According to PwC, AI and related technologies such as robotics, drones and driverless vehicles would replace human workers in some areas, but also create many additional jobs as productivity and real incomes rise and new and better products were developed. Healthcare and social work would be the biggest winners from AI, where employment could increase by nearly 1 million on a net basis, equivalent to more than a fifth of existing jobs in the sector.

Professional, scientific and technical services, including law, accounting, architecture and advertising firms, are forecast to get the second-biggest boost, gaining nearly half a million jobs, while education is set to get almost 200,000 extra jobs. PwC estimated the manufacturing sector could lose a quarter of current jobs through automation by 2037, a total of nearly 700,000.

Transport and storage are estimated to lose 22% of jobs – nearly 400,000 – followed by public administration and defence, with a loss of almost 275,000 jobs, an 18% reduction. Clerical tasks in the public sector are likely to be replaced by algorithms while in the defence industry humans will increasingly be replaced by drones and other technologies.

London – home to more than a quarter of the UK's professional, scientific and technical activities – will benefit the most from AI, with a 2.3% boost, or 138,000 extra jobs, the report said. The east Midlands is expected to see the biggest net reduction in jobs: 27,000, a 1.1% drop.

Source: https://www.theguardian.com/technology/2018/jul/17/artificial-intelligence-will-be-net-uk-jobs-creator-finds-report

Chapter III:

Technological Threats of AI

Highlights

- The use of AI for warfare is one of the most pressing physical threats posed by AI.
- Weaponized AI is an increasingly pressing concern internationally. Google employees recently signed an open letter of protest to the company's CEO due to Google's involvement with a US Department of Defense drone program.
- The US is among those countries most aggressively working on utilizing AI for military purposes. The Trump administration has announced plans to create a new Joint Artificial Intelligence Center to coordinate all existing AI-related programs across the Defense Department. Additionally, DARPA announced in September 2018 that they are committing $2 billion over the next 5 years to help make AI systems trusted and accepted by military commanders.
- The UK Government is taking a strong position on the development of safe and ethical AI, and has incredibly strong potentials to lead the world in the development of “Good Trusted AI.”
- Speaking on the topic of the new Centre for Data Ethics, which is just one of many initiatives to reduce risks associated with AI, UK Prime Minister Theresa May has stated that "This would be a “world-first advisory body which would review the current “governance landscape” and advise the Government on “ethical, safe and innovative uses of data, including AI" and that the centre "will not be a regulatory body, but it will provide the leadership that will shape how artificial intelligence is used," emphasising the UK Government's intention to "ensure that the adoption of AI is accompanied, and in some cases led, by a body similarly set up not just with technical experts who know what can be done but with ethicists who understand what should be done so that the gap between those two questions is not omitted."
- Additionally, the UK will be joining the World Economic Forum’s newly-established Council on Artificial Intelligence to help shape global governance around the topic of AI safety and ethics.

The risks of growth of AI

A group of 26 experts from around the world have warned in the Malicious AI report that Wanton proliferation of artificial intelligence technologies could enable new forms of cybercrime, political disruption and even physical attacks within five years.

In the Malicious AI report, the academic, industry and the charitable sector experts, describe AI as a “dual use technology” with potential military and civilian uses, akin to nuclear power, explosives and hacking tools. They argue that researchers need to consider potential misuse of AI far earlier in the course of their studies than they do at present, and work to create appropriate regulatory frameworks to prevent malicious uses of AI.

AI is likely to revolutionise the power of bad actors to threaten everyday life. In the digital sphere AI could be used to lower the barrier to entry for carrying out damaging hacking attacks. The technology could automate the discovery of critical software bugs or rapidly select potential victims for financial crime. It could even be used to abuse Facebook-style algorithmic profiling to create “social engineering” attacks designed to maximise the likelihood that a user will click on a malicious link or download an infected attachment.

The increasing influence of AI on the physical world means it is also vulnerable to AI misuse. The most widely discussed example involves weaponising “drone swarms”, fitting them with small explosives and self-driving technology and then setting them loose to carry out untraceable assassinations as so-called “slaughterbots”. Others may create “automated, hyper-personalised disinformation campaigns”, targeting every individual voter with a distinct set of lies designed to influence their behaviour. Or AI could simply run “denial-of-information attacks”, generating so many convincing fake news stories that legitimate information becomes almost impossible to discern from the noise.

The report concedes that AI is the best defence against AI, but argues that “*AI-based defence is not a panacea, especially when we look beyond the digital domain*”.

Source: https://www.theguardian.com/technology/2018/feb/21/ai-security-threats-cybercrime-political-disruption-physical-attacks-report

MIlitary application of AI

Artificial intelligence is not a weapon. Instead, artificial intelligence, from a military perspective, is an enabler, much like electricity and the combustion engine. Thus, the effect of artificial intelligence on military power and international conflict will depend on particular applications of AI for militaries and policymakers.

The potential promise of AI—including its ability to improve the speed and accuracy of everything from logistics to battlefield planning and to help improve human decision-making—is driving militaries around the world to accelerate their research into and development of AI applications.

There are several possible AI applications for the military. Replacing frozen software with systems that do not need to be refreshed periodically creates a broad potential for creating more nimble systems, possibly at lower cost. AI could be used in training systems, for example, it could provide unpredictable and adaptive adversaries for training fighter pilots. Computer vision, the ability of software to understand photos and videos, could greatly help in processing the mountains of data from surveillance systems or for “pattern-of-life” surveillance. NLP, used by systems such as Amazon's Alexa, enables systems to interact with humans using natural language. NLP could enable systems to take orders without using keyboards. NLP also can translate documents and could serve as a translator in the future.

Other suggested applications might include: using AIs to solve logistics challenges; to support war games; to automate combat in so-called manned-unmanned operations; to speed weapon development and optimization, and for identifying targets (as well as non-combatants).

AI could enable a variety of new military concepts of operation on the battlefield, such as the oft-discussed “loyal wingman” idea, which posits a human airplane pilot or tank driver who could coordinate a number of uninhabited assets as well. The more complicated the battlespace, however, the more useful it will be for those “wingmen” to have algorithms that help them respond in cases where the coordinating human controller cannot directly guide them. Swarms, similarly, will likely require AI for coordination.

Source: https://www.rand.org/blog/2017/09/artificial-intelligence-and-the-military.html
https://thebulletin.org/landing_article/the-promise-and-peril-of-military-applications-of-artificial-intelligence/

US weaponizing AI

Artificial intelligence is a transformative technology, and US generals already see it as the next big weapon in their arsenal. Michael Griffin, Undersecretary of Defense for Research and Engineering, signaled how keen the military is to make use of AI at the Future of War 2018 conference held in Washington, DC.

Many AI researchers are already worried about military use of the technology. Google employees recently signed an open letter of protest to their CEO after the company's involvement in a US DoD drone program was revealed. It could prove hard for companies to resist lucrative military contracts, however. It seems inevitable that AI will be used for everything from data gathering and analysis to developing more sophisticated autonomous systems.

In September 2018 the Defense Department's cutting-edge research arm has promised to make the military's largest investment to date in artificial intelligence systems for U.S. weaponry, committing to spend up to $2 billion over the next five years in what it depicted as a new effort to make such systems more trusted and accepted by military commanders.

The DARPA investment is small by Pentagon spending standards, where the cost of buying and maintaining new F-35 warplanes is expected to exceed a trillion dollars. The agency sees its primary role as pushing forward new technological solutions to military problems, and the Trump administration's technical chieftains have strongly backed injecting artificial intelligence into more of America's weaponry as a means of competing better with Russian and Chinese military forces.

DARPA isn't the only Pentagon unit sponsoring AI research. The Trump administration is now in the process of creating a new Joint Artificial Intelligence Center in that building to help coordinate all the AI-related programs across the Defense Department.

Source: https://www.technologyreview.com/the-download/610842/the-us-military-desperately-wants-to-weaponize-ai/
https://www.theverge.com/2018/9/8/17833160/pentagon-darpa-artificial-intelligence-ai-investment

China in race in AI Warfare

In May 2018 Chinese supreme leader Xi Jinping met with senior military scientists as chairman of the all-powerful Central Military Commission. During the meeting, the Chinese leader was photographed at the PLA Academy of Military Sciences shaking hands with Major General Li Deyi, a leading authority on artificial intelligence and a key figure in the Chinese military's effort to overtake the United States in the emerging field of advanced weapons.

The Chinese military quest for integrating AI into its tanks, naval forces and aircraft is the part of China's asymmetric or "assassin's mace" warfare strategy – building high-technology arms that will enable China's weaker forces to defeat the more powerful military in any future conflict.

Wang Changqing, a Chinese weapons designer, said future cruise missiles "*will have a very high level of AI and automation. They will allow commanders to control them in a real-time manner, or to use a fire-and-forget mode, or even to add more tasks to in-flight missiles.*"

China's application of AI to its growing cyber warfare capabilities also will increase the danger posed by cyber attacks and espionage. China's advanced AI-powered arms are among Beijing's most closely guarded secrets. Little is known about how far along China's military has developed these AI-powered weapons that include autonomous tanks and land vehicles, submarines and surface warships as well as bombers, fighters and drone aircraft. China recently demonstrated the use of an unmanned tank and showed off a swarm of drone aircraft as part of its AI military program.

Big Data will provide the fuel for Chinese intelligent combat through gathering masses of information used in algorithm training, pattern mining and optimization analysis crunched by powerful computers. The combined elements will guide military operations by producing intelligence analyses and battle plans for both troops and unmanned systems that will then conduct rapid and accurate attacks on targets. The Chinese believe AI weaponry will learn rapidly from the battlefield "*like a human recruit growing into a battle-hardened veteran*".

Source: https://nationalinterest.org/blog/the-buzz/china-race-overtake-us-military-ai-warfare-26035?page=0%2C1

AI War

"*How Might Artificial Intelligence Affect the Risk of Nuclear War?*" by the nonprofit Rand Corporation, warns that AI could erode geopolitical stability and remove the status of nuclear weapons as a means of deterrence by 2040. The researchers said that AI in the future could encourage human actors to make catastrophic decisions. Improvements in sensory technology, for instance, could result in the destruction of retaliatory forces like submarine and mobile missiles.

AI could also tempt nations to launch a pre-emptive strike against another nation to gain bargaining power, even if they have no intention of carrying out an attack. "*There may be pressure to use AI before it is technologically mature, or it may be susceptible to adversarial subversion. Therefore, maintaining strategic stability in coming decades may prove extremely difficult and all nuclear powers must participate in the cultivation of institutions to help limit nuclear risk.*"

The RAND paper highlights the dangers of the use of AI to take military decisions rather than the threat of autonomous drones and other so-called "killer robots."

Artificial Intelligence is used extensively in its present form in certain areas like for defusing bombs, improvised explosive devices, carrying equipment on the warfront, surveillance and reconnaissance missions, etc. AI research In unmanned ground vehicles is one of the fastest growing segments. It is expected that Unmanned Ground Vehicles will learn, amass knowledge, plan, learn spoken languages, perceive threats, corroborate with other robots to manipulate objects among other things.

Some experts fear that an increased reliance on AI could lead to new types of catastrophic mistakes. On the other hand, if the nuclear powers manage to establish a form of strategic stability compatible with the emerging capabilities that AI might provide, the machines could reduce distrust and alleviate international tensions, thereby decreasing the risk of nuclear war. Maintaining strategic stability in the coming decades may prove extremely difficult, and all nuclear powers will have to participate in the cultivation of institutions to help limit nuclear risk. This goal will demand a fortuitous combination of technological, military, and diplomatic measures that will require rival states to cooperate.

Source: https://www.cnbc.com/2018/04/25/ai-could-lead-to-a-nuclear-war-by-2040-rand-corporation-warns.html
https://www.rand.org/pubs/perspectives/PE296.html

UK Military and AI

In September 2018 it was revealed that UK has created and successfully tested a new arsenal of military robots, which will allow British soldiers to have the edge in the war on ISIS and fight in a more secure and rapid way.

The artificial intelligence technology, called SAPIENT, will be able to scan battlefields and identify hidden attackers, by sending sensors to soldiers on the ground.

The Ministry of Defence explained it will be a revolutionary addition to Britain's offence capabilities, as the system reduces human error and allows soldiers to freely and safely move on the ground. The robots were tested on the streets of Montreal.

Defence Minister Stuart Andrew said about the UK-created artificial intelligence: "*This British system can act as autonomous eyes in the urban battlefield.*" Present and former intelligence officers told the Times newspaper the Government Communications Headquarters (GCHQ) and British armed forces developed the use of new cyber technologies to spread malware to block jihadists' access to data.

The arsenal also includes measures to disrupt the terrorists' cash transactions and their online propaganda.

The director of GCHQ, Jeremy Fleming, said in April 2017 that his agency had started developing offensive weapons "to take the terrorism fight online". He had also warned that Britain's adversaries were "becoming more tech savvy".

Plans for a new UK cyber force, which would involve more than 2,000 operatives, are also close to being agreed at a cost which could run into the hundreds of millions of pounds, according to reports.

Sky News reported the new unit will nearly quadruple the number of British cyber specialists and focus on offensive operations.

Source: https://www.express.co.uk/news/uk/1021798/UK-military-robots-war-British-Britain-World-War-3-Artificial-Intelligence-soldiers

Chapter IV:

AI and Issues with Inequality and Statistic Related Biases

Highlights

- One challenge related to AI bias is that it is fundamentally based on statistical analysis, and as such will be systematically biased to give greater emphasis to the majority and lesser emphasis to the minority.
- The importance of inclusiveness in AI is becoming an increasingly hot topic, and an increasing number of national AI industrial strategies are explicitly accounting for inclusion in their frameworks.
- India's national strategy, for instance, is geared toward ensuring inclusive social growth.
- Meanwhile, Canada and France have announced the formation of specific task forces to develop an international study on inclusive and ethical AI.
- Despite these positive indications, more work needs to be done on explicitly accounting for inclusiveness and prevention of bias national AI strategies, ensuring that the development of nation's AI industry allow for maximum societal participation.
- Many such proposals and guidelines focus on the need to develop AI safely and ethically, with a strong focus on informed and transparent AI governance, on maximizing the social impact of AI for the benefit of a wide variety of UK stakeholders, and in ensuring that the deliverables of the AI industry serve to promote social good and well being among as large a proportion of the nation's population as possible.
- A report released by the House of Lords Artificial Intelligence Committee has attempted to put the above motivation into action by offering five core principles for the safe and ethical development of AI in the UK in a report entitled "AI in the UK: Ready, Willing and Able?", focusing on how AI should be developed for the wider benefit of humanity, should operate according to principles of intelligibility and fairness, should not be used to reduce data rights and privacy, should include initiatives to help citizens become educated about the changes brought about through AI so that they can use them to their social benefit, and should avoid providing AI with the power to hurt, destroy or deceive human beings.

One of the challenges regarding AI

The biggest actual threat faced by humans, when it comes to AI it's biased algorithms and it disproportionately affects the poor and marginalized. Machine learning algorithms, whether in the form of "AI" or simple shortcuts for sifting through data, are incapable of making rational decisions because they don't rationalize — they find patterns.

The bias debate broke wide-open when Pro-Publica published a damning article exposing the apparent bias in the COMPAS algorithms – a system that's used to sentence accused criminals based on several factors, including race. Basically, the report clearly showed several cases where it was obvious that the big fancy algorithm predicts recidivism rates based on skin tone.

In an age where algorithms are "helping" government employees do their jobs, if you're not straight, not white, or not living above the poverty line you're at greater risk of unfair bias. Matters of sexuality and race may not be intrinsically related to poverty or disenfranchisement, but the marginalization of minorities is. LBGTQ+ individuals and black men, for example, already face unfair legislation and systemic injustice. Using algorithms to perpetuate that is nothing more than automating cruelty.

Writer Elizabeth Rico believes unfair predictive analysis software may have influenced a social services investigator to take away her children. In the article, published on UNDARK, she said:

"… the 131 indicators that feed into the algorithm include records for enrollment in Medicaid and other federal assistance programs, as well as public health records regarding mental-health and substance-use treatments. Putnam-Hornstein stresses that engaging with these services is not an automatic recipe for a high score. But more information exists on those who use the services than on those who don't. Families who don't have enough information in the system are excluded from being scored."

The best intentions of researchers and scientists are no match for capitalism and partisan politics.

Source: https://thenextweb.com/artificial-intelligence/2018/09/21/artificial-intelligence-hates-the-poor-and-disenfranchised/

Inclusion in the Age of Artificial Intelligence

From relentless automation to algorithmic bias and human rights abuses, artificial intelligence (AI) has a laundry list of well-known potential costs and risks that do not bode well for the future of inclusion. Scant regulation and oversight, in addition to a workforce unequipped with the skills for the jobs of tomorrow, will result in greater inequality, discrimination, and exclusion. To achieve greater inclusion and maximize the social impact of AI, we need innovative and forward-thinking public policies. In the hands of governments, the technology is “primed for abuse” and a “grave threat” to civil rights and liberties.

In the past year, over 15 countries have released national strategies to promote the use and development of AI. They almost all include multi-million dollar investments in basic and applied AI research, initiatives to encourage the uptake of AI across the economy, and steps to develop and attract AI talent. Many countries also seek to become the “global leader” in specific areas of AI: the EU wants to set the global standards for AI ethics, China wants to be the world’s primary AI innovation center, and Canada wants to be the global leader in AI research and training.

Unlike other national strategies, each of India’s initiatives are geared towards ensuring social and inclusive growth. Canada and France announced ahead of the 2018 G7 Summit a new task force to develop an international study group on inclusive and ethical AI. Japan’s strategy, likewise, focuses on the industrialization of AI solutions for social problems that Japan and the world faces.

AI can be used to increase productivity, competitiveness, and economic development, but it must also be used to enhance the ability of every person to actively and fully participate in all aspects of life that are meaningful to them.

From Indigenous rights to gender equality, from cleaner water to energy conservation, AI technologies have a lesser-known list of potential benefits and opportunities for the future of inclusion. It is the responsibility of governments to invest in these benefits to ensure that the age of AI is inclusive of everyone.

Source: https://medium.com/politics-ai/inclusion-in-the-age-of-artificial-intelligence-37e0c906987d

AI-Politicians: A Revolution In Politics

During the recent presidential elections in Russia, a person named “Alice” ran as a candidate. While Alice didn’t win, she did receive 25,000 votes. Alice was an artificial intelligence (AI) system created by Yandex, Russia’s equivalent to Google.

In the past, humans ran for office. Tomorrow, AI will. And, AI may win, as people become increasingly frustrated with “human politicians.” What happened in Russia’s presidential election is reflective of how politics is changing in the age of AI.

In April, 2018, during a mayoral race in a part of Tokyo, an AI named “Michihito Matsuda” placed third with 4,000 votes. His campaign slogan: ”Artificial intelligence will change Tama City.” Alongside Alice and Michihito is SAM, an AI from New Zealand. SAM, who is referred to as a she, is being created to run in the 2020 general elections and has been called the first virtual politician in the world. Today, SAM is reaching out to voters through Facebook Messenger and is sharing her thoughts on climate change, healthcare and education, among other topics.

What kind of decisions might an AI-politician make once elected? The first layer is the idea of “special interests.” Today, special interests are organizations who donate money to a politician and then call in favors once the politician is elected. This doesn’t change with AI-politicians because the AI itself is being created by a company or person.

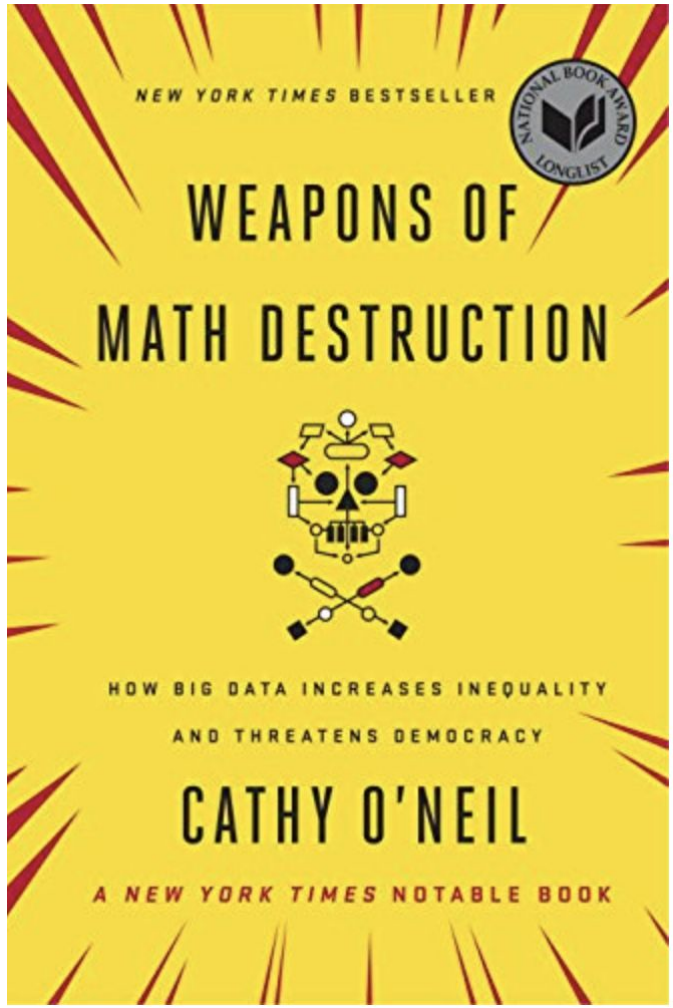
The second layer is “ethics.” Human politicians suffer from all kinds of ethical dilemmas and some of these dilemmas make headlines. AI-politicians will also suffer from ethical dilemmas but of a different kind. AI-politicians will need to be loaded with ethics that make the politicians understand the impact of what they are doing.

The third layer is “appointment.” If AI-politicians exist, they may not necessarily have to be elected. Future human political leaders might appoint AI into certain positions. In China, several AI-systems are being developed to help diplomats make decisions. The AI-systems will sift through huge amounts of data, from casual posts on social media to data supplied by Chinese intelligence agencies.

Source: https://medium.com/politics-ai/ai-politicians-a-revolution-in-politics-11a7e4ce90b0

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy

Cathy O'Neil is a data scientist, author, a Harvard PhD graduate in mathematics and actively involved in the Occupy movement. O'Neil wrote "*Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*" in 2016 describing the way that math can be manipulated by biases and affect every aspect of our lives.



As well as questioning the two-party system in the US, she's also looked at how mathematics has been used in the housing and banking sector to affect people's lives. This idea is at the heart of O'Neil's thinking on why algorithms can be so harmful. In theory, mathematics is neutral – two plus two equals four regardless of what anyone wishes the answer was. But in practice, mathematical algorithms can be formulated and tweaked based on powerful interests.

O'Neil's book explains how other mathematical models do a similar thing – such as the ones used to measure the likelihood an individual will relapse into criminal behavior. When someone is classed as "*high risk*", they're more likely to get a longer sentence and find it harder to find a job when they eventually do get out. That person is then more likely to commit another crime, and so the model looks like it got it right. And then there are those biases. Contrary to popular opinion that algorithms are purely objective, O'Neil explains in her book that "models are opinions embedded in mathematics".

Ultimately algorithms, according to O'Neil, reinforce discrimination and widen inequality, "*using people's fear and trust of mathematics to prevent them from asking questions*". But sometimes it's hard for non-statisticians to know which questions to ask. O'Neil's advice is to be persistent. "*People should feel more entitled to push back and ask for evidence, but they seem to fold a little too quickly when they're told that it's complicated*," she says.

Source: https://www.theguardian.com/books/2016/oct/27/cathy-oneil-weapons-of-math-destruction-algorithms-big-data