Technology and Society
TAS152

Note

The content of the STADIO Study Guides and teaching documents is not intended to be sold or used for commercial purposes. Such content is in essence part of tuition and constitutes an integral part of the learning experience, regardless of mode.

Links to websites and videos were active and functioning at the time of publication. We apologise in advance if there are instances where the owners of the sites or videos have terminated them. Please contact us in such cases.

A Glossary of terms is provided at the end of this study guide to clarify some important terms.

Any reference to the masculine gender may also imply the feminine. Similarly, singular may also refer to plural and vice versa.

Table of contents

Topic 4
Artificial intelligence ethics

4.1   INTRODUCTION

This topic relates to the following module outcome/s:

1.   Demonstrate an understanding of the ethical challenges posed by artificial intelligence.

This topic focuses on the potential impact of artificial intelligence (AI) and robotics within the context of the fourth industrial revolution. The topic has been divided into three distinct but related subsections:

Subsection 4.2 below (Bossman, 2016:102-107) provides an overview of key ethical issues related to the use of new AI-based technologies and recommends the analysis of potential risks.

Subsection 4.3 below (Smith, 2018: 108-110) highlights the importance of using AI for the common good, based on an overarching ethical vision.

Subsection 4.4 below (Gillen and Reddy, 2018: 110-114) discusses the ethical implications and potential risks of using AI-based algorithms to replace decision-making that was previously done by humans.

In this topic, you will gain knowledge in the following areas:

1.   Nine ethical issues in artificial intelligence
Five core principles to keep AI ethical
Ethics should inform AI

4.2   NINE ETHICAL ISSUES IN ARTIFICIAL INTELLIGENCE

Prescribed reading

Read Section 5.1 on pages 102-107 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

4.2.1   Introduction

Artificial intelligence (AI) is increasingly being incorporated in emerging and new technologies. However, this phenomenon has given rise to a number of challenges related to the upholding of ethical standards and the risks that may arise from the use of intelligent machine systems that incorporate AI-based technologies.

In this subsection, Bossman (2016:102-107) discusses nine ethical issues that are associated with the use of AI:
1.   Unemployment
2.   Inequality
3.   Humanity
4.   Artificial stupidity
5.   Racist robots
6.   Security
7.   Evil geniuses
8.   Singularity
9.   Robot rights

4.2.2   Unemployment: The impact of job losses

Increasing levels of automation within large manufacturing industries has resulted in job losses for many manual workers. Although some of those workers might undergo training that would prepare them for more responsible positions, an increasing number of manual workers face the possibility of unemployment, leaving them unable to provide for their families. Bossman (2016:102) cites the example of truck drivers in the United States, whose jobs are under threat from the development of autonomous, self-driving vehicles – which have a lower risk of accidents, since there is no driver to fall asleep at the wheel and veer off the road (Bossman, 2016:102).

In a recent research paper, Parschau and Hauge (2020) investigated the potential impact of automation on employment within the South African clothing industry. They found that rather than threatening existing jobs, automated production lowered manufacturing costs and increased productivity, leading to an increase in the demand for labour; while the use of digital technologies makes specialised production runs more economically feasible. In addition, the ability to

locate manufacturing facilities close to retail outlets (rather than outsourcing or offshoring) creates additional employment opportunities for local workers. Nevertheless, Parschau and Hauge have estimated that by 2030, about 10% of workers who are currently employed in manufacturing industries will need to find alternative employment in other sectors.

Parschau and Hauge (2020) subsequently interviewed stakeholders employed in a variety of roles within the South African clothes manufacturing sector. They also observed production processes in order to familiarise themselves with the organisation's policies and practices. Barriers to automation were mainly related to insufficient financial resources, limited availability of technology, lack of management and maintenance skills and inadequate infrastructure. However, most companies reported that automation had led to significant business growth and "allowed us to triple our business and almost triple the number of workers" (Parschau and Hauge, 2020). This suggests that employment opportunities in automated manufacturing industries may be better than previously predicted.

<div style="border: 2px solid #4a90d9; padding: 10px;">

Activity

Automated manufacturing processes could lead to increased unemployment levels in both first world and third world countries. Suggest two ways in which affected South African workers could re-skill themselves in order to support themselves and their families.

</div>

4.2.3   Inequality: Distributing the wealth created by machines

<div style="border: 2px solid #4a90d9; padding: 10px;">

Prescribed reading

Read Section 5.1.2 on page 103 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

</div>

Traditionally, business organisations have paid their employees on an hourly, daily, weekly or monthly basis. With increasing levels of AI being built into automated manufacturing equipment and even playing a role in business decision-making, increasing numbers of employees who previously brought home a regular income, are now joining the ranks of the unemployed. While the owners and senior management of AI-driven companies benefit from this situation, a corresponding price is paid by employees who are likely to lose their jobs as the incorporation of AI in machine tools becomes increasingly

widespread. This leaves us facing a 21st century dilemma: How do we structure a fair post-labour economy? (Bossman, 2016:103).

Many employees who were previously employed in low-skilled jobs have been replaced by automated systems that are controlled by AI technologies (Suri, 2017). Research into the employment patterns of displaced workers found that most of them were not equipped to assume more specialised roles; and a decade after they had been retrenched, those workers were still earning less than their wage at the time they were retrenched. Instead, organisations are employing specialised staff in roles such as IT service management, website development and the provision of consulting services (Suri, 2017). The Organisation for Economic Cooperation and Development (OECD) has estimated that up to 40% of low-skilled workers in the US could face unemployment within the next 10 years, most of whom are not academically equipped to train for more specialised jobs.

According to Suri (2017), one way to address the problem of mass unemployment would be to introduce a basic income grant; however, Collins (2014) claims that this approach could lead to a 'fiscal crisis' in light of the growing number of unemployed workers within the US population. A more effective intervention might be to review and adapt school and college curricula in order to equip a new generation of graduates for employment in a technology-based world.

Activity

Tasks that were traditionally performed by office workers and production line employees are increasingly being replaced by automated systems driven by AI algorithms. List the qualities and skills that you think a new generation of employees will need to function effectively in a technology-dominated workplace.

4.2.4   Humanity: How machines affect our behaviour

Prescribed reading

Read Section 5.1.3 on pages 103-104 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

Artificially intelligent 'bots' are increasingly being used to communicate with human customers via text input. The underlying algorithms can be automatically optimised based on the responses of human customers, to the extent that bots are able to develop 'relationships' with the customers they interact with. Other bots are being incorporated in computer games to make them more addictive, which is not necessarily a desirable outcome. However, bots also have the potential to exert a positive influence on human behaviour, for example by suggesting ways in which interpersonal conflict might be resolved (Bossman, 2016:103-104).

'Social bots' generally refer to social media sites that are controlled by AI and not by human beings (Pozzana and Ferrara, 2020). Social bots communicate via text content which has been created using natural language generation and which may be deliberately manipulated in order to influence public opinion. During the 2020 French presidential elections, Pozzana and Ferrara collected more than 16 million election-related tweets which were then classified as being created either by a human or by a bot; to be considered for analysis, an account had to have posted at least 5 tweets. After excluding borderline cases where the account owner could not be ascertained, the researchers were left with 19,000 accounts that were identified as bots and 290,000 accounts that were identified as humans (Pozzana and Ferrara, 2020). Subsequent data analysis revealed five key behavioural differences between humans and bots:

1. Humans re-tweeted more frequently than bots.

2. Humans replied to tweets more frequently than bots.

3. Tweets posted by humans included more mentions than tweets posted by bots.

4. Tweets posted by humans were generally shorter than tweets posted by bots.

5. Tweets posted by humans gradually decreased in length as the sessions progressed, while no similar pattern was detected among tweets that were posted by bots.

Overall, the method of analysis that was applied by Pozzana and Ferrara (2020) appears to have successfully differentiated between social media sites controlled by humans and social media sites controlled by bots.

Activity

Do you think that bots incorporating AI algorithms are deliberately designed to develop 'relationships' with human users in a way that promotes addictive behaviour? If so, then who should be held responsible for monitoring the algorithms that influence those relationships? Write down an explanation of your reasoning.

4.2.5 Artificial stupidity: Guarding against mistakes

Prescribed reading

Read Section 5.1.5 on page 105 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

Human intelligence and machine intelligence both emerge from a process of focused learning. Human intelligence develops at home, in school, in the workplace and in other social situations; machine intelligence is learned by presenting a machine with many thousands of different 'patterns' and training the machine to recognise which of those patterns perform best in practice. However ongoing training and testing is a crucial element of machine learning, as existing algorithms need to adapt themselves over time in order to accommodate new patterns of behaviour (Bossman, 2016:104).

4.2.6 Robotic judgements: How do we eliminate AI bias?

Although we expect AI-based systems to be fair and neutral, we must also remember that AI systems are created by humans who may be biased or judgemental; in which case the resulting AI systems may discriminate unfairly against certain behaviours or population groups. Moreover, if the data that is used to train an algorithm under-represents or over-represents particular social groups (for example, men vs. women, or older vs. younger age groups) then the predictions made by that model will be influenced by the unbalanced training data. A frequently-cited example relates to the Amazon recruitment algorithm, which automatically reduced the number of female applicants who were considered eligible for employment, based on the fact that most of the job applications submitted over the previous 10 years had been made by males (Lee, Resnick and Barton, 2019).

Ongoing training of AI technology is essential to ensure that any gaps in the underlying logic are swiftly identified and resolved, so that people who are affected by robotic decision-making will not be unfairly excluded or discriminated

against. Siwicki (2021) points out unless all possible instances of the relevant data have been included in the training dataset, the resulting algorithms will unavoidably be biased; the only question is the extent of the bias. He cites an example related to the healthcare services that are provided at US medical facilities, and draws attention to the fact that the corresponding data is likely to be biased due to the fact that many people without medical insurance will not have been included in the training dataset (Siwicki, 2021). Nevertheless, AI remains a critical tool in the pursuit of social progress and positive change (Bossman, 2016:105).

Activity

Kulkarni (2021) discusses two types of bias that can be found in AI algorithms.

(a)   Algorithmic AI bias occurs when <u>biased data</u> (which may discriminate against e.g. a particular race, gender or culture) is used to train AI algorithms.

(b)   Societal AI bias occurs when <u>unrealistic assumptions</u> and blind spots influence the thinking of algorithm developers, giving rise to algorithms that reflect social intolerance and discrimination.

Working together with a fellow student, describe one example of algorithmic AI bias and one example of societal AI bias that could result if inadequate or outdated datasets are used to train AI algorithms.

4.2.7   Security: Preventing malicious use of AI

Prescribed reading

Read Section 5.1.6 on page 105 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

AI systems can be used with malicious intent, ranging from hacking into the database of a business competitor to the creation of conflict between nations. The role of cybersecurity in protecting the rights of governments, business organisations and individual citizens is thus assuming an increasingly important role (Bossman, 2016:105).

Following a two-day workshop held in Oxford, UK, an extensive report was compiled outlining approaches that could be employed to mitigate threats presented by the malicious use of AI technologies (Cussins, 2018). These threats focus primarily on digital security, e.g. cyberattacks; physical security, e.g. drone attacks; and political security, e.g. circulation of propaganda and manipulated videos.

Fortunately, countermeasures are available to address and control the threats listed above. These include the responsible disclosure and remediation of code vulnerabilities; the encryption of documents containing sensitive information; providing certificates of authenticity for images and video clips; alerting the public to fake news; regulating the development of military hardware; and prohibiting the sale of lethal autonomous weapons (Cussins, 2018).

The Oxford report also recommended that attention should be paid to the following areas:

Policy makers and technical researchers should collaborate in the investigation, prevention and mitigation of malicious uses of AI.

Researchers and engineers working in AI should engage in research that is intended to prevent possible uses of AI for harmful purposes.

Best practices that have been incorporated in traditional approaches to computer security should be adapted for use in AI-based security tools.

Ongoing discussions should be held between stakeholders and domain experts in order to identify and monitor potential threats.

4.2.8 Taking responsibility: Protection from unintended consequences

Prescribed reading

Read Section 5.1.7 on page 106 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

The effectiveness of AI algorithms depends on the quality of the underlying code, and the training that those algorithms have been exposed to. An analogy can be drawn with the game of 'broken down telephone' that many of us played during childhood, when participants sit in a circle and the first participant whispers a message to the person beside them, which is then passed on the next person until it finally reaches the end of the chain. The message that is received by the

final participant usually differs significantly from the original message, with no one person being responsible for that difference.

A similar situation can arise when an algorithm has been insufficiently trained, or when the training rules have not been updated to cater for new circumstances. It then becomes impossible to predict whether or not the outcome of the algorithm is still valid (Bossman, 2016:106). Two key concerns identified by the European Union Agency for Fundamental Rights relate to errors of representation, i.e. does the data adequately represent the characteristics of the population; and errors of measurement, i.e. does the data measure the attribute that it was intended to measure? (FRA Focus, n.d.).

Activity

Think about who should assume responsibility for monitoring and controlling the quality of AI algorithms, and how the users of AI-based software can know that the algorithms have been adequately tested and/or updated. Discuss this question with a fellow student and then write down your views as to who should monitor the quality of AI algorithms.

4.2.9   Singularity: Can we control complex intelligent systems?

Prescribed reading

Read Section 5.1.8 on page 106 of Bossman (2016) *Top nine ethical issues in artificial intelligence*.

Human intelligence and ingenuity combined with education and experience has allowed us to dominate other species, whether animals or plants. We are able to control and manage the behaviour of animals by training them perform certain tasks, by farming them along with plants as a source of food for humans, or (if they are sufficiently exotic) by physically restraining them in cages where they can be viewed by members of the public. At the same time, however, we need to consider the possibility that in years to come, advanced AI machines will have learned to defend themselves against humans by self-modifying the algorithms that determine their own behaviour (Bossman, 2016:106).

4.2.10 The rights of robots: The legal status and humane treatment of AI

AI systems rely on 'reward' mechanisms associated with genetic algorithms to improve their decision-making performance. If one algorithm leads to a better outcome than other similar algorithms, then the most successful algorithm will be retained (and thus rewarded) while less effective instances of the algorithm will be deleted. However, algorithmic success is temporary: a previously successful algorithm may become less effective when changes occur in its environment, and will in turn be replaced by a more effective competing algorithm.

Bossman (2016:107) asks us to consider a future scenario in which machines that rely on the logic embedded in genetic algorithms can perceive, feel and act. Does this mean they should be regarded as intelligent beings? Future generations of humans may perhaps be faced by a new set of ethical questions needing to be resolved.

4.3   FIVE CORE PRINCIPLES TO KEEP AI ETHICAL

4.3.1   Introduction

Technology leaders working in the fields of robotics and AI have voiced concerns about potential risks may be associated with the use of intelligent machines (Smith, 2018:108-110). For example, intelligent weapons that are intended to defend civilian populations against threats of potential conflict, could in the future

be trained to attack human populations – or might even develop the capability of training themselves for this purpose. (Of course, we have no way of knowing whether any such scenarios would be likely to arise in the future.)

Kai-Fu Lee (2021) suggests that a number of benefits are associated with the use of autonomous weapons, based predominantly on the claim that fewer soldiers' lives would be lost if wars were fought by intelligent machines. However, this view appears to be logically flawed, since the loss of fewer soldiers' lives might well be accompanied by a considerably higher number of civilian deaths as well as extensive damage to infrastructure, for which nobody might be held accountable. More importantly, the moral implications associated with the use of autonomous weaponry go far beyond discussions relating to the number of lives lost. According to UN Secretary-General António Guterres (quoted in Lee, 2021): "The prospect of machines with the discretion and power to take human life is morally repugnant"; while Lee (2021) ends his article by stating that "Autonomous weapons are the AI application that most clearly and deeply conflicts with our morals and threatens humanity".

Note

In April 2018, the UK Government's Select Committee on Artificial Intelligence published a 180-page report entitled *'AI in the UK: ready, willing and able?'*. Interested students are encouraged to read pages 5-7 of that publication, which highlights a variety of AI-related challenges and opportunities, and suggests steps that could be taken to make the most of the innovative opportunities that AI has to offer (UK Government, 2018:5-7).

Smith (2018:108-110) highlights five core principles contained in that report which are intended to guide and inform the ethical use of AI. Those five core principles are outlined in subsections 4.3.2 to 4.3.6. below.

4.3.2   Using AI as a force for good

PRINCIPLE 1: AI should be developed for the common good and benefit of humanity.

This principle argues for the development of a shared framework that will support the ethical design and implementation of AI and will help to clarify the ways in which AI technologies can make a positive contribution to society. The proposed

ethical framework should cater for as diverse a population as possible, with close attention being paid to the elimination of historical prejudices.

<div style="border: 2px solid #4a86e8; padding: 10px;">

Activity

Identify an example of unfair discrimination against a particular population group. Write down a summary of the historical prejudice that needs to be addressed; then consider who would have benefited from that discriminatory practice and suggest how an appropriate form of compensation might be determined.

</div>

4.3.3   Intelligibility and fairness

PRINCIPLE 2: AI systems should be intelligible and should be fairly applied.

Organisations whose processes or products incorporate AI-based technologies should require the developers of those technologies to ensure that they are intelligible to users and are applied fairly to all members of society.

4.3.4   Data protection

PRINCIPLE 3: AI systems should not undermine either the data rights or the privacy of individuals, families or communities.

Within large commercial organisations, details of customers and their transactions are frequently stored in large databases. This information can subsequently be manipulated and analysed in order to generate valuable strategic information, which can be used for e.g. targeted marketing aimed specific groups of consumers. However, consumers have a right to the protection of their privacy, and it is the responsibility of every organisation that has access to their data, to ensure that published reports reflect only anonymised and aggregated information.

4.3.5   Working with AI

PRINCIPLE 4: All people should have the right to be educated about the potential impact of AI technologies.

While they are still at school, children should learn about the potential benefits (and risks) associated with the use of AI technologies, and should be exposed to

simple AI applications in order to prepare themselves for the workplaces of the future. For adults who joined the labour market before the advent of AI, additional training focusing on the theory, application, benefits and risks associated with the implementation of AI technologies should be provided to help consolidate their career trajectories.

4.3.6   Regulating the destructive power of AI

PRINCIPLE 5: AI researchers must ensure that the design, development, regulation and deployment of AI is governed by international norms, and will not be used for harmful purposes against human beings.

This principle speaks for itself. AI technologies should never be invested with the power to deceive, hurt or destroy human beings. All AI researchers need to consider the ethical implications of their work based on international norms for the design, development, regulation and deployment of AI across different sectors and industries.

Activity

Make a list, in order of importance, of the five core principles outlined above, and write down a brief explanation of the basis on which you determined their relative importance.

4.4   ETHICS SHOULD INFORM AI

4.4.1   Introduction

Prescribed reading

Read Section 5.3.1 on page 111 of Gillan & Reddy (2018) *Ethics should inform AI – but which ethics?*

Large datasets play an important role in facilitating the development of machine learning algorithms, which are an essential component of AI-based applications (Gillen and Reddy, 2018:110). However, it remains the responsibility of organisations that implement AI-based applications, to ensure that those

applications are used ethically and appropriately, and that the privacy of individuals is protected. These and other related issues are discussed in the subsections that follow (Gillen and Reddy, 2018:111-114).

4.4.2   Deontological vs teleological ethical standards

Ethical standards governing the responsible use of AI-based technologies need to be established on a global level, based on consultations with key stakeholders in business organisations and in national governments. In this subsection we consider the difference between deontological and teleological ethical standards.

Deontological ethical standards focus on the intentions of the technology creator and the means by which those intentions will be achieved. This approach ensures that both the process and the product will be aligned with the desired ethical standards (Gillen and Reddy, 2018:111).

Teleological ethical standards focus on the desired ends and outcomes that the technology creator wishes to achieve, without taking into account the ethical implications of how those outcomes will be achieved. However, this approach might encourage technology creators to take risks during the development process as long as their desired outcome is achieved (Gillen and Reddy, 2018:111).

For readers who are not persuaded that a deontological approach is the better of the two options, consider the situation of a hospital patient awaiting surgery. Would you prefer a surgeon who carefully performs each step in the surgical procedure and makes sure that it has been successfully completed before moving on to the next step in the process; or would you be happy with a surgeon who is willing to cut corners during the process in order to complete the surgery faster?

Activity

"When determining what is right and what is wrong, you should focus on the moral decision that needs to be made and ignore the potential consequences of that decision." Explain in your own words whether this statement represent a deontological or a teleological perspective.

For more information on the topic of deontological vs. teleological ethical standards, refer to DifferenceBetween.com (2019).

4.4.3   The golden rule of ethics

The 'golden rule' of ethics states that you should treat other people in the way that you would like them to treat you. Unfortunately, this principle does not necessarily apply in the case of AI-based technologies, since data that has been gathered from a large number of different sources and subsequently transformed via algorithmic learning before being incorporated in AI-based applications, will not necessarily be used for the benefit of the original contributors of that data (Gillen and Reddy, 2018:112).

4.4.4   Cultural relativism vs universalism

National cultures play an important role in influencing the ethical practices of different countries.

Cultural relativism is based on the theory that morality is subjectively based on the values that are practised within a particular culture. This approach has been adopted by countries such as China, that are unwilling to adopt universal principles which might impede the growth of their own economies (Gillen and Reddy, 2018:112.

Universalism is an ethical ideology that seeks the adherence of all nations to a mutually agreed set of ethical standards (Gillen and Reddy, 2018:112). However, many difficulties have been identified in relation to the development and implementation of a universal set of ethical ideals, including the challenge of defining a single moral code that is acceptable to all nations.

Activity

In some countries, people may snack on fried insects, a man may have several wives, or kissing in public may be forbidden.

1.   Explain whether this example illustrates cultural relativism or universalism.
2.   Describe a behaviour that you take for granted in your own country, but which might appear unusual to citizens of other countries.

4.4.5   Developing ethical as well as technical standards

"AI presents opportunities to complement and supplement human intelligence and enrich the way industry and governments operate" (Gillen and Reddy, 2018:113). At the same time, a variety of ethical risks need to be considered in relation to the development of AI-based technologies. How will the learning processes of cognitive machines be controlled and monitored, and who will assume responsibility for the associated outcomes? Predictability and trust will be important attributes of the new AI-based technologies, and will play an important role in encouraging further investment in the use of AI (Gillen and Reddy, 2018:113). To support this vision, a new legal framework based on a robust and universally accepted code of ethics will need to be developed and implemented. Government representatives, industry spokespersons, and members of the private sector, consumer groups and academia should all be invited to participate in drafting the new legal framework and its corresponding code of ethics (Gillen and Reddy, 2018:113).

In the context of agricultural and marine farming, Galaz *et al*. (2021:6) note that although production levels may increase due to the implementation of AI and related technologies, the associated impact on other related ecosystems such as forests and rivers must also be considered. Biodiversity needs to be maintained in order to provide resilience within the ecosystem, enabling it to resist environmental threats such as extreme weather patterns and the introduction of new diseases. Furthermore, research has shown that in general, smaller farms tend to be more productive and biodiverse, and play an important role in the development of resilient ecosystems (Galaz *et al*., 2021:6). For this reason, Galaz *et al*. (2021:6) recommend the development of sector-specific guidelines, product and process standards, based on new or amended legal-regulatory frameworks and monitored using AI technologies.

Ballard (2020) believes that Internet connectivity combined with the use of technologies such as the Internet of Things (IoT) could provide important benefits for rural farmers, for example by monitoring water supplies and water quality in real time. Internet connectivity would also allow people living in rural areas to benefit from online resources ranging from education programmes, banking facilities and healthcare advice to methods for improving agricultural

practices; and would also provide opportunities for them to earn an income while working from home (Ballard, 2020).

South Africa's National Development Plan 2030 was prepared by the National Planning Commission (2012) following an inclusive and participatory process involving all stakeholders. Chapter 6 of this document outlines three key areas that will be the focus of a new differentiated strategy for rural development (National Planning Commission, 2012:219).

Agricultural development based on successful land reform, employment creation and strong environmental safeguards.

Quality basic services, particularly education, healthcare and public transport.

In areas with greater economic potential, industries such as agro-processing, tourism, fisheries (in coastal areas) and small enterprise development should be developed with market support.

Activity

Discuss the potential impact of Internet connectivity, increasing levels of automation and the use of AI-based technologies on rural farms in remote areas of South Africa, with someone who might be interested in this topic. Write down a brief summary of your discussion.

Note

For more information about South Africa's new differentiated strategy in support of rural development, refer to:
National Planning Commission of South Africa. 2012. Chapter 6. An integrated and inclusive rural economy. In: *National Development Plan 2030: Our future – make it work*. Pretoria: Government Printer. [Online] Available from https://www.gov.za/issues/national-development-plan-2030 [Accessed 2022-03-24].

4.4.6   Ethical challenges

Finally, Gillen and Reddy (2018:114) highlight five ethical challenges related to the use of complex machine learning systems which will need to be addressed if new AI-based technologies are to receive universal acceptance.

1. Decision-making and liability: Who will assume responsibility for the outcomes of AI-based decisions or algorithmically-developed outcomes?

2. Transparency: Will it be possible to determine accountability for a decision by stepping through the AI-based algorithms that were applied during the process that was followed to reach a final outcome?

3. Bias: If the developers of a machine learning system unintentionally incorporate algorithmic bias in their software, will the underlying software system be able to identify and eliminate that bias?

4. Human values: Common sense is not an inherent attribute of AI systems. Gillen and Reddy (2018:114) emphasise that "Humans, not robots, are the responsible agents; it should be possible to find out who is responsible for any robot and its behaviour".

5. Data protection and intellectual property (IP): Large data sets play an important role in supporting the effective implementation and use of AI. Developers of machine learning systems need be aware of the implications of legislation that is intended to protect IP: for example, an AI system that was trained on one data set, may not apply its learning to a second data set without first obtaining permission from the owner of the second data set. Humans, not robots, are the responsible agents; it should be possible to find out who is responsible for any robot and its behaviour (Gillen and Reddy (2018:114).

Phoenix (2020) lists advantages and disadvantages associated with the use of AI technology. Advantages include:

The efficient and accurate performance of a wide variety of processing tasks.

The elimination of human errors caused by lapses in concentration.

The ability to perform repetitive tasks 24 hours a day.

Reduced employment costs for the organisation as human tasks are increasingly performed by machines.

The ability to rapidly perform and update predictive calculations regarding e.g. targeted marketing, new product development, or distribution routes.

The development of AI-based prosthetics that can be integrated with living tissue and controlled via the recipients own nervous system.

The rapid extraction, interpretation and transformation of data.

However, unexpected AI behaviours may emerge as AI systems learn to adapt their own algorithms (Phoenix, 2020).

The incorporation of AI in military weapons (including robots) would allow them to improve their combat strategies over time and thus increase the risk to human life during times of conflict.

AI systems are unable to judge the implications of their decisions based on a broader perspective that takes other factors into account.

AI systems are unable to take emotions such as kindness and compassion into account when making a decision.

The physical components within AI systems may fail over time, leading to unpredictable results.

Activity

AI technologies are increasingly being incorporated in medical devices ranging from pacemakers to prosthetic limbs. Think about who should be held responsible for errors in the algorithms that control those devices, and who should be responsible for addressing them. Write down your answers to these questions.

Other disadvantages associated with the use of AI include reduced employment opportunities for humans as their jobs are taken over by robots, with an accompanying loss of income; the high cost of the technologies that are needed to support AI systems; the inability of AI systems to think creatively; and uncertainties associated with the ownership and use of data that has been collected by AI systems (Phoenix, 2022).

Summary

In subsection 4.2 of this study guide, Bossman (2016) discusses the incorporation of artificial intelligence (AI) in new and emerging technologies, new

challenges that are likely to arise regarding the upholding of ethical standards, and the risks associated with intelligent machine systems. She claims that increasing automation of factory production lines will leave many workers facing unemployment, and points out that artificially intelligent 'bots' are increasingly taking over roles that were previously performed by call centre agents. A more immediate concern raised by Suri (2017) is that most retrenched workers lack the knowledge and skills that are needed to equip them for specialised roles. However, research conducted in South Africa by Parschau and Hauge (2020) found that within the SA clothing industry, lower manufacturing costs due to automation have allowed businesses to expand, thus creating additional employment opportunities.

Bossman (2016) also points out that data errors, as well as under or over-representation of the populations that are used to train AI algorithms, may result in unintended discrimination against particular groups, and ongoing training is essential to accommodate changing patterns of behaviour. Of even greater concern is the risk that AI systems may be used for malicious purposes via cyberattacks, drone attacks, manipulated videos, and (in extreme cases) the use of autonomous weapons; and Bossman (2016) raises the possibility that in a future world, advanced AI machines may have learned to defend themselves against humans.

In subsection 4.3 of this study guide, Smith (2018) raises the possibility that intelligent weapons could in future be trained to attack human populations. He then highlights five core principles identified by the UK Government's Select Committee on Artificial Intelligence, which are intended to guide the ethical use of AI (Smith, 2018:108-110). In particular, Smith highlights the importance of using AI for the common good, based on an overarching ethical vision supported by international norms for the design, development, regulation and deployment of AI across different sectors and industries.

In subsection 4.4 of this study guide, Gillen and Reddy (2018) discuss the ethical implications and potential risks of using AI-based algorithms to replace decision-making that was previously done by humans. They outline key differences between the deontological approach, in which both the process and the product are aligned with ethical standards; and the teleological approach, in which the final product is created without taking into account any related ethical implications. Gillen and Reddy (2018:112-113) then compare the ideologies of cultural relativism versus universalism, and raise questions as to how the future use of AI-based technologies will be monitored and controlled. They conclude that a new legal framework based on a robust and universally accepted code of ethics will need to be developed and implemented (Gillen and Reddy, 2018:113).

In the final section of this text, Gillen and Reddy (2018:114) highlight five ethical challenges related to the use of AI technologies: (a) Who will assume responsibility for the outcomes of AI-based decisions? (b) Will it be possible to 'step through' the underlying AI-based algorithms? (c) Will it be possible to identify and eliminate bias caused by unrepresentative or inadequate training data? (d) Who will be held responsible for the behaviour of AI systems or robots? (e) Will intellectual property legislation be able to prevent the transfer of learning between AI-based systems?

Self-Assessment Questions

Write an essay of approximately 350-450 words on each of the topics outlined below:

1. Search online for at least one example of algorithmic AI bias and at least one example of societal AI bias. Then discuss the importance of training and testing the algorithms that are incorporated in intelligent machines, and the problems that may arise if either algorithmic AI bias or societal AI bias is present in the training data.
   [Refer to subsection 4.2 of Topic 4 when answering this question.]

2. Lee (2021) suggests that a number of benefits are associated with the use of autonomous weapons, based predominantly on the claim that fewer soldiers' lives would be lost if wars were fought by intelligent machines. In contrast, Smith (2018:110) states that "the autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence". Write an essay in which you attempt to reconcile these differing points of view.
   [Refer to subsection 4.3 of Topic 4 when answering this question.]

3. Discuss the potential benefits of using AI-based algorithms to replace human decision-making, taking into account any associated risks. You may want to consider the bullet points below when constructing your answer.

   Benefits:
   AI-based algorithms can 'work' consistently for 24 hours a day.
   AI-based algorithms can rapidly extract, transform and analyse data.

Given appropriate guidelines, AI-based algorithms can control manufacturing equipment, autonomous vehicles, etc.

Risks:

Unrepresentative or inadequate training data may result in algorithmic bias.

AI technologies are likely to make bad decisions if the supporting data is flawed or incomplete.

[Refer to subsection 4.4 of Topic 4 when answering this question.]

Glossary of terms

AI-based prosthetic: An artificial device that is used to replace a missing body part.

Algorithm: A set of instructions that are used to perform a specific computation.

Algorithmic AI bias: Errors in the definition of an algorithm that lead to unfair outcomes.

Autonomous car: A vehicle that uses information obtained from environmental sensors to navigate, with little or no human involvement.

Autonomous weapon: A lethal weapon that operates either partly or entirely without human supervision.

Autonomy: The ability of an individual human being to make a rational informed decision.

Artificial intelligence: The ability of a computer to 'think for itself' based on logical tools that include knowledge representation, reasoning and machine learning.

Bandwidth: The amount of data that can be transmitted between two points within a given time frame (usually measured as bits per second).

Blockchain: A system that distributes transaction information across a network of linked computers, making it almost impossible to edit or delete any single entry.

Bot: A software application that has been programmed to perform a specific automated task, often replacing a human user.

Cascading failure: A process in which the failure of one component triggers the failure of other related components.

Cognitive behavioural therapy: A method of addressing anxiety-related disorders by developing appropriate coping strategies.

Contextual ethics: Takes into account the particular context within which an action was performed, rather than judging the action based only on absolute moral standards.

Cultural relativism: The idea that a person's beliefs and behaviour should be based on their own particular cultural context.

Cyber bullying: Undesirable behaviour in which somebody bullies other people on the Internet, often on social media sites.

Cybercrime: Illegal computer-based activity that takes place in cyberspace.

Cyberspace: A global network of IT infrastructures, telecommunications systems and computers.

Data anonymisation: The removal of personally identifiable information from data records.

Data privacy: Data privacy legislation ensures that personal information cannot be disclosed to other parties without the owner's consent.

Data protection: Ensures that important data is protected from loss or corruption and can be restored to a functional state if it becomes unusable.

Deontological ethics: A theory in which human actions are judged as being morally right or morally wrong based on their adherence to a set of underlying rules.

Digital pirate: Somebody who downloads online material from the Internet without the permission of the copyright holder.

Disruptive technology: An innovative technology that changes the way in which particular industries or businesses operate.

Domain: A network of computers and other devices that are controlled by a specific organisation and are linked to a particular IP address.

Drone attack: An attack by one or more unmanned aerial vehicle (UAVs) which may be equipped with military weapons.

Entropy: Measures the level of disorganisation within a particular system.

Ethos: The set of beliefs and behaviours that are representative of a particular social group.

Fake news: Inaccurate or misleading content that is intended to influence or deceive the public.

Fintech: Technologies that are used to automate the provision and use of financial services.

Fiscal crisis: A situation where the expenditure of a state or country is greater than its tax revenues.

Forensic specialist: An investigator who collects and analyses evidence related to a crime.

Genetic algorithm: A search method that is frequently used in AI applications to identify optimised solutions to complex problems.

Hierarchy of Needs (Maslow): A set of human psychological requirements progressing from physiological needs (level 1) to self-actualisation needs (level 5).

Human agency: The ability of an individual to achieve their potential based on their choices and actions.

Informational ecosystem: A theoretical structure that uses communication networks to meet the information needs of a population or community.

Intellectual property: Intangible creations or inventions that are legally protected against unauthorised duplication or use.

Internet access provider (IAP): An organisation that provides users with access to the Internet, but does not usually offer additional ancillary services.

Knowledge society: A society in which invests in education and knowledge creation in order to build human capital and foster innovation.

Legal norms: Rules that govern social behaviour by regulating the rights and duties of citizens and imposing sanctions when citizens do not adhere to prescribed rules.

Livelihood: The sources of income that are available to individuals, families and communities, such as wages or the proceeds of entrepreneurial activities.

Machine intelligence: A combination of pre-programmed problem-solving techniques with deductive logical abilities that together allow a machine to demonstrate original behaviour.

Machine learning technologies: Technologies that use large data sets to develop predictive algorithms, which can then be applied to autonomously solve data-based business problems.

Machine tool: A power-operated machine used in manufacturing facilities to shape metal components.

Net neutrality: The idea that Internet service providers should deliver the same level of service to all their customers.

Normative guidelines: Generally accepted standards that are used to evaluate behaviours, actions and outcomes.

Post-labour economy: An economy in which automation or the relocation of facilities to third-world countries has reduced the demand for skills that were previously supplied by human workers.

Predictive modelling: A statistical technique in which mathematical models are used to predict the probability of future outcomes.

Product development partnership (PDP): Collaboration between a not-for-profit organisation and a government or manufacturing partner, in order to deliver drugs and diagnostic devices to inhabitants of low-income countries.

Relation-based decision: The decision that will lead to the most optimal outcome is identified by evaluating the outcomes of all likely alternative decisions.

Robotisation: The use of industrial robots to perform automated tasks, usually as part of a manufacturing process.

Socially responsible investment (SRI): An investment strategy that takes into account the environmental, ethical or social impact of that investment.

Societal AI bias: A problem that arises when algorithm designers are influenced by, and then unwittingly reinforce, pre-existing social, cultural or institutional norms.

Stakeholder: A person or group of people who will be affected by the outcome of a business activity or project.

Supply chain: A process that encompasses the sourcing of raw materials, the manufacturing of finished goods, and the transportation of the final product to the distributor.

Sustainable Development Goals: A set of common goals identified by the United Nations that are intended to protect the planet and improve the social wellbeing of its inhabitants.

Targeted marketing: A marketing approach in which personalised advertising is used to target potential purchasers based on e.g. their income group or hobbies.

Techno-social disruption: The disruption of traditional practices following the introduction of new technologies or production methods within a particular market or industry.

Teleological ethical standards: Ethical standards which are based on the principle that the rightness or wrongness of an action can only be determined by its consequences.

Tweet: An online message created and saved by a Twitter user.

Universalism: The belief that generalised norms, values or concepts can be applied to all people and cultures.

Virtual environment: A computer-generated environment that allows a user wearing virtual reality video goggles to manoeuvre within a simulated 3D space.

References

Agrawal, S. 2021. Autonomous weapon systems: Our new soldiers, or a disaster waiting to happen? *Viterbi School of Engineering, USC*. [Online] Available from https://vce.usc.edu/volume-5-issue-1/autonomous-weapon-systems-our-new-soldiers-or-a-disaster-waiting-to-happen/ [Accessed 2022-05-22].

American Library Association. 2018. Net neutrality: An intellectual freedom issue. *American Library Association*, February 2018. [Online] Available from https://www.ala.org/advocacy/intfreedom/netneutrality [Accessed 2022-04-30].

Ammanath, B. 2021. Thinking through the ethics of new tech… before there's a problem. *Harvard Business Review*, November 2021.[Online] Available from https://hbr.org/2021/11/thinking-through-the-ethics-of-new-techbefore-theres-a-problem [Accessed 2022-02-16].

Anneroth, M. 2021. AI bias and human rights: Why ethical AI matters. *Ericsson Blog*, November 2021. [Online] Available from https://www.ericsson.com/en/blog/2021/11/ai-bias-what-is-it [Accessed 2022-05-17].

Anon. Not dated. Check if your email or phone is in a data breach. [Online] Available from https://haveibeenpwned.com/ [Accessed: 2022-03-18].

Arthur, K. & Owen, R. 2019. A micro-ethnographic study of big data-based innovation in the financial services sector: Governance, ethics and organisational practices. *Journal of Business Ethics*, 160:363-375. [Online] Available from https://link.springer.com/article/10.1007/s10551-019-04203-x [Accessed 2022-03-28].

Aschendbrand, J., Proctor, J. & Trebilcock, B. 2018. The ethical supply chain. *Supply Chain Management Review*, Nov. 2018. [Online] Available from https://www.scmr.com/article/the_ethical_supply_chain [Accessed 2022-03-31].

Avelino, F. 2021. Theories of power and social change. Power contestations and their implications for research on social change and innovation. *Journal of Political Power*, 14:3, 425-448. [Online] Available from https://doi.org/10.1080/2158379X.2021.1875307 [Accessed 2022-05-02].

Ayim, C., Kassahun, A., Addison, C. & Tekinerdogan, B. 2022. Adoption of ICT innovations in the agriculture sector in Africa: A review of the literature. *Agriculture and Food Security*; 11, 22. [Online] Available from https://agricultureandfoodsecurity.biomedcentral.com/articles/10.1186/s40066-022-00364-7 [Accessed 2022-04-19].

Ballard, B. 2020. New tech could halt the decline of rural communities. *The New Economy*, January 17 2020. [Online] Available from

https://www.theneweconomy.com/technology/silicon-villages-how-new-tech-could-halt-the-decline-of-rural-communities [Accessed 2022-03-24].

BBC News. 2017. *What is net neutrality and how could it affect you?* [Online video]. Available from https://www.youtube.com/watch?v=zq-2Yk5OgKc. [Accessed 2022-03-13].

Bossman, J. 2016. Top nine ethical issues in artificial intelligence (Chapter 5.1). In: Stückelberger, C. & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. Free PDF downloadable from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

Brown, S. 2020. A new study measures the actual impact of robots on jobs. It's significant. *MIT Management, Sloan School*. [Online] Available from https://mitsloan.mit.edu/ideas-made-to-matter/a-new-study-measures-actual-impact-robots-jobs-its-significant [Accessed 2022-04-20].

Bussel, J. 2013. Cyberspace. In: *Britannica Online Encyclopedia*. [Online] Available from https://www.britannica.com/topic/cyberspace [Accessed: 2022-02-18].

Çelik, C.B. 2016. Educational intervention for reducing Internet addiction tendencies. *Addicta: The Turkish Journal on Addictions*, Winter 2016, 3:375-386. [Online] Available from https://www.addicta.com.tr/en/educational-intervention-for-reducing-internet-addiction-tendencies-13108 [Accessed 2022-03-23].

Cemiloglu, D., Almourad, M.B., McAlaney, J. & Ali, R. 2022. Combatting digital addiction: Current approaches and future directions. *Technology in Society.* 68:1-11.[Online] Available from https://www.sciencedirect.com/science/article/pii/S0160791X21003079?via%3Dihub [Accessed 2022-03-23].

CEPEJ. 2018. European ethical charter on the use of artificial intelligence in judicial systems and their environment. *European Commission for the Efficiency of Justice, Strasbourg*. [Online] Available from https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c [Accessed 2022-04-18].

Chetty, M., Sundaresan, S., Muckaden, S., Feamster, N. and Calandro, E. 2013. Investigating Broadband performance in South Africa. *Towards Evidence-based ICT Policy and Regulation*, Vol. 2. [Online] Available from https://www.researchictafrica.net/docs/RIA_policy_paper_measuring_broadband_performance_South_Africa.pdf [Accessed 2022-04-29].

Cloudian. Not dated. *Data protection and privacy: 12 ways to protect user data*. [Online] Available from https://cloudian.com/guides/data-protection/data-protection-and-privacy-12-ways-to-protect-user-data/ [Accessed: 2022-03-17].

Collins, R. 2014. The end of middle-class work: No more escapes. In: *Does Capitalism have a Future?* (pp. 39-51). New York: Oxford University Press.

COMEST. 2019. Preliminary study on the ethics of artificial intelligence. *Extended Working Group on the Ethics of Artificial Intelligence*. [Online] Available from https://unesdoc.unesco.org/ark:/48223/pf0000367823 [Accessed 2022-04-17].

Congressional Research Service. 2021. *Defense primer: Cyberspace operations*. [Online] Available from https://fas.org/sgp/crs/natsec/IF10537.pdf [Accessed: 2022-03-14].

Convention on the Rights of the Child. 1989. *General Assembly resolution 44/25 of 20 November 1989*. [Online] Available from https://www.ohchr.org/documents/professionalinterest/crc.pdf [Accessed: 2022-02-28].

Cussins, J. 2018. How to prepare fore the malicious use of AI. *Future of Life Institute*, February 2018. [Online] Available from https://futureoflife.org/2018/02/27/how-to-prepare-for-malicious-ai/ [Accessed 2022-04-02].

Deloitte Financial Advisory Services LLP. Not dated. *5 insights on cyberattacks and intellectual property*. [Online] Available from https://www2.deloitte.com/content/dam/Deloitte/us/Documents/finance/us-fas-five-insights-on-cyber-attacks-and-intellectual-property.pdf [Accessed 2022-03-30].

Desmond, P. 2021. Data privacy laws in South Africa. *De Rebus*, July 2021. [Online] Available from https://www.derebus.org.za/data-privacy-laws-in-south-africa/#:~:text=1.,the%20subject%20has%20given%20consent [Accessed 2022-03-21].

DifferenceBetween.com. 2019. *Difference between teleological and deontological ethics*. [Online] Available from https://www.differencebetween.com/difference-between-teleological-and-deontological/ [Accessed 2022-03-22].

Dizikes, P. 2020. How does using robots in manufacturing impact employment? *Massachusetts Institute of Technology*. [Online] Available from https://news.mit.edu/2020/how-many-jobs-robots-replace-0504 [Accessed 2022-04-20].

Dodd, A. 2019. What is the Internet. *InformIT*, April 2019. [Online] Available from https://www.informit.com/articles/article.aspx?p=2963467&seqNum=4 [Accessed 2022-04-29].

Ethical Systems. Not dated. *Contextual Influences*. [Online] Available from https://www.ethicalsystems.org/contextual-influences/ [Accessed 2022-03-29].

Floridi, L. 2008. Information ethics, its nature and scope. In: v. d. Hoven, J. and Weckert, J. (eds.) *Information technology and moral philosophy*. Cambridge: Cambridge University Press.

FRA Focus. Not dated. Data quality and artificial intelligence–mitigating bias and error to protect fundamental rights. *European Union Agency for Human Rights*. [Online] Available from

https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-data-quality-and-ai_en.pdf [Accessed 2022-04-01].

Gaiani, S. 2008. Information and communication technologies (ICTs) for rural development in developing countries. *New Medit*, 7(1):50-57. [Online] Available from https://newmedit.iamb.it/2008/03/22/information-and-communication-technologies-icts-for-rural-development-in-developing-countries/ [Accessed 2022-04-19].

Galaz, V., Centeno, M.A., Callahan, P.W., Causevic, A., Patterson, T., Brass, I., Baum, S., Farber, D., Fischer, J., Garcia, D., McPhearson, T., Jiminez, D., King, B., Larcey, P. & Levy, K. 2021. Artificial intelligence, systemic risks and sustainability. *Technology in Society*, 67:1-10. [Online] Available from https://www.sciencedirect.com/science/article/pii/S0160791X21002165?via%3Dihub [Accessed 2022-03-23].

Garvey, M.D. 2021. A philosophical examination on the definition of cyberspace (Chapter 1). In: Carnovale, S. & Yeniyurt, S. (eds.) *Cyber security and supply chain management: Risks, challenges and solutions*, World Scientific Publishing Company. Downloadable from https://www.worldscientific.com/doi/pdf/10.1142/9789811233128_0001 [Accessed: 2022-02-14].

Gautier, A. & Somogyi, R. 2020. Prioritization vs zero-rating: Discrimination on the internet. *International Journal of Industrial Organization*, 73 (2020) 102662. [Online] Available from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7495184/ Accessed 2022-05-01.

Gillen, M. & Reddy, S. 2018. Ethics should inform AI – but which ethics? (Chapter 5.3). In: Stückelberger, C. & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. Free PDF downloadable from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

Globethics.net. 2018. Ethics in the information society: The nine P's. In: Stückelberger, C. & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. Free PDF downloadable from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

Golden, J. 2019. AI has a bias problem: This is how we can solve it. *WorldQuant Predictive*. [Online] Available from https://www.weforum.org/agenda/2019/01/to-eliminate-human-bias-from-ai-we-need-to-rethink-our-approach [Accessed 2022-05-21].

Green, B. 2021. The contestation of tech ethics: A sociotechnical approach to technology ethics in practice. *Journal of Social Computing,* 2(3):209-225. [Online] Available from https://scholar.google.co.za/scholar_url?url=https://arxiv.org/pdf/2106.01784&hl=en&sa=X&ei=I6lCYsybCpLeyQSQoL7oCQ&scisig=AAGBfm3XxbQ_Bh4Lj3XYZU2gWzBvEN-wDQ&oi=scholarr [Accessed 2022-02-15].

Hermstrüwer, Y. & Werkmeister, C. 2016. Using the Internet's fast lane: Exemptions from Net neutrality. *Cambridge International Law Journal*, May

2016. [Online] Available from http://cilj.co.uk/2016/05/12/using-internets-fast-lane-exemptions-net-neutrality/ [Accessed 2022-04-29].

Holling, C.S. & Meffe, G.K. 1996. Command and control and the pathology of natural resource management. *Conservation Biology,* 10(2):328-337. https://doi.org/10.1046/j.1523-1739.1996.10020328.x

Hopper, E. 2020. Maslow's hierarchy of needs explained. ThoughtCo, February 2020. [Online] Available from https://www.thoughtco.com/maslows-hierarchy-of-needs-4582571 [Accessed 2022-04-01].

Hopster, J. 2021. What are socially disruptive technologies? *Technology in Society*, 67:1-8. [Online] Available from https://www.sciencedirect.com/science/article/pii/S0160791X21002256?via%3Dihub [Accessed 2022-03-24].

Horwitz, G. 2019. Understanding net neutrality. *ThirdWay.org*, April 2019. [Online] Available from https://www.thirdway.org/memo/understanding-net-neutrality [Accessed 2022-04-30].

International Committee of the Red Cross (ICRC). 2018. Ethics and autonomous weapon systems: An ethical basis for human control? (Chapter 18). In: Stückelberger, C. & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. Free PDF downloadable from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

iWMS. 2020. Building an ethical supply chain. *WMS Supply Chain Solutions*. [Online] Available from https://iwms.co.za/2020/01/31/building-an-ethical-supply-chain/ [Accessed 2022-04-27].

Kulkarni, A. 2021. Bias and AI in machine learning: Sources and solutions. *Lexalytics*. [Online] Available from https://www.lexalytics.com/lexablog/bias-in-ai-machine-learning [Accessed: 2022-03-17].

Lee, K-F. 2021. The third revolution in warfare. *The Atlantic*, September 2021. [Online] Available from https://www.theatlantic.com/technology/archive/2021/09/i-weapons-are-third-revolution-warfare/620013/ [Accessed 2022-03-22].

Lee, N., Resnick, P. & Barton, G. 2019. Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. *The Brookings Institution*, May 2019. [Online] Available from https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/#footnote-9 [Accessed 2022-04-01].

Legal Information Institute. Not dated. *Intellectual Property*. Cornell Law School. [Online] Available from https://www.law.cornell.edu/wex/intellectual_property [Accessed 2022-03-30].

Makaula Z. 2021. Information and communication technologies (ICT) towards agricultural development in rural areas: Case of smallholder farmers in Umzimvubu local municipality of the Eastern Cape Province in South Africa.

South African Journal of Agricultural Extension; 49(1):81-90. [Online]. Available from: https://sajae.co.za/article/view/10779 [Accessed 2022-04-19].

McCammon, C. 2018. Domination. *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.) [Online] Available from https://plato.stanford.edu/archives/win2018/entries/domination/. [Accessed 2022-05-02].

McKenna, M. 2019. Three notable examples of AI bias. *AI Business*, October 2019. [Online] Available from https://aibusiness.com/document.asp?doc_id=761095# [Accessed 2022-05-17].

Miller, S. 2020. Freedom of political communication, propaganda and the role of epistemic institutions in cyberspace. In: M. Christen et al. (eds.), *The Ethics of Cybersecurity*, The International Library of Ethics, Law and Technology 21, https://doi.org/10.1007/978-3-030-29053-5_11 [Accessed 2022-03-30].

Moyer, D.M. & Hedden, S. 2020. Are we on the right path to achieve the sustainable development goals? *World Development*, 127, March 2020. [Online] Downloadable from https://www.sciencedirect.com/science/article/pii/S0305750X19303985 [Accessed: 2022-03-15].

National Planning Commission of South Africa. 2012. Chapter 6. An integrated and inclusive rural economy. In: *National Development Plan 2030: Our future – make it work*. Pretoria: Government Printer. [Online] Available from https://www.gov.za/issues/national-development-plan-2030 [Accessed 2022-03-24].

Omowole, A. 2021. Research shows AI is often biased. *World Economic Forum*, July 2021. [Online] Available from https://www.weforum.org/agenda/2021/07/ai-machine-learning-bias-discrimination/

Parschau, C. & Hauge, J. 2020. Is automation stealing manufacturing jobs? Evidence from South Africa's apparel industry. *Geoforum*, Vol. 115:120-131. [Online] Available from https://www.sciencedirect.com/science/article/pii/S0016718520301871 [Accessed 2022-03-31].

Paturel, A. 2014. Game theory: The effects of video games on the brain. *Brain & Life*, June/July. [Online] Available from https://www.brainandlife.org/articles/how-do-video-games-affect-the-developing-brains-of-children [Accessed 2022-04-21].

Pazzanese, C. 2020. Trailblazing initiative marries ethics, tech. *The Harvard Gazette*, October 16, 2020. [Online] Available from https://news.harvard.edu/gazette/story/2020/10/experts-consider-the-ethical-implications-of-new-technology/ [Accessed: 2022-02-15].

Phoenix, J. 2020. *The pros and cons of artificial intelligence*. [Online] Available from https://understandingdata.com/artificial-intelligence-advantages-and-disadvantages/ [Accessed 2022-03-24].

Pozzana, I. & Ferrara, E. 2020. Measuring bot and human behaviour dynamics. *Frontiers in Physics*, April 2020. [Online] Available from https://www.frontiersin.org/articles/10.3389/fphy.2020.00125/full [Accessed 2022-03-31].

Puddephat, A. 2016. Freedom of expression and the internet. *UNESCO*. [Online] Available from https://unesdoc.unesco.org/ark:/48223/pf0000246670 [Accessed 2022-04-16].

Rochel, J. 2018. Towards a just Internet: A republican Net neutrality. In: Stückelberger, C, & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. [Online] Available from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

Saruta, F. 2021. Effects of vertical integration on Internet service providers' zero-rating choice. *MPRA Paper No. 110288*, November 2021. [Online] Available from https://mpra.ub.uni-muenchen.de/110288/. [Accessed 2022-04-29].

Sax, M. 2016. Big data: Finders keepers, losers weepers? *Ethics and Information Technology,* 18:25-31. [Online] Available from https://link.springer.com/article/10.1007/s10676-016-9394-0 [Accessed: 2022-03-28].

Schuelke-Leech, B. 2018. A model for understanding the orders of magnitude of disruptive technologies. *Technological Forecasting and Social Change*, 129:261-274. [Online] Available from https://www.sciencedirect.com/science/article/abs/pii/S0040162517314610 [Accessed 2022-03-28].

Scott, B. Heumann, S. & Kleinhans, J. 2015. Landmark EU and US net neutrality decisions: How might pending decisions impact internet fragmentation. *Global Commission on Internet Governance*, July, 2015. [Online] Available from https://www.cigionline.org/documents/949/no18.pdf [Accessed 2022-04-30].

Security.org. Not dated. *How secure is my password?* [Online] Available from https://www.security.org/how-secure-is-my-password/ [Accessed: 2022-03-18].

Shapiro, A. 2019. Autonomous weapon systems: Selected implications for International security and for Canada. *Library of Parliament*, Publication No. 2019-55-E. [Online] Available from https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/201955E [Accessed 2022-05-22].

Siwiki, B. 2021. How AI bias happens–and how to eliminate it. *Healthcare IT News*, November 2021. [Online] Available from https://www.healthcareitnews.com/news/how-ai-bias-happens-and-how-eliminate-it [Accessed 2022-04-01].

Smith, A. 2021. Creating a sustainable and ethical supply chain. *PlugandPlay*. [Online] Available from

https://www.plugandplaytechcenter.com/resources/creating-sustainable-and-ethical-supply-chain/ [Accessed 2022-04-27].

Smith, R. 2018. Five core principles to keep AI ethical (Chapter 5.2). In: Stückelberger, C. & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. [Online] Available from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

Stahl, B., Timmermans, J. & Flick, C. 2017. Ethics of emerging information and communication technologies. *Science and Public Policy*, 44(3), 369-381. [Online] Available from doi: 10.1093/scipol/scw069 [Accessed 2022-02-16].

Stevens, H. & Huys, I. Innovative approaches to increase access to medicines in developing countries. *Frontiers in Medicine*, 4:218. [Online] Available from https://www.frontiersin.org/articles/10.3389/fmed.2017.00218/full [Accessed 2022-03-30].

Stückelberger, C. 2018. Cyber society: Core values and virtues. In: Stückelberger, C, & Duggal, P. (eds.) *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. [Online] Available from http://hdl.handle.net/20.500.12424/169317 [Accessed: 2022-02-14].

Stückelberger, C. & Duggal, P. (eds.) 2018. *Cyber ethics 4.0: Serving humanity with values*. Geneva: Globethics. [Online] Available from http://hdl.handle.net/20.500.12424/169317 [Accessed 2022-02-14].

Suri, A. 2017. Artificial intelligence and the rise of economic inequality. *Towards Data Science*, August 2017. [Online] Available from https://towardsdatascience.com/artificial-intelligence-and-the-rise-of-economic-inequality-b9d81be58bec [Accessed 2022-04-01].

Triberti, S., Milani, L. Villani, D., Grumi, S. Peracchia, S., Curcio, G. & Riva, G. 2018. What matters is when you play. *Addictive Behaviors Reports* 8 (2018) 185-188. [Online] Available from https://doi.org/10.1016/j.abrep.2018.06.003https://reader.elsevier.com/reader/sd/pii/S235285321830035X?token=7653AF56CBA36446822DC7FE09D078B23E07F4649C1931F5C7127AF523E7C9D6C8E8A97755EB8B858A4DA5024596FDE1&originRegion=eu-west-1&originCreation=20220425153048 [Accessed 2022-04-25].

Turja, T., Särkikoski, T., Koistinen, P. & Mellin, H. 2022. Basic human needs and robotization: How to make deployment of robots worthwhile for everyone. *Technology in Society*, 68:1-10. [Online] Available from https://www.sciencedirect.com/science/article/pii/S0160791X22000586 [Accessed 2022-03-23].

Turner Lee, N., Resnick, P. & Barton, G. 2019. Algorithmic bias detection and mitigation. *Governance Studies, Brookings* (May 2019). [Online] Available from https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/   [Accessed 2022-05-17].

UK Government. 2018. *AI in the UK: ready, willing and able?* Paper 100 published by the authority of the House of Lords. [Online] Available from https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf. [Accessed: 2022-03-10].

UK Rehab. Not dated. Gaming addiction explained. *UK Rehab*. [Online] Available from https://www.uk-rehab.com/behavioural-addictions/gaming/ [Accessed 2022-04-22].

UNESCO. Not dated. *Artificial intelligence: Examples of ethical dilemmas*. [Online] Available from https://en.unesco.org/artificial-intelligence/ethics/cases [Accessed: 2022-03-15].

UNESCO. 2012. *Inclusive Knowledge Societies for Sustainable Development*. [Online] Available from https://www.un.org/en/development/desa/policy/untaskteam_undf/groupb_unesco_knowledge_societies.pdf [Accessed 2022-03-29].

UNESCO. 2016. How we could triple the availability of textbooks. *Global Education Monitoring Report*. [Online] Available from https://world-education-blog.org/2016/01/19/how-we-could-triple-the-availability-of-textbooks/ [Accessed 2022-03-30].

UNESCO. 2021. Promoting open access to information for all and supporting multilingualism. [Online] Available from https://en.unesco.org/ci-programme/open-access [Accessed 2022-03-31].

UNICEF Office of Research. 2018. *Policy guide on children and digital connectivity*. United Nations Children's Fund, New York. [Online] Available from https://www.unicef.org/esa/media/3141/file/PolicyLab-Guide-DigitalConnectivity-Nov.6.18-lowres.pdf [Accessed: 2022-02-28].

Valentiner, T. 2018. 8 guiding principles for building ethical global supply chains. GreenBiz. [Online] Available from https://www.greenbiz.com/article/8-guiding-principles-building-ethical-global-supply-chains [Accessed 2022-04-27].

Vallor, S., Raicu, I. & Green, B. 2020. Technology and engineering practice: Ethical lenses to look through. *The Markula Center for Applied Ethics at Santa Clara University*. [Online] Available from http://www.scu.edu/ethics/. [Accessed 2022-02-15].

VCEWeb. 2022. Maslow's hierarchy of needs. [Online video] Available from https://www.youtube.com/watch?v=ei26_TjBlhI [Accessed 2022-04-15].

Von der Heiden, J., Braun, B., Muller, K. & Egloff, B. 2019. The association between video gaming and psychological functioning. *Frontiers in Psychology*, July 2019. [Online]. Available from https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01731/full [Accessed 2022-04-25].

Wang, J-L., Sheng, J-R. & Wang, H-Z. 2019. The association between mobile game addiction and depression, social anxiety, and loneliness. *Frontiers in*

Public Health, Vol. 7:Article 247. [Online] Available from doi: 10.3389/fpubh.2019.00247 [Accessed 2022-04-25].

Westland, J. 2022. An introduction to supply chain management (SCM). *Project Manager*, February 2022. [Online] Available from https://www.projectmanager.com/blog/supply-chain-management [Accessed 2022-04-28].

Wihbey, J. 2019. Net neutrality. *Blackwell Encyclopedia of Sociology*. [Online] Available from https://camd.northeastern.edu/wp-content/uploads/2018/02/Net-Neutrality-entry_Blackwell-Encyclopedia-of-Sociology-1.pdf [Accessed 2022-04-30].

Wikipedia. Not dated. *Industrial revolution*. [Online] Available from https://en.wikipedia.org/wiki/Industrial_Revolution [Accessed: 2022-02-21].

Wikipedia. Not dated. *Digital revolution*. [Online] Available from https://en.wikipedia.org/wiki/Digital_Revolution [Accessed: 2022-02-21].

Wikipedia. Not dated. *Fourth industrial revolution*. [Online] Available from https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution [Accessed: 2022-02-21].

Wikipedia. Not dated. *Net neutrality*. [Online] Available from https://en.wikipedia.org/wiki/Net_neutrality [Accessed 2022-04-29].

Wikipedia. Not dated. *Republic*. [Online] Available from https://en.wikipedia.org/wiki/Republic [Accessed: 2022-04-01].

Wikipedia. Not dated. *Sustainable development goals*. [Online] Available from https://en.wikipedia.org/wiki/Sustainable_Development_Goals [Accessed: 2022-02-22].