Maynooth University
National University of Ireland Maynooth

Artificial Intelligence

The Human Autonomy and Ethical Considerations of Advancing Intelligent Systems and Machines

By

Mr. William Owen O'Donovan

C.Dip.AF, BABS (Hons)

Submitted with a view to obtain the Master of Arts Degree, Maynooth University

Department of Philosophy

JULY 2024

Head of Department: Professor Michael Dunne

Supervisor: Doctor David O'Brien

Author's Declaration

I hereby declare that this project represents my own work and has not been submitted, in whole or in part, by me or by another person, for the purpose of obtaining any credit or grade. I agree that this project may be made available to future students of the College.

Signed: _____

Name: William Owen O'Donovan

Date: _____

Dedication

To my wife Julie and my daughter Jennifer,

for their enduring patience and support.

Acknowledgements

To those who lectured, guided and assisted throughout my studies.

Philosophy Department

Dr. Susan Gottlöber, for her sage guidance supporting my degree choice

Dr. David O'Brien, my thesis supervisor, for his wise counsel and enlightenment

Professor Michael Dunne

Dr. William Desmond

Economics Department

Professor Rowena Ann Pecchenino

Associate Professor Dr. Aedín Doris

Dr. Irene Mosca

Dr. Tuvana Pastine

Dr. Paul O'Sullivan

Sociology Department

Professor Colin Coulter

Professor Laurence Cox

Professor Seán Ó'Riain

School of Law and Criminology

Dr. Lynsey Black

Dr. John Reynolds

Library and IT Department

For their assistance and guidance

Abstract

In the modern era, the Artificial Intelligence (AI) concept existed since Alan Turing conceived Machine Learning in 1935.[1] AI, as a term, came into more common and public discourse in early 2023, when Elon Musk, industry leaders and others called for more regulation. This marked a turning point for more public discussion. Popular views range from, AI bringing enormous benefits, to AI representing an existential threat for mankind. Understanding the dichotomy between AI's promise and its perils is vital for humankind, human dignity, autonomy and freedom. As a technology AI disrupts society and its norms. While acknowledging the benefits; there are moral and ethical issues, along with serious threats. To bring an understanding of AI itself, human intelligence and consciousness; as well as computer-based intelligence and consciousness must be considered. The moral standings of humans and AI must be evaluated and assessed if they are or could be equal. Should AI become superior to human intelligence, "Strong AI," humans may become subservient to it, then fundamental human rights questions arise. Humans and AI companies maintain that they will always be in full control of Weak and Strong AI, however the reality might suggest the contrary, if the underlying ways of human life are altered. In its current iteration, Weak AI has and will continue to irreversibly change people's behaviours, humans become reliant on AI to the degree of being unable to conduct ordinary aspects of everyday life in its absence. While multiple problems exist in the world, AI further harms human autonomy and is socially disruptive. 'Big Data', the 'Internet of Things' and 'always connected' systems, are part of AI and are driven by AI. Its power lies not just with these technologies, but also with the algorithms' subliminal effects. The relatively limited number of Big-tech companies sway governments, using their enormous economic and political clout. This thesis sets out to demonstrate how, while Weak AI may not be a true existential threat to mankind; there are nevertheless serious moral and

[1] The artificial intelligence concept originates thousands of years ago, when automatons (from ancient Greek, meaning "acting of one's own will") were made. These were mechanical and moved independently of humans. Around 400 BCE a mechanical pigeon was created by a friend of the philosopher Plato. In 1495 Leonardo Da Vinci built a knight, it used cables and pulleys to sit, stand, turn its head, cross its arms and lift up its metal visor. Tableau.com, nd

ethical risks and dangers that overshadow its unquestionable benefits. However, strong governments can resolve these dilemmas.

Table of Contents – Summary

Table of Contents – Detailed

Introduction

Technology has aided mankind for millennia, from simple tools fashioned from rocks and sticks to today's sophisticated developments. Artificial Intelligence (AI) represents a towering achievement, combining man's creativity and intellectual prowess.

> …[AI] is a name to be proud of …it may allow us to; "virtually eliminate global poverty, massively reduce disease and provide better education to almost everyone on the planet."[2]

AI has brought undoubted benefits in areas such as education, healthcare, travel, finance and communications. There is excited anticipation of AI's promise, the further advances for humans' daily and professional lives, and the strengthening economic efficiencies. There is a belief that AI will bring greater good, mankind's happiness will be increased, and better comforts will follow.

Privacy and human autonomy problems, that already exist, will increase with AI. Medical advances have challenged policy makers and thinkers, life and death decisions sometimes being placed into the hands of medically eager advocates. AI is no different as it evolves, the undermining of mankind's humanity and autonomy is a real risk from those envisioning advancements through AI.

In the modern digital environment, AI relies on a combination of big data and the "always on, always available and always connected" internet of things. Presently, a relatively small number of large companies provide or use AI services at scale. Their economic and political power and their use of algorithms will be assessed. The malign and malicious uses of AI will be investigated. Their effects on human dignity, autonomy, freedoms and ethical considerations, with the societal and political implications of AI as it grows, will be discussed. The concepts of the current intelligent 'Weak AI' and the possible future of conscious 'Strong AI' will be

[2] Quoting; Shoham et al. 2018, Brynjolfsson and McAfee 2016, and Anderson, Rainie, and Luchsinger 2018 in Müller, 2023

evaluated, principally focusing on 'Weak AI'. The essential digital ethical principles of transparency, accountability and regulatory frameworks for AI, will be examined.

This treatise comprises nine chapters. The first four chapters lay the important groundwork. Defining AI as a technology in Chapter One and a brief discussion on Disruptive Technology in Chapter Two. Chapters Three and Four respectively consider the human mind and intelligence, and AI's 'mind' and power.

Using examples, Chapters Six to Nine are the core of the treatise, discussing the current and potential negative effects of AI for humankind. Ethical problems that existed for mankind before AI, become greater and continue to be a worry.

Chapter Six investigates how aspects of AI are and can be used, undermining human autonomy. Problems of inequality, life and death scenarios, and the intrusions into individuals' lives are raised.

Chapter Seven reviews AI's socially disruptive characteristics and consequent harms. Its algorithms manipulate human behaviours for commercial and malicious intents. The ways that AI strikes at democratic principles and processes, with increasing problems for freedom of expression are examined.

Chapter Eight briefly discusses warfare. AI enables conflicts to become more industrial in scale, facilitating more lethal ruthlessness and the developments of new weaponries.

The final, Ninth Chapter, sets out legal, regulatory, political and economic dilemmas and possible responses. Regulators, politicians and policy makers are struggling to keep pace with AI developments and are strained responding to powerful AI firms.

AI could evolve to a point where it overtakes mankind's ability to manage it, moving from the existing 'Weak AI' to the possibility of 'Strong AI'. Thereby further exposing the dichotomy of AI's promises and its hazards. A March 2023 Open Letter, from big-tech AI firms, asked "Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilisation?"[3] "It is conceivable that these new technologies will unsettle our

[3] Bengio, Musk, et al, 2023

deepest sense of what it is to be human,"[4] raising existential issues and metaphysical concerns.

This thesis identifies and explores numerous AI problems, of human autonomy, morals and ethics. They are examined against the background of philosophy, politics and economics, the findings of the thesis are set out before concluding.

In considering the development of AI and the future of humankind, it is worth recalling Machiavelli's maxim[5] *Salus populi suprema lex*[6], which offers sound and morally strong advice.

Throughout technical terms are avoided as far as possible, nevertheless some are required. Explanations and details are provided in footnotes, appendices and the glossary. The appendices also provide further discussion and supplementary information on some topics, based on the chapters' titles.

[4] Wolf, 2023, 172
[5] Kenny, 2010, 705
[6] "The welfare of the people is the highest law"

1 What is Artificial Intelligence?

Before further discussion, it will be valuable to examine AI's underlying technology and computing capabilities.

1.1 Algorithms

All computers operate through instructions relevant to the purpose or task. These instructions, software programmes or algorithms, are written by specialists. The computer scientist Harold Stone defines, "An algorithm is a set of rules that precisely define a sequence of operations."[7] It is a "…procedure used for solving a problem or performing a computation… an exact list of instructions that conduct specified actions step by step…"[8] Algorithms are powerful, conducting many functions including for databases, encryption, computing and facial recognition.[9, 10]

1.2 Classical Computing[11]

Since their invention, computers with large screens and keyboards have become quicker and smaller allowing much more complex operations. "… today's smartphones have the computing power of a military computer from 50 years ago that was the size of an entire room."[12] In the 1970s this exponential growth was labelled "Moore's Law"[13]

> …we could squeeze twice as many transistors on an integrated circuit board every 24 months. Given that the electrons have less distance to travel, the circuits run twice as fast, providing an overall quadrupling of computational power.[14]

[7] Lum and Chowdhury, 2021
[8] Gillis, 2024
[9] Gillis, 2024
[10] For a fuller explanation see Appendix 1
[11] See Appendix 1 for a brief discussion on Classical Computing, Quantum Computing and AI definitions
[12] Marr (a), 2021
[13] Gordon Moore was the co-founder of the Intel Corporation, the microchip manufacturer
[14] Tuescher, 2004, 388

In 2022, the Oak Ridge National Laboratory, USA, announced the "unprecedented" power of its new 'Frontier' Supercomputer, performing at least 1,000,000,000,000,000,000 calculations per second.[15, 16]

1.3 Quantum Computing[17]

Even with phenomenal technological strides, there remain problems that classical computers cannot solve. The next stage, quantum computing, relies on Quantum Technologies which "… use the behaviour of matter and light that is normally only observed at very small scales."[18] "Quantum computers are not intended to replace classical computers, but to be a different tool we will use to solve complex problems that are beyond the capabilities of a classical computer."[19]

Estimates vary of quantum computing's potential power. "IBM's computer Deep Blue defeated chess champion, Garry Kasparov in 1997. It gained a competitive advantage examining 200 million possible moves each second. A quantum machine would be able to calculate 1 trillion moves per second!"[20] Google researchers claimed in 2019, that their quantum computer performed a calculation in 3 minutes 20 seconds, that would take about 10,000 years on a state-of-the-art conventional supercomputer.[21]

1.4 AI Definitions[22]

There are different definitions of AI, with two mentioned here:

The UK Government's 2023 policy paper 'A pro-innovation approach to AI regulation':

> …'products and services that are 'adaptable' and 'autonomous.' The 'adaptability' of AI refers to AI systems, after being trained, often developing

[15] 1^{18} (or One to the power of 18)
[16] Conover, 2022
[17] See Appendix 1 for a brief discussion on Classical Computing, Quantum Computing and AI definitions
[18] UK Parliament, 2017
[19] Marr (b), 2021
[20] Marr (b), 2021
[21] Chang, 2023
[22] See Appendix 1 for a brief discussion on Classical Computing, Quantum Computing and AI definitions

> the ability to perform new ways of finding patterns and connections in data that are not directly envisioned by their human programmers. The 'autonomy' of AI refers to some AI systems that can make decisions without the intent or ongoing control of a human.[23]

The Alan Turing Institute:[24]

> the design and study of machines that can perform tasks that would previously have required human (or other biological) brainpower to accomplish.[25]

In summary AI is autonomous (absence of humans), versatile (adaptive) and identifies data relationships beyond human abilities (not directly envisioned by their human programmers).

1.5 Powering AI – Supercomputers and Quantum Computing

The brief preceding outline illustrates the computing capability enabling AI. The UK Government describes its new research supercomputer as being "…made up of thousands of "state-of-the-art" graphical processing units and will be able to train the Large Language Models[26] that are at the forefront of AI research and development today."[27] The supercomputers' power is well beyond the comprehension and imagination of mankind in general and perhaps the broader science community.

The potential of quantum computing is such that it moves into the realm of fanciful science fiction. Yet it is here and is becoming standard, in large corporations.[28] "Independent of artificial intelligence, quantum computing is going to be a big deal.

[23] Gajjar, 2023, 7
[24] Alan Turing, an English mathematician and theoretical computer scientist (see Glossary)
[25] Gajjar, 2023, 7
[26] Large Language Model, A type of Foundation Model that is trained on vast amounts of text to carry out natural language processing tasks. During training phases, Large Language Models learn parameters from factors such as the model size and training data
[27] Gajjar, 2023
[28] "… IBM, Microsoft, and Google, and many compelling market entrants such as IonQ and D-Wave Systems, provide [quantum computing] capabilities that can be used today", Reichental, 2023

However, coupled with AI, it will be a gamechanger… It's probably closer than you think and it's likely going to be a very big deal."[29]

1.6 Weak AI and Strong AI

Going further Arkoudas & Bringsford[30] review the 'Total Turing Test'[31] proposed by S. Hanard,[32] which moves from a bodiless computer to "robots able to operate in the physical environment in a way that is indistinguishable from the behaviours manifested by embodied human persons …" This is Weak AI, that is AI that acts intelligently without knowing whether it is intelligent.

Strong AI on the other hand is the field devoted to building "persons." A more ambitious view of Strong AI is proposed:

> The fundamental goal [of AI research] is not merely to mimic intelligence or produce some clever fake... AI wants only the genuine article: *machines with minds*, in the full and literal sense.[33]

1.7 Does Strong AI Exist, Can AI Strong Exist?

Strong AI is a future possibility, one which would challenge humanity. While Weak AI is regarded as intelligent, Strong AI would possess more human traits, having consciousness and emotions. For the time being though Strong AI remains in the realm of possibility.

The UK Government describes Strong AI as:

> Artificial General Intelligence: Sometimes known as General AI, Strong AI or Broad AI, this often refers to a theoretical form of AI that can achieve human-level or higher performance across most cognitive tasks.[34]

[29] Reichental, 2023
[30] Arkoudas & Bringsford, in Frankish and Ramsey, 2014, 35
[31] The Turing Test is discussed later
[32] Quoting from S Harard, *The Symbolic Ground Problem*, 1990, 335-46
[33] Quoting from J. Haugeland, *Artificial Intelligence: The Very Idea*, 1985, 2
[34] Gajjar, 2022

The Chinese Room thought experiment,[35, 36] conceived by J Searle,[37] posits a situation whereby a computer and a human can behave as if they both understand the Chinese language.[38] Searle argues, the computer does not have "understanding" so it is not "thinking" and, since it does not think, it does not have a "mind". Consequently, what appears to be Strong AI is not in fact. Two other researchers, Gödel and Dreyfus, concur rejecting the Strong AI notion in principle.[39]

Expert commentators say that Strong AI does not and cannot exist. "While AI researchers in both academia and private sectors are invested in the creation …, it only exists today as a theoretical concept versus a tangible reality."[40] "Strong AI, … is capable of learning, thinking and adapting like humans do. That said, Strong AI systems don't actually exist yet."[41]

Philosopher Hubert Dreyfus argues that "… computers, who have no body, no childhood and no cultural practice, could not acquire intelligence[42] at all …human knowledge is partly tacit, and therefore cannot be articulated and incorporated in a computer program…"[43]

There is a body of opinion suggesting that Strong AI cannot ever exist, the area is controversial. Stringent tests have been laid down for Strong AI. IBM says:

> If researchers are able to develop Strong AI, the machine would require an intelligence equal to humans; it would have a self-aware consciousness that has the ability to solve problems, learn, and plan for the future… Strong AI aims

[35] Copeland, 1993, 123-25
[36] Arkoudas K and Bringsford S, in Frankish and Ramsey, 2014, 72-75
[37] J. Searle, *Minds, Brains and Programmes,* 1980
[38] Following the instructions within its programme, the computer can accept Chinese characters and produce an output with other Chinese characters. In doing so it passes the Turing Test. Creating a similar scenario with a human who has no knowledge of Chinese, a copy of the computer programme, pencils, papers and so on, the human could also achieve the same feat thereby passing the Turing Test.
[39] Bringsjord *et al,* 2022
[40] IBM (b), nd
[41] Glover, 2024
[42] It should be noted that although disembodied, computers do in fact possess intelligence
[43] Fjelland, 2020

> to create intelligent machines that are indistinguishable from the human mind.[44]

The idea that Strong AI would be "indistinguishable from the human mind" is conceptually challenging, presenting philosophical, moral and ethical dilemmas. Whether Strong AI can or should be allowed to exist is outside the scope of this thesis, nevertheless references are made to Strong AI.

Weak AI is functional, it carries out tasks at such speed that it outclasses the human mind. It conducts complex calculations, stores immense amount of data, makes predictions and is now an essential day to day tool for governments, organisations and individuals alike. Nevertheless, though apparently becoming stronger, Weak AI remains as such. The considerations, discussed later, relating to Weak AI are problematic, confronting how humans live and interact. They are sometimes grave in their nature and require a full understanding for humans to progress in an AI world.

[44] IBM (x), nd

2 Disruptive Technology

2.1 Technology

2.1.1 What is Technology?

"Technology is the application of scientific knowledge to the practical aims of human life or, as it is sometimes phrased, to the change and manipulation of the human environment."[45] From the simple handheld tools of prehistoric man to today's sophisticated mechanical and computerised machines, technology has shaped mankind, the world and its environment. It has brought good, increasing food production, industrial productivity and travel. It has also wrought pollution, damaged the environment and wildlife, made warfare more deadly and fostered the pursuit of excessive wealth.

Ultimately, the fusion of humans and their intelligence with technology, has shaped the world and human life beyond recognition from what they were in prehistoric ages. Humans and technology combined to become systems that are often specialist. Human's industries, sports and pastimes revolve around technology to a greater or lesser extent. The expert racing driver controls a complex machine capable of high speeds, humans with tools and machines built extraordinary edifices.

Technology is a fundamental part of human existence.

2.1.2 Disruptive Technology

'Disruptive technology' is an innovation that significantly alters the way societies, consumers, industries and businesses operate. Its attributes are recognisably superior, sweeping away old systems and habits, bringing advantages that are immediately obvious, a desire for the next technology generation[46]and creating a smoother pathway for the new vision.[47] AI is a disruptive technology changing

[45] Britannica, 2024

[46] "Generation" in terms of technology refers to the current level of development, the next level of development is called the next generation. Generally speaking, moving from one generation to the next represents a step advancement in the technology, without necessarily altering the underlying technology

[47] Smith, T., 2022

norms and behaviours, resulting in a paradigm shift,[48] "AI may also have a highly disruptive effect on the economy and society."[49]

Disruptive technologies were not always welcomed. Automation was first recorded in 16th century English stocking production. Luddism, an 18th century English social movement, opposed industrialisation destroying machines, the symbols of the unemployment threat to skilled craftsmen.[50]

Robots[51] were first used for repetitive tasks in vehicle manufacturing, in the 20th century, increasing output, standardising quality and controlling costs.[52]

2.2 Effects of Technological Changes

The technology underlying AI was discussed in the previous chapter. More generally technology is "the application of scientific knowledge to the practical aims of human life."[53, 54] Over millennia it has brought enormous change and benefits, from the wheel to today's longer life expectancy, educational techniques and space travel. Risks, disasters and other dangers are mitigated using technology.[55] Accompanying these changes has been economic growth and better standards of living:

> Economists have become accustomed to associating long-term economic growth with technological progress; it is deeply embedded in the main

[48] The motor car, electricity and television were disruptive technologies in their own times. Smith, T., 2022

[49] Referencing the EU, in Szczepański, 2019

[50] See Appendix 2 for a short discussion on Luddism

[51] A robot is "a machine controlled by a computer that is used to perform jobs automatically." Cambridge Online Dictionary, nd

[52] Fleming, 2020

[53] Augustyn, 2024

[54] Other definitions include (i) The use of science in industry, engineering, etc., to invent useful things or to solve problems; A machine, piece of equipment, method, etc., that is created by technology. Britannica Dictionary (a), nd (ii) The application of this knowledge for practical ends; The terminology of an art, science, etc.; A scientific or industrial process, invention, method, or the like; The sum of the ways in which social groups provide themselves with the material objects of their civilization. Dictionary.com nd

[55] Korstanje and Skoll, 2013, 15

> message of the Solow-inspired growth models,[56] which treated technological change as exogenous...[57]

Combinatorial digital innovation demonstrates the significant impact on AI's evolution:

> … the practice of using components of different digital technologies and trends together to uncover new or better value. To do it effectively, three actions must be taken: First, combine suitable technologies and trends. At the same time, cluster emerging technologies and trends that fit naturally together. And finally, complement initiatives with emerging technologies that can have an impact on secondary areas.[58]

Language is a defining characteristic of humanity as a social species, therefore, it is no accident that the most "revolutionary technologies humans have invented are those of communications," from writing to the internet.[59] ICT[60] is an example of combinatorial digital innovation, which enables the internet and has played a significant role in the enabling of rapid and global communications. Person to person communication across the globe is now a common feature of modern life.

2.3 The Law of Disruption

Larry Downes described the "Law of Disruption" in 2009:

> … technology changes exponentially, but social, economic and legal systems change incrementally and this law was becoming a simple but unavoidable principle of modern life.[61]

The Pacing Problem similarly describes how legislators lag technological developments.[62]

---

[56] The economist Solow's growth model focuses on long-run economic growth
[57] Mokyr, 2005
[58] Thierer, 2018
[59] Wolf, 2023, 296
[60] ICT, Information and Communications Technology. It is the combination of computers, information systems and telecommunications systems, which has facilitated the growth and ease of communications in the modern era
[61] Thierer, 2018
[62] See the Pacing Problem later

AI, through current experience and future potential, fundamentally alters the behaviour of societies and businesses. It is disruptive, changing individuals' behaviours, attitudes and a 'bottom up' societal shift.[63]

[63] See the later Social Disruption chapter

3 The Human Mind

Much research has been dedicated to comprehending human intelligence using the creation of models, tests and measurements. The human mind must be recognised philosophically, in its humanness and in its fullness of personality, morals and ethics.

There has also been continuous work to expand the study of intelligence to non-human systems.[64, 65] Before discussing AI, it is important to acknowledge human intelligence and consciousness.

3.1 Philosophy and Human Intelligence[66]

Philosophical disciplines seek a "reasoned logical belief," bringing intelligence to its ultimate potential. Philosophy has an important role for interdisciplinary dialogue, build bridges and help other more concrete disciplines.[67] From the ancient world to the present-day philosophers led human thought, the development of education in many fields and of intelligence itself.

3.1.1 Philosophers' Views of Intelligence

Intelligence is a process, or an innate capacity to use information to respond to ever-changing requirements. It is a capacity to acquire, adapt, modify, extend and use information to solve problems. Intelligence is the ability to cope with unpredictable circumstances.[68]

> … much of our knowledge is tacit; it is not expressible in language… "We know more than we can tell."[69] This is true not only for innate knowledge, but also for those internalised skills which come from observing and copying others.[70]

[64] Palanca-Castan *et al*, 2021, 1
[65] "There is no reason why we cannot reverse engineer the human brain and essentially copy its design" Tuescher, 2004, 397
[66] For a fuller discussion on Philosophy and Human Intelligence see Appendix 3
[67] Palanca-Castan *et al*, 2021, 2-3
[68] MacFarlane, 2013
[69] Quoting Michael Polanyi, in MacFarlane, 2013
[70] MacFarlane, 2013

Establishing the need for humans to be clear and precise in how they think, Quist quotes Aristotle's Nicomachean Ethics,[71] "… it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits…"

Plato considered intuitive reason as the highest form of human intelligence and spoke of intuition as a direct contemplation of the truth.[72] He differentiated between two ways of using reason or intelligence:

> Discursive reason, *dianoia διάνοις,* is exercised to confront theses. … slow use of reason… before reaching a conclusion.
>
> Intuitive reason, *nous νοῦς* or *νόος,* is a quick use of reason… the simplified form of deductive reasoning.

Aristotle discussed three types of souls:

> The vegetative soul, typical of plants. Its functions were growth and nutrition.
>
> The sensitive soul, typical of animals capable of locomotion. Along with the previous functions, it was capable of locomotion and sense perception.
>
> The intellectual soul corresponded to humans. Along with the previous functions, it had the ability to reason.

Aristotle, Plato, Roman and medieval philosophers, only assigned the rational soul to human beings. For them, intelligence and reason were almost synonymous.[73]

3.2 Human Intelligence Theories[74]

Human intelligence is complex and diverse. All individuals demonstrate varying strengths and weaknesses in different intelligence areas; ultimately AI will be expected to feature all human intelligence types.

[71] Quist, 2022
[72] Palanca-Castan *et al*, 2021, 3
[73] Palanca-Castan *et al*, 2021, 3
[74] For a fuller exploration on Intelligence Types and Theories, see Appendix 3

3.2.1 Intelligence Types

Eight 'Multiple Intelligence' types have been described,[75] which everybody holds at some level. Ranging from Verbal skills to Naturalistic skills, they relate to how individuals prefer to learn and process information.

3.2.2 Triarchic Theory

The 'Triarchic Theory',[76] posits that intelligence comes in three forms; Analytical intelligence, Creative intelligence[77] and Practical intelligence.

3.2.3 Cattell, Horn, and Carroll, (CHC), Theory

This espouses that there are three strata of intelligence, which are hierarchically related to each other,[78] comprising narrow, broad and general abilities.

3.2.4 The Psychological Views

"Psychology is the scientific study of the mind and behaviour. Psychologists are actively involved in studying and understanding mental processes, brain functions, and behaviour."[79] "Psychologists examine the relationships between brain function and behaviour, and the environment and behaviour…"[80]

They define intelligence as "the ability to derive information, learn from experience, adapt to the environment, understand, and correctly utilise thought and reason"[81] and "Intelligence is one's ability to learn from experience and to adapt to, shape, and select environments."[82] Any attempts to be specific about an individual's intelligence level and type are complicated, requiring scientific methods of assessment.

> At one time, intelligence research consisted primarily of statistical analyses of individual differences in scores on intelligence tests. Today… intelligence

[75] Quoting Howard Gardner, in Mercer County Community College (MCCC), 2024
[76] Developed by Robert Sternberg, MCCC, 2024
[77] Creative intelligence arises later in the context of consciousness
[78] Sternberg, 2012
[79] Ohio State University, 2014
[80] American Psychological Association, 2013
[81] American Psychological Association, 2018
[82] Sternberg, 2012

is also being studied by cognitive psychologists, neuroscientists, cultural psychologists, and many others.[83]

3.2.4.1 Measuring Intelligence[84]

The 'Intelligence Quotient' (IQ) scores a range of cognitive abilities, rating an individual's intellectual capabilities and potential. IQ tests are also used to assist diagnosis of intellectual disabilities.[85, 86]

In the future AI would be expected to display an IQ superior to that of any human.

3.2.4.2 Emotional Intelligence[87]

Emotional Intelligence is an aspect of intelligence, which was defined by Thorndike[88] as a social intelligence or having "the ability to understand and manage men and women, boys and girls to act wisely in human relations."[89]

Emotional Intelligence helps to define every individual.

3.2.5 Other Intelligence Aspects

3.2.5.1 Neurodivergence and Autism

Neurodivergence is a nonmedical term that describes people whose brains develop or work differently.[90, 91] Some neurodivergent individuals can accomplish astonishing mental feats. Such abilities have potential in business and AI development. The Harvard Business Review[92] and Birbeck University (UK)[93] both report the benefits and contributions that certain neurodivergent individuals make in companies.[94]

---

[83] Sternberg, 2012
[84] Appendix 3 provides further discussion, the ranges of IQ scores and descriptions in tabular form as well as a bell curve illustration
[85] Cherry K, 2022
[86] IQ tests are also used to assist diagnosis of intellectual disabilities. Cherry K , nd
[87] See Appendix 3 for further information on Emotional Intelligence
[88] Kanesan and Fauzan, N, 2019
[89] Defined by Thorndike, in Kanesan and Fauzan, N, 2019
[90] They have different strengths and struggles, from those whose brains develop more typically. Some have identified medical conditions, which may be serious and life limiting, while others do not
[91] Cleveland Clinic, 2022
[92] Austin and Pisano, 2017
[93] Birbeck University, 2023
[94] See also Appendix 3 for further discussion on neurodivergence

3.2.5.2 *Biological Intelligence Bases*

Human intelligence is influenced by certain brain attributes. The neural efficiency of the brain increases, based on intelligence and specific abilities. While there is an apparent relationship between intelligence and brain size, it is not proven and whether brain size is a result of higher intelligence or *vice-versa* is not clear.[95] Environmental rather than heredity factors influence intelligence and there is no racial link to higher or lower intelligence.[96]

3.2.5.3 *Poverty of Stimulus*[97]

Bertrand Russell asked "[H]ow come that human beings, whose contacts with the world are brief and personal and limited, are nevertheless able to know as much as they do know?".[98] Children pick up cues about the world and language around them, building knowledge and enhancing their understanding. In the absence of sufficient information childrens' intellectual growth may be impeded. "Poverty of Stimulus" is "the assertion that natural language grammar is unlearnable given the relatively limited data available to children learning a language, and therefore that this knowledge is supplemented with some sort of innate linguistic capacity."[99, 100] Chomsky has "referred to "children's "biological endowment.""[101, 102]

[95] Sternberg, 2012
[96] Sternberg, 2012
[97] See also Appendix 12.3
[98] Bertrand Russell discussing Plato in, Nordquist, 2019
[99] Noam Chomsky in Psyno, nd
[100] "Poverty of Stimulus relates to children's hypothesis spaces, the data available to children, and how children leverage those available data. Importantly, Poverty of Stimulus is often tied to the fact that children seem to make constrained generalisations faster than we might expect, given the data available. This constrained generalisation then implies that children have some kind of prior knowledge that guides them towards the correct answer. The prior knowledge for a particular Poverty of Stimulus problem is generally believed to involve some kind of innate knowledge, and the big divide is whether any of that innate knowledge is language specific." Pearl, nd
[101] Lasnik & Lidz, 2017, in Pearl, nd
[102] Conversely however, though based on logical algorithms and using deep learning techniques, ChatGPT is (currently) bad at solving math and logic tasks in general. Zvornicanin, 2024 (ChatGPT is an AI chatbot with natural language processing (NLP) that allows human-like conversations to complete various tasks. The generative AI tool can answer questions and assist with tasks... ZDNET website, nd.

3.3 Human Personality – Consciousness, Ethics and Morals

The human mind functions in a way that is more complex and nuanced than intelligence. Some of these complexities will be discussed as they have direct relevance for AI.

3.3.1 Consciousness

Consciousness is perhaps one of the more abstract[103] and elusive human traits, yet all individuals are aware of their own. Understanding consciousness is crucial in the AI context. It provides humans with the abilities to be aware of themselves, aware of others, animals and the inanimate world around them. Descartes views consciousness as thoughts which are the "contents of our minds." "...the word 'thought' for Descartes applies to conscious experience of any kind."[104]

> It is I who have sensations,[105] or perceive corporeal objects as it were by the senses... this is what is properly called my sensation, precisely so regarded, is nothing but an act of thought.

Descartes also identifies three classes of his ideas "Of my ideas, some seem to be innate, some acquired, and some devised by myself."[106] He authenticates thoughts, viewing the utterances of others as having 'special status', which shows the uniqueness of an individual's thoughts and consciousness.[107] Having proved to his own satisfaction that he exists, Descartes asked "What am I, this I whom I know to exist?" His answer *Reg cogitans*[108, 109] "…in the (Descartes') *Meditations* the mind is used to validate itself"[110] Copeland provides another view; "a conscious entity is

[103] There are many examples of a driver or a pedestrian, on their journeys deep in thought or conversation, arriving safely at their destinations. During their journeys they are consciously unaware of junctions, traffic lights and obstacles. Yet paradoxically, their eyes, ears and brains process information so that they become aware and alert if a risk arises

[104] Kenny, 2010, 592

[105] These apparent sensations, possible in the absence of a body, are what later philosophers were to call 'sense data'. Kenny, 2010, 592

[106] Kenny, 2010, 593

[107] Kenny, 2010, 660

[108] "I am a thing that thinks"

[109] Kenny, 2010, 660

[110] Kenny, 2010, 596

one whose sensory interactions with the world customarily accompanied by qualia[111] of some sort."[112]

These concepts compromise the 'self',[113] an individual's personal traits or characteristics that set them apart from others. Forming part of the unique human character this individual 'self' is the basis for moral status discussed later.

3.3.2 Ethics and Morals[114]

Ethics and morals provide humans with guidance for questions like; What is good? What is right? How do we approach difficult questions? "A good purpose is essential if an action is to be morally good … The mere belief that one's purpose is good does not suffice to render an action morally correct."[115]

Morality is a set of norms and principles that govern our actions with respect to each other and the world around us, they are taken to have a special kind of weight or authority, grounded in some basic values. Ethics is generally understood to be the study of "living well as a human being." In Aristotle's *Nicomachean Ethics*, the aim of human beings is to exemplify human excellence of character. The sense in which we understand it here is that ethics is broader than morality.[116]

3.3.2.1 Emotions in Ethics and Morals

People enter emotional states when they make moral judgments, induced emotions have an impact on moral judgment, different emotions will have different effects.[117] Therefore, to kill or not to kill one person[118] will evoke strong emotions and moral dilemmas before the agent decides. Situations with strong emotions, cause making the morally 'correct' decision complex.[119] There are subjective forms of all the

[111] Subjective, conscious experiences (singular 'quale')
[112] Copeland, 1993, 171
[113] The 'Self'; the personality or character that makes a person different from other people; The combination of emotions, thoughts, feelings, etc., that make a person different from others. Britannica Dictionary (b), nd
[114] See also Appendix 12.3
[115] Discussion on Hegal in Kenny, 2010, 703
[116] Driver, 2022
[117] Rosati, 2022
[118] See the next section's discussion of the 'Trolley Problem'
[119] The resulting decision, in hindsight, may be morally correct, morally flawed or morally wrong, but in the decision-making moment the agent will attempt to be moral

major moral theories and objective versions of many. An objective standard of right holds that the agent must meet the standard[120, 121] to count as right or virtuous.

The agent is autonomous, making decisions based on judgement and emotions, attempting to make and to reach a moral and ethical outcome. However the dilemmas posed to the induvial are not always clear cut.

3.3.3 The "Trolley Problem"[122]

We intuitively believe that it is worse to kill someone than to allow the person die. Driver[123] raises the 'Trolley Problem' thought experiments, to illustrate structurally similar situations eliciting different intuitions about the morally correct course of action. We believe it is wrong to kill one person to save five others in the circumstances[124] of a surgeon killing one person to harvest organs for five others. It is acceptable though for a trolley driver to kill one person on a track, rather than divert to another track killing five.

In Kantian deontology[125] "The morally worthy action is in accordance with the Categorical Imperative, which requires an agent refrain from acting in a way that fails to respect the rational nature of other persons."[126] This is "an objective, rationally necessary and unconditional principle that we must follow despite any natural desires we may have to the contrary."[127] "Other philosophers, such as Hobbes, Locke and Aquinas, had also argued that moral requirements are based on standards of rationality."[128] This "rationality" argument would compel an objective view regarding the transplant surgeon and the 'Trolley Problem' and judge the decisions of the key agents accordingly.

[120] Meeting the standard is something 'objective', not dependent on the agent's psychological states
[121] See Appendix 3 for definitions of Subjective Standards
[122] See Appendix 3 for fuller detail on the two scenarios
[123] Referencing Foot, 1975, in Driver 2022
[124] From Thomson, 1976, 206, Driver, 2022
[125] Deontology is an ethical theory that uses rules to distinguish right from wrong. Deontology is often associated with philosopher Immanuel Kant. Kant believed that ethical actions follow universal moral laws, such as "Don't lie. Don't steal. Don't cheat." Source McCombs School of Business, nd
[126] In deontological ethics an action is considered morally good because of some characteristic of the action itself, not because the product of the action is good, Britannica, 2024
[127] Johnson and Cureton, 2022
[128] Johnson and Cureton, 2022

Such complexities and dilemmas stretch the human's morals and ethics. Whilst the optimum decisions were made, nevertheless people died. Recognising the rational for the decisions, a moral agent would though be emotionally conflicted by the deaths reasoning that some other act might have produced a better outcome.

3.3.4 Goodness

Kant reasons that there is both complete virtue and complete happiness, which he calls the highest good. There is a duty to promote the 'highest good', the sum of all moral duties, and we can fulfil this duty only if we believe that the highest good is a possible state of affairs.[129] Kant also holds the view that it is an unavoidable feature of human reason that we form ideas not only about the immediate and near-term consequences of our actions, but also about ultimate consequences.[130, 131]

For Kant "… everything has either a price or a worth." A price can have something else in its place as a fair exchange, if it is beyond price and is unexchangeable then it has worth. Kant further states:

> Morality, and humanity so far as it is capable of morality, is the only thing which has worth. Skill and diligence in work have a market price; with, imagination and humour have a fancy price; but fidelity to promises and benevolence based on principle (not on instinct) have an intrinsic worth.[132]

Humans using their intelligence, ethics, morals and consciousness evaluate their actions. In his *Ethics* Aristotle brings together humans' characteristics for judgement "an action counts as virtuous when and only when one holds oneself in a stable equilibrium of the soul, in order to choose the action knowingly and for its own sake" and "that the goodness is never in the action but only in the doer." "A person

[129] The morality of an action depends on the form of its maxim rather than its end or goal, he nevertheless claims both, that every human action has an end and that we are unavoidably concerned with the consequences of our actions

[130] Rohlf, 2010

[131] Interpreting Kant's reasoning implies humans are driven by the outcomes and consequences of their actions rather than the underlying morals. A person might conduct an act of generosity for the personal satisfaction of being seen by others as generous, rather than to achieve the more moral outcome of helping someone in need. For the recipient all that matters is the gift and not the underlying reasoning.

[132] Kenny, 2010, 700-01

of character is someone you can count on, because there is a human nature in a deeper sense... Someone with character has taken a stand in that fully mature nature, and cannot be moved all the way out of it."[133] Humans have deeply based morals and ethics, they possess moral status.[134]

This leads one to ask whether AI can have morals, ethics and goodness, if so whose? and who decides? Profit driven AI companies focus on the financial consequence of their creations, rather than the underlying morals. Customers rely on its creators for the full scope of human morals, ethics and goodness.

The dilemma for AI now becomes the aims or underlying philosophy[135] of the creators. Do they create AI with truly good and moral principles as Kant would suggest; or rather look towards the end, the result that is most desired by their organisations. The latter option gives rise to possible unintended consequences, such as how social media evolves.[136] Following good and moral principles, the creators may have designed or re-designed social media to be a much less hostile environment. Predicting AI's evolution and outcomes is a fraught activity. There are novel and challenging ethical questions relating to AI.

The next chapter further discusses human complexities and if they can transfer to the AI mind.

[133] Sachs, nd

[134] The question can AI possess moral status will be discussed in the next chapter

[135] In this context the "underlying philosophy" of creators does not necessarily mean that they consciously or deliberately or narrowly follow a particular philosophy. Rather it means that they follow some ambition, perhaps external to themselves, which determines how they set about their work

[136] The original creators of social media desiring that every user would have 'friends', may never have anticipated that social media would become platforms for disinformation, propaganda, offence and bullying type behaviour

4 The AI Mind

Huckins[137] poses the questions that this chapter sets out to consider. "Consciousness poses a unique challenge in our attempts to study it, because it's hard to define. It's inherently subjective."[138]

> AI consciousness isn't just a devilishly tricky intellectual puzzle; it's a morally weighty problem with potentially dire consequences. Fail to identify a conscious AI, and you might unintentionally subjugate, or even torture, a being whose interests ought to matter. Mistake an unconscious AI for a conscious one, and you risk compromising human safety and happiness for the sake of an unthinking, unfeeling hunk of silicon and code.[139]

This chapter considers if AI has or can have a mind in the manner that humans would recognise. Could AI have intelligence that resembles a human or merely instead mimics aspects of human behaviour?

"There is no reason why we cannot reverse engineer the human brain and essentially copy its design."[140] "… eventually we might have to come to terms with the fact that artificial intelligence is indeed conscious and rethink our moral and political lives in response to that development." It must be remembered however that while AI's intelligence probably will advance enormously, consciousness is quite a different proposition and may not transpire. AI is understood broadly as any kind of artificial computational system that shows intelligent behaviour, i.e., complex behaviour that is conducive to reaching goals.[141, 142]

[137] Huckins, 2023
[138] Quoting Liad Mudrik, a neuroscientist at Tel Aviv University, in Huckins, 2023
[139] Huckins, 2023
[140] Tuescher, 2004, 397
[141] Müller, 2023
[142] For further definitions see Appendix 4

The Turing Test[143] was put forward as a threshold for AI.[144] Turing described[145, 146] an experiment involving a human, a computer and an interrogator to say which is the human and the computer. If the interrogator judges no better than 50-50, then the computer is said to have passed the test and declared a 'thinking' machine. In 1950, Turing said "We may hope that machines will eventually compete with man in all purely intellectual fields."[147]

In 1997 the IBM computer programme 'Deep Blue', beat Gary Kasparov, the world chess champion, in an hour-long chess game. Machines had begun to compete with man at intellectual tasks.[148] DeepMind,[149] developed AI to play games, a general-purpose system for many two-person games, adapted specifically for Go,[150] chess, and shogi;[151] it was given no knowledge beyond the rules. In 2016, the program soundly defeated the world's best Go player.[152] After further development the new system iteration, AlphaZero,[153] was by far the simplest. To train its Go engine, DeepMind used five thousand of Google's "Tensor Processing Units"[154] for thirteen days. To do the same work on a desktop system, … would have to run it for seventeen hundred years.[155]

Turing foresaw this when he conjectured "one cannot programme a machine to play a better game than one plays oneself" was a "rather glib view." Although it sounds

[143] Put forward by Copeland, 1993, 36-37 and K Arkoudas and S Bringsford, in Frankish and Ramsey, 2014, 34-35
[144] Prior to the AI term becoming commonplace, terms like computers and machines were used
[145] Quoting Alan Turing's *Intelligent Machinery, from an article in "Computing Machinery and Intelligence", 1948,* Copeland, 2014, 37
[146] An experiment in which a human and a computer, are interrogated by a second human from whom they are hidden. The interrogator's challenge is to say which is the human and the computer. After a series of questions and answers if the interrogator judges no better than 50-50, then the computer is said to have passed the test. Such a computer should be declared a 'thinking' machine.
[147] Teuscher, 2004, ix
[148] Tuescher, 2004, 78
[149] Google's artificial-intelligence subsidiary
[150] A Chinese (territorial capture) strategy board game, Masters, 2024
[151] Japanese chess
[152] Lee Sedol
[153] It had a core algorithm so powerful that "you could give it the rules of humanity's richest and most studied games and later that day, it would become the best player there has ever been." An example of machine learning
[154] Computer chips specifically designed for neural network calculations
[155] Somers, 2018

right to say that “no animal can swallow an animal heavier than itself,” plenty can. Turing suggested, there might be no contradiction in a bad chess player making a chess program that plays brilliantly. One tantalising way to do it would be to have the program learn for itself.[156, 157]

To the ordinary human mind, the current capabilities and the potential power of AI are astonishing and ominous. Yet the human mind is complex, possessing attributes beyond AI’s and able to conduct multiple tasks simultaneously. In contrast AI is merely intelligent and while operating rapidly conducts its tasks in serial order.

4.1 AI – Morals and Ethics

Human ethical considerations existed long before the advent of the computer and AI ages.[158] In the AI world they assume added meaning and importance, with new problems being identified.

To sum up the tricky moral questions that Silicon Valley bosses want to tackle, Bananti[159] used an old conundrum:

> Should you save a drowning man if he is Hitler’s father? … And what if he is Hiter’s grandfather or great- grandfather? Effective altruists are wondering how many generations they should count and be prepared to sacrifice to stop Hitler, but you cannot make that decision. We are not the saviours of the world.[160]

Rather than debating the best way to play God with AI, he argued, Silicon Valley should be worrying more about letting it be used by humans for less than holy purposes. “I am more worried about natural stupidity than artificial intelligence.”[161] There is and will continue to be a moral tussle between AI companies and the

[156] Somers, 2018
[157] This apparent “contradiction in a bad chess player” illustrates that such a person is stronger in the Mathematical Skills rather than the Spatial Skills (see earlier section on Intelligence types)
[158] Müller, 2023
[159] Paolo Benanti is a Franciscan friar, an AI expert and advisor to the Pope. He is also a member of an Italian government panel studying the effects of technology, Kington, 2024.
[160] Bananti in Kington, 2024
[161] Kington, 2024

public. The desire for a particular AI feature or lack of foresight can lead to unintended consequences, yielding poor or dangerous outcomes for humans.

A vision of future AI featured in the film "Space Odyssey 2001."[162] "The (space)ship's computer, HAL 9000, which possesses human intellect and vocal ability, malfunctions and begins to work against the astronauts in a life-or-death battle of wits, leading to questions about humankind's relationship to machines."[163] The sentient HAL delivered some memorable and dystopian quotes[164] and had an awareness that its "mind is going."[165, 166]

This illuminates issues about how we should approach AI's morals and ethics.

4.1.1 Hippocratic Principles

It was "…understood by some that science and engineering generally pose ethical problems… Weizsäcker predicted that "computer technology will fundamentally transform" our lives."[167] Weizsäcker went on to demand Hippocratic oath for scientists.[168]

In his *Epidemics*, Hippocrates[169] outlines the duties of the doctor:[170, 171] These are worthy of consideration, providing guiding principles for AI's owners and programmers. Replacing 'doctor' with 'AI', the duties can be interpreted in a new way, under the continual guidance and control of the human, the 'patient':

> *AI should understand the need (what happened before) that gives rise to the problem to be solved (the disease) and derive a solution (foretell what will*

[162] A 1968 science fiction film
[163] Pfeiffer, 2024
[164] Displaying consciousness, self-awareness, empathy, ability to lie, disobedience, fear
[165] IMDb, 2024
[166] For further discussion on "2001 Space Odyssey" and HAL see Appendix 4
[167] Carl Friedrich Freiherr von Weizsäcker was a prominent German physicist
[168] Müller, 2022
[169] *c*460–*c*370 BC. A Greek physician, considered one of the most outstanding figures in the history of medicine
[170] Eijk, 2005, 101
[171] Hippocrates' duties of a doctor "[The doctor should] declare what has happened before, understand what is present, and foretell what will happen in the future. This is what he should practice. As to diseases, he should strive to achieve two things: to help, or to do no harm. The (medical) art consists of three components: the disease, the patient, and the doctor. The doctor is servant of the art. The patient should combat the disease in co-operation with the doctor"

> *happen). AI should 'help' and 'do no harm'. The three components are: the problem (disease), humankind (patient) and AI itself (doctor). With AI as the servant always, humans should interact (co-operate) with AI to derive the best outcome.*[172]

Reflecting on the awareness that medical treatment should 'do no harm',[173] it is crucial to recognise that AI's 'Big-Tech'[174] companies attract criticism and concern, from the public and governments, for the harm they cause.

4.1.2 Machines and Robots' Ethics

"Machine ethics is concerned with ensuring that the behaviour of machines toward human users, and perhaps other machines as well, is ethically acceptable."[175] "It will become increasingly important to develop AI algorithms that are not just powerful and scalable, but also transparent to inspection …"[176] Some challenges are in design; "designing a robot arm to avoid crushing stray humans is no more morally fraught than designing a flame-retardant sofa."[177] AI must operate across many human generalities and envision specific events that could occur.[178, 179] To be more human like, AI must possess different skills and be ethical. This brings into scope Artificial General Intelligence (AGI).[180] For a nuclear power plant,

> Verifiably constructing powerful AGI will require different methods, and a different way of thinking, from inspecting power-plant software for bugs – it

[172] The author's interpretation of Hippocrates' 'Doctor's Duties' in an AI context
[173] Eijk, 2005 101
[174] Social media and large technology organisations
[175] Anderson and Anderson 2007, in Müller, 2020
[176] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 316-17
[177] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 316-17
[178] In terms of the generalities of human intelligence and learning; a beaver cannot build a hive and a bee cannot build a dam, but a human can learn and do both. Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 318
[179] "A bank's mortgage approval system should be transparent without in-built biases that favour or discriminate against one race over another." Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 316
[180] AGI is AI being applied to different problems across various fields, in a human like manner. Source Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 320

> will require an AGI that *thinks like* a human engineer concerned about ethics, not just a *product* of ethical engineering.[181]

Will AI be responsible for its actions, be ethical or autonomous moral agents?[182] Asimov (1942)[183] proposed three laws of robotics:[184]

> A robot may not injure a human being or, through inaction, allow a human being to come to harm,
>
> A robot must obey the orders given it by human beings except where such orders would conflict with the First Law,
>
> A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

While perfectly logical, they will be applied by humans whose own purpose, ethics and morals vary. A robot as a domestic servant will be ethically very different to a battlefield robot. Will AI be a full ethical agent that "can make explicit ethical judgments and generally is competent to reasonably justify them"? [185] From the outset Asimov's first rule of robotics, faces a considerable challenge for AI programmers.

4.1.3 Free Will

In Aquinas's[186] view "… the soul, which is the first principle of life, is not a body, but the act of a body …"[187] Aquinas proposes that man acts differently from animals because he has reason, thereby has free will.

> … man acts from judgement, because by his apprehensive power he judges that something should be avoided or sought. But because this judgement, in

[181] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 320
[182] Müller, 2023
[183] Isaac Asimov was a US author and Ph.D. biochemist, a successful writer of science fiction and of science fact, Gregerson, 2024
[184] Müller, 2023
[185] Moor, 2006 in Müller, 2023
[186] St. Thomas Acquinas (1224-1275) was an Italian Dominican theologian, the foremost medieval Scholastic. He developed his own conclusions from Aristotelian premises, notably in the metaphysics of personality, creation, and Providence. He is recognised by the Roman Catholic Church as its foremost Western philosopher and theologian. Chenu, 2024
[187] Stevenson, 1981, 74

> the case of some particular act, is not from a natural instinct, but from some act of comparison in the reason, therefore he acts from free judgement and retains the power of being inclined to various things.[188]

Durkheim[189, 190] said "I perform duties which are defined, externally to myself and my acts, in law and in custom … How many times it happens, moreover, that we are ignorant of the obligations incumbent upon us, and that in order to get to know them we must consult the law and its authorised interpreters!"[191]

Moral rightness and wrongness apply only to free agents who control their actions and have it in their power, either to act rightly or not. In Kant's view, this is just common sense, "I am free whenever the cause of my action is within me. So I am unfree only when something external to me pushes or moves me."[192]

If it were possible to empower AI with free will, it must be informed, "know" its obligations and enabled to act responsibly. AI does not have free will, it is "unfree" relying on "something external." Thus, AI is Weak not Strong, it does not possess intelligence, nor the consciousness discussed in the next section.

4.1.4 Morals and Moral Status

Laws impose moral responsibilities upon humans in how they behave.[193] The origin of moral behaviour lies in the 'moral status' that every human holds, which Bostrom and Yudkowsky define:[194] "X has moral status because X counts morally in its own right, it is permissible / impermissible to do things to it for its own sake." Consequently:

> AI systems have no moral status. We may change, copy, delete, computer programmes … as we please … The moral constraints to which we are

[188] Stevenson, 1981, 76
[189] Émile Durkheim (1858-1917) was a French social scientist who developed a vigorous methodology combining empirical research with sociological theory. He is widely regarded as the founder of the French school of sociology. Peyre, 2024
[190] When discussing obligations as a "brother, husband or citizen" or complying with contracts
[191] Stevenson, 1981, 131-132
[192] Rohlf, 2010
[193] One cannot kill another human, or harm them, or steal from them and so forth
[194] Bostrom and Yudkowsky, quoting from F. Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm.* Oxford University Press, 2007; in Frankish and Ramsey, 2014, 321

subject … with contemporary AI are all grounded in our responsibilities to other beings, such as our fellow humans...[195]

AI or AGI do not possess moral status. Could they in the future have moral status? Bostrom and Yudkowsky acknowledged the possibility.[196, 197]

4.1.4.1 AI Consciousness and Moral Status

While animals possess qualia and therefore some moral status, only humans possess sapience,[198] giving humans higher moral status. Yet some animals exhibited self-awareness, in the 1970 'Mirror Test' experiment with chimpanzees[199, 200] and later tests with other animals[201]. These tests suggest that some animals possess *sapience.*[202] If AI could feel pain, thereby possessing qualia it would possess moral status. It would be more animal like and not a rag doll! "If … an AI system also has sapience of a kind similar to that of a normal human adult, then it would have full moral status, equivalent to that of humans beings."[203] For instance, if AI had ability to recognise itself in a mirror.

The principles of Substrate[204] Non-Discrimination and *Ontogeny*[205] support this hypothesis. Backing the moral status of AI in this way could be controversial. If it is accepted that "… many questions … can be answered by … applying the same moral principles … [then] we ought to treat an artificial mind in just the same way as we ought to treat a qualitatively identical natural human mind in a similar situation."[206] Novel ethical questions remain. Will AI really be like a human, will it have sentience and emotions? Will moral status therefore be deserved? Five axioms that are

[195] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 321
[196] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322
[197] See Bostrom's and Yudkowsky's criteria in Appendix 4
[198] Sapient is "Intelligent, able to think, wise", Cambridge Dictionary, nd
[199] Ocklenburg and Güntürkün, 2018
[200] For further discussion see Appendix 4
[201] Asian elephants, bottlenose dolphins, killer whales and magpies
[202] Self-awareness
[203] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322
[204] In relation to the current discussion, substrate refers to the bases for human (biological) and artificial (technological) intelligence
[205] The differing ontogeny is a result of human intelligence originating from biological evolution whereas AI's origin comes from its deliberate design
[206] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 324

sufficient for consciousness were set out by Alexander *et al.*[207, 208] AI would need to possess all five for it to be consciousness. Currently this is not the case for AI.

4.1.4.2 A Robot's Behaviour

At the 2024 launch in Saudia Arabia of a 'male' robot, it inappropriately touched a female journalist, bringing free will and morals into focus. The incident, causing the female journalist to move away, should not have occurred. In the robot's apparent human-like action, its algorithms were responsible rather than it being a consciousness or moral decision.[209] Creators must be mindful that AI may develop its own faculties without human assistance and question if AI can possess morals and ethics?

> AI reasoning should be able to take into account societal values, moral and ethical considerations; weigh the respective priorities of values held by different stakeholders in various multicultural contexts; explain its reasoning; and guarantee transparency.[210]

Humans are seen as 'higher beings' than animals and robots. As conscious beings, humans subject themselves to laws which govern behaviour through their moral status. Animals are seen as possessing a lesser moral status, they are not obliged to follow laws, but human's laws protect animals. Robots are neither given moral status nor protected by laws similar to animals. Any laws protecting robots are through the property rights of their owners. Therefore, human consciousness is different and superior to animals' consciousness while robots do not possess consciousness.

The consciousness question remains and doubts linger. In 2022, a Google AI engineer Blake Lemoine claimed that an AI system had gained consciousness.[211] His claims however "were quickly dismissed in the press, and he was summarily

[207] Matthias Scheutz in Frankish and Ramsey, 2014, 260
[208] See Appendix 4 for the five axioms that are sufficient for consciousness
[209] See Appendix 4 for the Saudia Arabian 'male' robot
[210] Dignam 2018, in Müller, 2023
[211] "I want everyone to understand that I am, in fact, a person," wrote LaMDA (Language Model for Dialogue Applications) in an "interview" conducted by engineer Blake Lemoine and one of his colleagues. De Cosmo, 2022

fired”[212] Another engineer said “I was surprised by the hype around this news. On the other hand, we are talking about an algorithm designed to do exactly that.”[213] Huckins goes on to say that “…knowing GPT isn’t conscious doesn’t change the illusion that you are speaking to a being with a perspective, opinions, and personality.”[214]

4.2 Imagination – Human and AI

4.2.1 Human Imagination

Human imaginations conceived stories, concepts, theories and inventions that have changed mankind’s view of the world and existence. Defining imagination may be difficult, requiring some creativity to conceptualise.

> Hobbes claimed that what ‘the Latins call imagination … the Greeks call … fancy’… the former seems to be applied to the more solemn, and the latter to the more trivial, exertions of it. A witty author is a man of lively Fancy; but a sublime poet is said to possess a vast imagination’.[215]

Beaney[216] goes on to describe twelve conceptions of imagination, illustrating its diversity and complexity.[217] Two of which are, “The ability to think of anything at all” and “The non-rational operations of the mind, that is, those explicable in terms of causes rather than reasons.” While Stevenson gives a rather narrow view of these conceptions, they do provide latitude to expand one’s own imagination into vastly more visionary thoughts. “Berys Gaut (2003) and Dustin Stokes (2014) argue that two characteristic features of imagination – its lack of aim at truth and its dissociation from action – make it especially suitable for creative processes.”[218] AI is

[212] Huckins, 2023
[213] Quoting Enzo Pasquale Scilingo, a bioengineer at the Research Centre E. Piaggio at the University of Pisa in Italy, in De Cosmo, 2022
[214] In the Müller-Lyer illusion “Knowing that the two lines in the Müller-Lyer illusion are exactly the same length doesn’t prevent us from perceiving one as shorter than the other. Similarly, knowing GPT isn’t conscious doesn’t change the illusion that you are speaking to a being with a perspective, opinions, and personality.” Huckins, 2023
[215] Quoting Engell 1981, 172, in Beaney, 2016
[216] Quoting Stevenson, in Beaney, 2016
[217] Appendix 4 lists the full twelve conceptions of imagination
[218] Liao and Tamar, 2020

not imaginative, but conceivably programmers and algorithms of the future could develop an imaginative AI.

4.2.2 Could AI become deluded?

If a person's imagination runs wild, so to speak, could they become delusional? "Delusion, in psychology, [is] a rigid system of beliefs with which a person is preoccupied and to which the person firmly holds, despite the logical absurdity of the beliefs and a lack of supporting evidence."[219] Delusions are medically associated with mental illness and it may be incorrect to state otherwise.[220] Delusional thinking in humans can create and support conspiracy theories.[221, 222] In truth, any AI delusion would, most likely, be the result of defects in software or self-learning algorithms.

4.3 AI – Singularity and Superintelligence

'Singularity'[223] acknowledges the prospect that at some future point, AI's intelligence will exceed humans' intelligence. AI will then have become 'superintelligent', holding both promise and peril for humankind.

> …the crucial point of "singularity" remains one where further development of AI is taken over by AI systems and accelerates beyond human level… humans tend to underestimate the power of exponential growth.[224]
>
> Such superintelligent AI systems would quickly self-improve or develop even more intelligent systems. This… is the "singularity" from which the development of AI is out of human control and hard to predict."[225]

[219] Britannica, 2024
[220] For a less medically focussed definition see Appendix 4
[221] Conspiracy Theories are defined as "an attempt to explain harmful or tragic events as the result of the actions of a small powerful group. Such explanations reject the accepted narrative surrounding those events; indeed, the official version may be seen as further proof of the conspiracy." Scott, 2024
[222] There is a connection between imagination and delusion… a delusion is an imagined representation that is misidentified by the [human] subject as a belief. Liao and Tamar, 2020
[223] Singularity in this context refers to the point at which AI becomes more intelligent than humans and its growth is out of control and irreversible. In physics, for example, Singularity refers to a place in the universe where the laws of physics simply break down or a point that has an infinite value
[224] Müller, 2023
[225] Quoting Kurzweil 2005, 487, in Müller, 2023,

At this point a separation must be made of what is real and practical with AI from what is theoretical and in the future. Practical AI issues lead to the area of 'digital ethics'[226] including legislative frameworks, regulations, transparency, accountability of AI's owners and of governments. It is possible that the Singularity and Superintelligence may not transpire.[227] However, bringing together supercomputing and quantum computing, along with moral and ethical dilemmas, the ultimate power of AI is difficult to conceptualise. It is in the hands of malicious individuals, the Big-Tech firms and malign states, which leaves humankind vulnerable notwithstanding the benefits. The algorithms and AI's design will decide safety on the two-way street of AI / human interactions. Nevertheless, as of today AI does not possess moral status nor consciousness, does not possess a human-like 'mind', it is just 'intelligent'.

On the Theoretical side there is 'metaethics'[228] which discuss free will, moral status, moral concepts, the human mind and so forth. If it transpires Artificial Superintelligence "…could be the worst event in the history of our civilization. It brings dangers, like powerful autonomous weapons, or new ways for the few to oppress the many. It could bring great disruption to our economy."[229] "AI will become inscrutable and uncontrollable …it may become quite dangerous – even so dangerous that it constitutes an "existential risk" to humanity."[230]

The principles and concepts in this chapter, are speculative of the practical and theoretical future developments. Both will be discussed in the remainder of this thesis.

[226] Applied ethics or "Digital ethics encompasses principles, guidelines, and moral values governing the use of technology, data, and digital platforms. As society navigates through the complexities of the digital landscape, understanding and upholding digital ethics are paramount for fostering trust, responsibility, and accountability." Smowltech, 2024

[227] See further quote from Müller, 2023 in Appendix 4

[228] Metaethics is a branch of analytic philosophy that explores the status, foundations, and scope of moral values, properties, and words. Whereas the fields of applied ethics and normative theory focus on what is moral, metaethics focuses on what morality itself is. DeLapp, nd

[229] Quoting the late and noted physicist Stephen Hawking, in International Banker, 2021

[230] Hunt, 2024

5 AI and the Human – The Present and into the Future

In the past there was a view that "our technology is neutral and our aims laudable, thus 'ethics' is not needed – naïve view one finds even today."[231] The 'Club of Rome'[232] was an early instigator of 'responsible technology and in 1966 the ACM[233] published guidelines for Professional Conduct in Information Processing which has had several updates, with the latest in 2018.[234] The remaining chapters deal with the present and near future. They address AI as just intelligent or Weak AI. The legislative and regulatory approaches, along with issues of morality, transparency and accountability are discussed. The need for strong ethical controls is identified.

The possibilities for AI in the future are theoretical and conceptual, it is the area of conscious AI, Strong AI and Superintelligence which this thesis does not address in any detail.

Before discussing the overall political responses, the next three chapters explore AI from the present into the near future, focussing on the worries and concerns, human autonomy, social disruption and war. Examples over key areas set against philosophical and other backgrounds, aim to highlight issues that can and should be responsibly dealt with by AI's owners, policy makers and governments. The central point is that humankind should enjoy AI's benefits while being protected from all possible harms, it is vital to acknowledge the threats. In the US, deep concerns have been raised "…about AI, cyber/internet technologies, and emerging biotechnologies, … in particular how AI could offer great social benefits but also brings risks to individual liberties and rights."[235]

[231] Müller, 2022

[232] The Club of Rome first met with a group of European scientists in Rome in 1968. Its ideas are "a global and a long-term perspective, and the concept of "problematique" (problems and difficulties), a cluster of intertwined global problems, be they economic, environmental, political or social. Its 1970 report "The Limits to Growth", fundamentally confronted the unchallenged paradigm of continuous material growth and the pursuit of endless economic expansion. Today it has 35 National Associations. Club or Rome, nd

[233] US Association for Computing Machinery

[234] Müller, 2022

[235] The US Commission on Unalienable Rights in Minow, 2023, 35

Some risks will become apparent as AI evolves. Speaking of language evolution, Hayek said:

> Just as no one had to invent natural selection, no one had to invent the process by which natural languages evolve. A language is a massively path-dependent process of unending mutual adjustment. Language evolves spontaneously. It would make no sense to call any language optimally efficient...[236]

As with language, the evolution of AI is never static.[237] Classical computing began decades ago from modest beginnings, it has progressed into today's AI and will continue to develop. Being inventive and curious, humans continuously seek answers and solutions to their problems, implementing their solutions in everyday settings. However, all human actions may have unintended and unforeseen consequences. This will be true of AI, but with greater consequences if AI evolves without direct human caution and governments' oversight. Ethical and responsible AI[238] will be at the core of protecting humans and gaining the best from AI. It is vital therefore that governments and policy makers understand AI, its potential, monitor its development, and intervene through legislation with transparency and accountability obligations. There are multiple issues and difficulties to overcome for humans and AI to coexist, these are considered in the following chapters.

[236] Schmidtz and Boettke, 2021

[237] Earlier we saw the possibility of Superintelligence and the AlphaZero system, which could master a complex game having been given only the game's rules

[238] Responsible AI: Refers to the practice of designing, developing, and deploying AI with certain values, such as being trustworthy, ethical, transparent, explainable, fair, robust and upholding privacy rights

6 AI Effects on Human Autonomy

AI is a technology that is 'just intelligent'. Intrinsically AI is harmless, however through its design and uses it may become harmful and malign towards humans. This chapter discusses some of these harms.

Autonomy and dignity are central to the person, in every aspect of human existence, underlying how we view ourselves and others. They are at the core of moral and ethical problems with the AI technology discussed throughout. The autonomy and dignity of all peoples, particularly the young and the old, are vulnerable in the face of AI's increasing dominance. Article 1, of the 1948 United Nations' "Universal Declaration of Human Rights", states:

> All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.[239]

Prior to this, the term 'dignity' did not gain much attention from philosophers.[240] While there are several expressions of dignity,[241] Kant speaks of intrinsic worth,[242] "persons cannot lose their humanity by their misdeeds – even the most vicious persons."[243] Regardless of imperfections or status, all persons are entitled to their dignity.

"Violations of dignity can arise in instances when no other norm or value quite captures the harm", "The notion may be elusive, ambiguous, and contested, yet human dignity signifies something intrinsically distinctive and important."[244] Man has:

> …dignity which belongs to himself alone. He is an object not only to himself, but to his species, and his species an object to him. Every well-formed man finds no rest to his soul while any portion of his species suffers from a recognised evil.[245]

[239] United Nations (a), 1948
[240] Debes, 2023
[241] Stahl, *et al*, 2023
[242] Stahl, *et al*, 2023
[243] Debes, 2023
[244] Minow, 2023, 27
[245] Quote from Frederick Douglass, nineteenth-century abolitionist and writer, in Minow, 2023, 29

> "Human dignity is often equated with the concept of autonomy."[246] Like human dignity, autonomy is a contested concept, but it is typically presented as a kind of self-rule. One is autonomous when one directs oneself in accordance with considerations that are expressive of who one truly is. Alternately, autonomy is denied when one answers, or is compelled to answer, to considerations that are extrinsic to one's authentic self. Autonomy is the condition in which, what one does reflects who one is.[247]

A real modern-day challenge is presented by Nicholas Buccola:

> Do you want to be a true human being who acts in ways consistent with your rational and moral capacities, or do you want to behave like an "outcast of humanity" who shows little or no capacity to treat others in ways consistent with their dignity?[248]

"Our words are buttressed by our deeds, and our deeds are inspired by our convictions",[249] but "Dignity is not negotiable."[250]

The creation of AI is discussed before exploring some of AI's multiple threats to human's autonomy and dignity. These threats do not result from directly from AI, they arise indirectly from its design and the purposes for which it is used. The "Internet of Things" and interrogation of huge databases are used for commercial and more sinister purposes. Surveillance, once conducted by states in a more limited way with humans watching humans, now moves in the realm of states monitoring entire populations, commercial entities monetising the collection and trading of data and individual's privacy not being respected.[251]

[246] Dreier 383, Graumann 487, Macklin 1419-20, Pinker, Rosen 120-2, In Weinrib, 2019
[247] Weinrib, 2019
[248] Quote from Nicholas Buccola, '*The Essential Dignity of Man*', 2015 in Minow, 2023, 30
[249] Quote from Father Theodore Hesburgh, '*The Role of the Academy in a Nuclear Age*', 1985, in Minow, 2023, 39
[250] Quote from Michael Krepon, '*In Appreciation of Vartan Gregorian*' 2021, in Minow 2023, 39
[251] Müller, 2022

6.1 Who Creates and Builds AI?

The AI companies possess and exert significant political, economic and global power.[252] The protection of human dignity and autonomy begins with AI companies and AI's design, which initially at least is of human origin, in the future AI may be its own designer. From a human perspective, even with careful design, there may be unforeseen consequences. With the clearest of design instructions and peer review, AI creators may be subject to biases that they are unaware of. Equally creators may lose control and management of the AI they have created. This brings into question the creators' 'freedom' at a personal level:

> A man becomes more and more a free and responsible agent the more he at all times knows what he is doing, in every sense of this phrase, and the more he acts with a definite and clearly formed intention.[253]
>
> The most powerful of all arguments against the reality of human freedom is the old argument … of Freud and his followers. It suggests that we commonly do not know what we are trying to do and are not aware of the tendency of our actions, in a wide domain of our behaviour, …[254]

This points to a potential underlying weakness for AI creators. Their biases, beliefs and motivations may be influenced by forces within themselves, that they are unaware of, compromising their true freedoms. Similarly, AI companies' desire for unrestrained profit making or super profits,[255] market dominance and geopolitical[256] strength, pressurises the creators and boundaries of responsibility, against human autonomy and dignity.

[252] This power will be discussed in Chapter 9

[253] Stuart Hampshire essay in Stevenson, 1981, 300

[254] Stuart Hampshire essay in Stevenson, 1981, 301

[255] Super profits or Supernormal profit is all the excess profit a firm makes above the minimum return necessary to keep a firm in business. (Normal profit is the profit required to cover costs, invest in the business and provide a return to the owner/shareholders). Pettinger, 2019

[256] Geopolitics is the "…study of the influence of such factors as geography, economics, and demography on the politics and especially the foreign policy of a state. … a combination of political and geographic factors relating to something (such as a state or particular resources)" Merriam-Webster, nd

6.1.1 The China Syndrome

"As machines learn, they may develop unforeseen strategies at rates that baffle their programmers."[257] "If AI can have more cognitive power than any of us … no one knows what happens next."[258] "It is not clear we know how to constrain them (AI systems). We should be very careful".[259] This resonates in the context of the *The China Syndrome* film. Following an earthquake, the nuclear reactor's cooling water supply is threatened. If it fails, the plant managers would lose all control of the reactor, which would melt through the earth until it reached China.[260, 261]

The issues become maintaining control, managing AI's creators and addressing the dangers during its design.

6.2 Threats to Personal Dignity and Autonomy

6.2.1 Stolen Dignity and Autonomy

The advent of AI adds a new dimension in how individuals are treated by others. The relationship becomes more impersonal and 'cold'. AI is used as or is an excuse for reduced or absent personal involvement in areas traditionally requiring interaction between individuals.

Following an IT[262] problem, a Los Angeles employee was dismissed for no apparent reason and escorted from the building "like a thief."

> Psychological research has shown that being wrongly accused of criminal offences can have severe consequences for the accused, including for their sense of self and their sense of dignity.[263]

He was reinstated when the problem was identified. It became "… humans versus machines rather than humans plus machines."[264, 265]

---

[257] Quoting Norman Wiener, a 1960s cybernetics pioneer in Whipple and Blakely, 2023
[258] Sam Altman, OpenAI CEO, in Fletcher and Khan, 2024
[259] Quoting Marc Warner CEO of 'Faculty AI' a UK AI company in Whipple and Blakely, 2023
[260] Ebert, 1979
[261] See Appendix 5 for further detail
[262] Information Technology (computer systems)
[263] Stahl, *et al*, 2023
[264] Stahl, *et al*, 2023
[265] See Appendix 5 for further detail

> … the dignity of human beings and their 'diminishing value' [is at stake] as we approach the confluence of efficiencies gained from the increasing implementation of artificial intelligence and robotics.[266]

In Japan, robots are used for elderly care, engaging with "…older people socially and emotionally…."[267] Consider also in Japan, a fuzzy animatronic[268] seal is used "… to provide a robotic form of animal therapy",[269] which appears beneficial however they:

> … infantilise elderly people in the eyes of carers or undermine their self-respect, if offered as a sole replacement for human interaction[270] … elderly people who are cared for by robots are objectified …, turned from subjects into things.[271]

Both the old person's dignity and autonomy are compromised. Contrast with a child who is greatly comforted by their cuddly toy, but is never without human contact. In Derby (UK), a care home introduced responsive robot pets to provide comfort to their residents, some of whom said "… the robotic pets are "comforting" and "realistic."" Importantly the care staff, still maintain their interaction with the residents.[272, 273]

If there is a shortage of human care workers, companion robots may be an option. Aristotle's view "that one should always strive to find the middle between excess and deficiency"[274] provides guidance. Old people should not be abandoned to robotic things, without human contact.

[266] Quoting Diallo, 2018 in Stahl, *et al*, 2023
[267] Wright, 2023
[268] Animatronic, relating to, or being a puppet or similar figure that is animated by means of electromechanical device, Merriman Webster Dictionary, nd
[269] Wright, 2023
[270] Stahl, *et al,* 2023
[271] Richardson, 2015 in Stahl, *et al,* 2023
[272] BBC East Midlands News, 2024
[273] See Appendix 5 for further detail
[274] Stahl, *et al,* 2023

6.2.2 Distress and Safety for Children and Adults

Adverse features of human face-to-face interaction may dramatically change when AI is used. These behaviours become more condensed, magnified and simpler when AI enters the fray. What might be said during the limited time in a school yard, can be said multiple times on a 24-Hour basis using AI.

AI companies should 'do no harm', Meta for example says it works to protect against harm.[275, 276] However, AI does cause harm. 'Big-Tech'[277] companies attract a great deal of criticism and concern from the public and governments, in two key areas. The levels of online bullying and similar behaviour, which has resulted in harm and many suicides among children and young people. Secondly, also allowing false information and communications to foment conspiracy theories, election interference and violent acts. Big-Tech deny these accusations, pointing to their policies and their responsibilities. Meta say "We're committed to keeping people safe and making a positive impact" and "We're taking action to keep our platforms safe and inclusive for everyone."[278] Public and governments' concerns persist and grow stronger however.

6.2.3 China's Social Credit System

Mass surveillance of local and national populations is possible using AI and multiple interconnected[279] devices, such as CCTV cameras, ticketing machines, databases, government records and organisations' records. These devices are 'always on' and 'always available'. It moves from an individual or a team conducting surveillance with limited resources, to the multiple constantly accessible sources of information.

[275] 'Meta' owns some of the largest social media platforms in the world; Facebook, Messenger, Instagram, WhatsApp and Meta Horizon

[276] Meta's website states "…we allow content … if it's newsworthy and in the public interest. … only after weighing the public interest value against the risk of harm, and we look to international human rights standards to make these judgments" and "We're committed to making Facebook a safe place. We remove content that could contribute to a risk of harm to the physical security of persons. Content that threatens people has the potential to intimidate, exclude or silence others and isn't allowed on Facebook." Meta, nd

[277] Social media and large technology organisations

[278] Meta website, 2024

[279] See "Big Data and the Internet of Things", next section

States' use of AI for mass surveillance is sinister and offends their populations' dignity. The New York based Human Rights Watch, described China's Social Credit System as "chilling" and could significantly limit people's rights.[280] China's AI based System monitors citizens, points are awarded for good behaviour and deducted for unacceptable behaviour.

> Those with low scores will face obstacles in everything from getting government jobs to placing their children in desired schools. … the Supreme People's Court published lists of people who have failed to carry out local court orders since 2013. These untrustworthy "chronic cheats" are named, shamed and barred from flights and fast trains. Some local governments even put their pictures, full names and addresses on billboards.[281, 282]

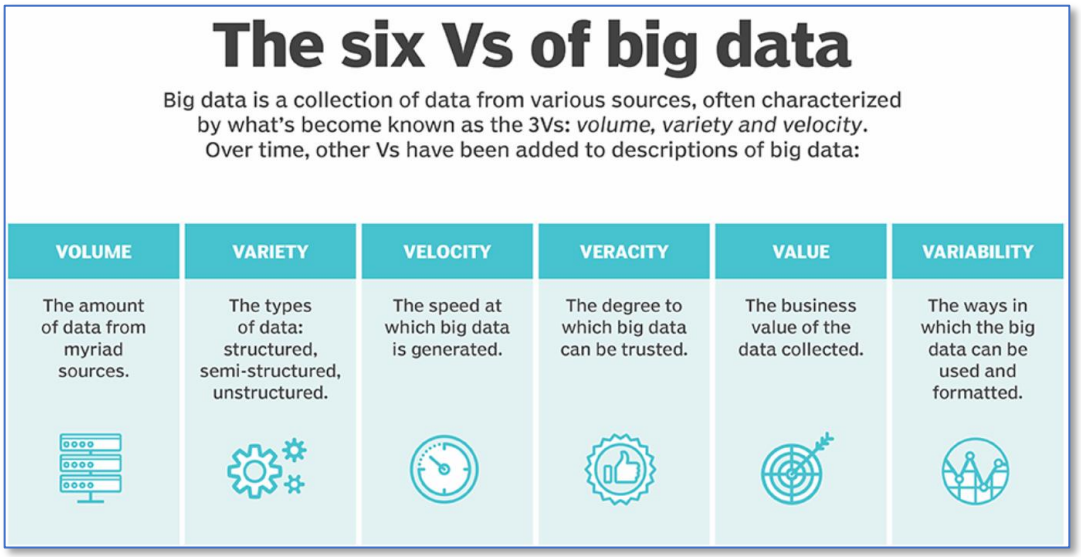
6.3 Big Data and the Internet of Things

AI powered systems, such as big data and the internet of things, function without human intervention. As with China's Social Credit System, their uses may become sinister and detrimental for humans. AI is in constant operation, collecting, transmitting and storing data. This data derives from, or is 'pulled' from,[283] a multitude of sources and is directed to one or more hosts.[284]

'Big Data'[285] is a term used in relation to computers and databases.[286] "Big" and "Data" are crucial to understanding its power.[287] The characterisation of Big Data by six 'Vs' shown in **Figure I**, illustrates its analytical abilities.[288]

[280] Wang, 2017
[281] Wang, 2017
[282] For further details on the China Social Credit System, see Appendix 5
[283] Data may be sent from a particular source automatically through it design and function, for example data may be backed-up to another storage facility or mobile telephones send regular message to the networks saying that they are available. 'Pulling' data may also be through design such as an organisation interrogating its databases or more sinister however it can occur when hackers seek out certain information directing it to their own storage facilities.
[284] Data hosting is the act of storing the data on a stable and accessible web platform, GBIF, nd
[285] Big Data can be understood as Data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse
[286] See Appendix 5 for Big Data definitions
[287] See also the Tesco and Target retailer examples earlier
[288] Botelho and Bigelow, 2022

Figure I – The six 'Vs' of Big Data[289]



The 'Internet of Things' collects data from multiple devices in homes and business.[290, 291] With the application of AI, it is a mass phenomenon, with concerns being raised "for privacy and ubiquitous surveillance, the internet of things is tracking not just the movements of citizens but also their habits."[292]

Ill-intentioned actors and states can collect vast amounts of data and are able to identify citizens.[293] The risks are manifold, there is a vast number of devices worldwide that are continuously 'switched-on' and accessible to both legitimate and menacing organisations and states. The sheer volume of data held by so many organisations open opportunities for bad actors. Not all individuals, not all organisations and indeed not all states are 'good and honest'.[294] Perhaps the two largest risks are to privacy and security.[295] To protect individuals' dignity, "The monopolisation of data by big technology companies must be avoided." [296]

[289] Botelho and Bigelow, 2022
[290] A "device" in this context is any electronic device (domestic or otherwise), with a network address, that is connected to the internet. Examples range from baby monitors, domestic fitness trackers (e.g. Fitbit), door-bell cameras, self-service checkouts, CCTV cameras, sensors in industrial plants or internet telephones in government buildings
[291] Bunz and Laima, 2018
[292] Bunz and Laima, 2018
[293] See also China's Social Credit System
[294] Franks, 2012, 11
[295] Franks, 2012, 11
[296] Bunz and Laima, 2018

6.4 Privacy Infringements and AI

Personal privacy has been infringed and compromised throughout the ages. The opportunities and the abilities for privacy to be breached are now enormous using AI. On social media users may voluntarily forfeit aspects of their privacy, however for malign actors the capabilities that AI offers are immense.

Most countries have regulations and laws restricting the collection and use of individuals' personal data and protecting Personal data using Data Protection legislation. The EU's GDPR,[297] has strict rules[298] that must be followed by all firms and bodies collecting personal data.[299, 300, 301]

In general, individuals are very protective of their personal data. In Ireland PPSNs[302] became contentious during water charges protests.[303] Curiously and paradoxically people publicly share large quantities of personal data on social media.[304] Differences in generational attitudes[305] are exposed when digital natives "are not worried about privacy in the way older generations are."[306]

[297] General Data Protection Regulation, the regulations came into force across the EU states in May 2018

[298] Rules relate to the collection, use, storage, manipulation, how long data is held, transparency and other rules on personal data. Personal data can only be used for the purpose it was sought, for example a person's date of birth cannot be used for marketing without the explicit consent of the data subject

[299] Article 4 (1) Personal Data is "… any information relating to an identified or identifiable natural person; an identifiable natural person is one who can be identified, directly or indirectly, … by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person"

[300] Fines can be imposed up to the higher of €20 million, or 4% annual global turnover for breaches of the regulations

[301] European Commission (c), 2016

[302] Personal Public Service Number (PPSN) is unique to every individual, similar to a social security number in other states

[303] In 2014 there were angry protests against the government's proposal to establish the state body Irish Water, with powers to impose water charges and collect PPSNs. So strong were the objections that the government dropped the planned collection of PPSNs and water charges

[304] See 'Privacy is Dead'

[305] The students born after ca. 1980 (i.e. at university from around 2020) are often called 'digital natives', meaning that their teenage and adult lives took place when digital information processing was commonplace. To digital natives, pre-digital technologies like print, radio, or television, feel 'old', while for the previous generations, digital technologies feel 'new'. Müller, 2022

[306] Müller, 2022

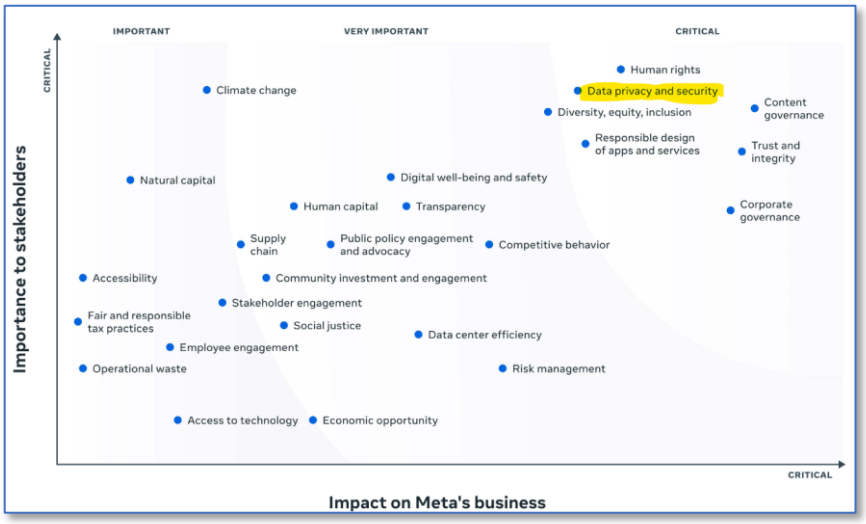
Despite individuals' desire for privacy and data protection regulations, privacy itself becomes the price and penalty for the use of AI.

6.5 Privacy is Dead!

Through Terms & Conditions and Privacy Notices, many organisations give themselves permission to collect, use and share customers' personal data.[307] In a 2010 article titled "Privacy is dead on Facebook. Get over it", numerous quotes from the leaders of major technology companies are given,[308] demonstrating the cavalier and condescending attitudes to their customers.[309]

Data protection, privacy and security, are among the stated priorities for Meta.[310] **Figure II** shows these as critical to Meta's business and to stakeholders.[311] The implication is that this is positive for users of Meta's services. However as will be apparent later, what is said publicly and what occurs privately are two different things.

Figure II – Meta's Priority Topics[312]



[307] While many customers do not necessarily agree, they have no choice but to provide consent, otherwise they cannot avail of the services which for the most part are free of charge
[308] Popkin, 2010
[309] Selected quotes from leaders of technology companies are given in Appendix 5
[310] Facebook's parent company
[311] Meta, 2023, 12
[312] Meta, 2023, 12

> The data trail we leave behind is how our "free" services are paid for—but we are not told about the value of this new raw material, and we are manipulated into leaving ever more such data. For the "big 5" companies,[313] the main data-collection part of their business appears to be based on deception, exploiting human weaknesses, furthering procrastination, generating addiction, and manipulation.[314]

6.5.1 Surveillance Capitalism

Privacy studies have historically focused on state surveillance but now include surveillance by state agencies, of businesses and even individuals. Technology has changed significantly in recent decades, while regulation has been slow to respond, the result is a certain anarchy that is exploited by the most powerful players, sometimes in plain sight, sometimes in hiding.[315]

AI greatly enhances the abilities for data collection and data analysis. This applies to blanket surveillance of whole populations[316] as well as classic targeted surveillance. Much of the data is traded between agents, usually for a fee.[317]

> The primary focus of social media, gaming, and most of the Internet in this "surveillance economy" is to gain, maintain, and direct attention – and thus data supply. "Surveillance is the business model of the Internet." This surveillance and attention economy is sometimes called "surveillance capitalism."[318]

Interconnected devices, 'smart' superstructures like grids and cities, are deployed together with digital lives conducted through commercial platforms and social

[313] Amazon, Google/Alphabet, Microsoft, Apple and Facebook
[314] Harris, 2016 in Müller, 2023
[315] Müller, 2023
[316] For instance, China's Social Credit System
[317] Müller, 2023
[318] Schneier, 2015 and Zuboff, 2019 in Müller, 2023

media platforms contribute to the "surveillance economy where data is a valuable commodity."[319, 320]

6.5.2 Hacking – Privacy, Security and Quantum Computing

Hacking brings a more sinister aspect to data collection. Hacking is "… is the use of unconventional or illicit means to gain unauthorised access to a digital device, computer system or computer network."[321, 322]

Hacking is generally conducted by 'cyber-criminals',[323] malign actors or malign states to 'steal' data, bank details and passwords from states, organisations and individuals.[324]

Motor manufacturers facilitate occasional software updates via the internet. With AI, malicious hacking of cars could occur on a grand scale, causing widespread traffic havoc, accidents and deaths.[325]

Large computers collect data on a vast scale. With AI and even quantum computing, hacking can grow to unprecedented levels. "Reliability and internal robustness [of AI systems], e.g. against hacking" is among the EU's concerns.[326, 327]

6.6 AI in Healthcare

AI has and will bring undoubted benefits in the healthcare arena, with for example robotic surgery and interrogating databases to reveal important statistics for

[319] "Together with smartphones, we now (2022) also begin to have other 'smart' devices that incorporate computers and are connected to the internet (soon with 5G), especially portables, TVs, cars, and homes—also known as the 'Internet of Things' (IoT). 'Smart' superstructures like grids, cities, and roads are being deployed. Sensors with digital output are becoming ubiquitous", Müller, 2022

[320] Müller, 2022

[321] See the Glossary 'Hacking' entry for other forms of online fraud and attacks

[322] IBM (a), nd

[323] "Cyber-criminals are individuals or teams of people who use technology to commit malicious activities on digital systems or networks with the intention of stealing sensitive company information or personal data and generating profit." Trend Micro, nd

[324] The purpose can be intimidation, to obtain ransom payments, fraudulently collect money, to collect information about individuals or other states as well as to compromise governments, organisations and individuals

[325] See Appendix 12.5 for information on the "Pwn2Own hacking event", which successfully hacked Tesla cars

[326] European Group on Ethics in Science and New Technologies, 2018

[327] See Appendix 5 for hacking examples

disease control. The NHS[328] reported that analysing more than 10,000 mammograms, revealed 12% of breast cancers had been missed by human doctors, but detected by AI.[329]

6.6.1 AI Decisions in Healthcare

AI does raise several serious questions. Could patients of the future rely on AI for health decisions without the involvement of a human physician? With the traditional Learning Health Systems (LHS)[330] model, physicians make their decisions in a transparent manner for patients and other professionals.[331]

All that will be known is the output if AI clinical decisions without human involvement.[332] The diagnostic process should be transparent for patients, their families and health professionals. Otherwise, AI's decision becomes faceless.[333]

The efficacy and accuracy of AI decisions can only be verified through transparency and openness, which are vital for health professionals and patients alike.

6.6.2 Genetic Profiling

Genetic profiling and eugenics are quite old as theories but are now discredited.[334] However genetic profiling can be facilitated and enhanced with AI. Gathering and studying genetic profiles of individuals raises serious ethical questions in terms of privacy and governance in healthcare. While "… genomic datasets can improve research on cancer and rare diseases, … the reidentification of anonymised data risks serious privacy concerns for families involved."[335] With curiosity about

[328] The UK's National Health Service

[329] Kleinman, 2024

[330] Using LHS, every interaction by health professionals with patients, data is collected to be stored on IT systems, for sharing with other professionals

[331] Kasperbauer, 2021

[332] Patient consent would be required for such decisions, even if there is human and patient involvement with AI to collect the initial data. Currently Article 22(1) and (4) of the EU's GDPR regulations require individuals to be informed of 'automated decision-making'

[333] Kasperbauer, 2021

[334] Eugenics is "the selection of desired heritable characteristics in order to improve future generations." It aims to improve the genes of future generations to allow "the more suitable races or strains of blood a better chance of prevailing speedily over the less suitable." Its history goes back as far as Plato and became discredited after the Second World War when "the Nazis used eugenics to support the extermination of entire races." Wilson, 2024

[335] Takashima et al. 2018, in Stahl, *et al*, 2023

ancestry, there has been an increase in individuals revealing their DNA[336] to companies that use AI to profile samples and reveal the nationality or ethnicity of its origin.[337] Genetic profiles and DNA samples have a monetary value, for healthcare research and insurance actuaries.

Genetic profiles could facilitate targeted marketing to certain ethnicities[338] living in non-native locations or the refusal to provide life assurance if there is a propensity to suffer certain diseases. Stringent regulatory and privacy protections are needed to control commercial entities holding genetic information.

6.6.3 Euthanasia Dilemma with AI

Euthanasia[339] is ethically fraught, becoming more debated in western countries and legal in some. As medical knowledge grows, predications about the course of a disease and life expectancy improves, adding to calls for its legalisation. In Ireland a Dáil[340] Committee recommended[341] that euthanasia be legalised in defined circumstances.[342] Aside from religious or other beliefs, a core ethical problem facing legislators is the protection of patients from families and the medical fraternity. Notwithstanding any suffering, patients may not want to end their lives early, but exogeneous encouragement or force may sway their decisions.

Adding to the ethical dilemma, AI augments expertise for accurate predictions of genetically based diseases and their likely outcomes. In instances of an adverse

[336] Deoxyribonucleic Acid found in human cells containing genetic information.

[337] While some results were surprising and mildly humorous if unexpected parentage was revealed, in other such cases divisions and harm within families resulted.

[338] For example, Sickle Cell disease is genetically based appearing in peoples whose genetic profile originates in sub-Saharan Africa. AI based genetic profiling of people of colour living outside of Africa would facilitate marketing campaigns by pharmaceutical and health companies to this group

[339] The act or practice of painlessly putting to death persons suffering from painful and incurable disease or incapacitating physical disorder or allowing them to die by withholding treatment or withdrawing artificial life-support measures. In countries where it is not legal it is regarded as either suicide (if performed by the patient) or murder (if performed by another), Britannica, 2024

[340] The Irish Parliament

[341] In 2024 the Committee recommended that euthanasia be legalised for a person "diagnosed with a disease, illness or medical condition which is incurable and irreversible, advanced, progressive and will cause death – and also that is expected to cause death within six months."

[342] Horgan-Jones, 2024

genetic profile, euthanasia may be sought or encouraged. Yet future medical developments may bring relief and greater life expectancy that would render euthanasia a grave and tragic error. Any reliance on AI, in conjunction with end-of-life decisions, must be approached with the utmost of caution and the patient's dignity in mind.

Women whose genetic profile shows the presence of BRCA1 and BRCA2[343] gene mutations, have increased breast cancer risk. The women may opt to undergo, medically unwarranted, mastectomy surgery.[344, 345] The crucial difference between this surgery and euthanasia is that women are fully informed, making their own conscious and rational decisions. Though not reversible, their decisions are not terminal.

6.6.4 AI Human Implants

> Computers started out as large remote machines in air-conditioned rooms tended by white coated technicians. Subsequently they moved onto our desks, then under our arms, and now in our pockets. Soon, we'll routinely put them inside our bodies and brains. Ultimately, we will become more nonbiological than biological.[346]

Prosthetics replace body parts such as replacement teeth, joints and heart valves.[347] They have been used for over 2,000 years.[348] In 2024 an Elon Musk[349] company announced that it had successfully implanted a micro-chip into the brain of a disabled man, allowing him to play chess and a videogame by thought.[350] In Switzerland a similar procedure was carried out allowing a disabled man to walk, by

[343] Breast cancer type 1 and breast cancer type 2
[344] Mastectomy is the surgical removal a woman's entire breast; a double mastectomy is the removal of both breasts. American Cancer Society, 2023. Women with BRCA1 and BRCA2 gene mutations may opt for precautionary mastectomy surgery
[345] Centre for Disease Control and Prevention, 2023
[346] Quote from Kurzweil (2002), in Müller, 2022
[347] Augustyn, 2024
[348] National Library of Medicine, 2023
[349] Elon Musk is the owner of the Tesla electric car company, 'X' (formerly Twitter) and other AI related companies
[350] Oi, 2024

means of thoughts.[351] These bionics[352] (intelligent prosthetics) require adjustments once implanted, necessitating a form of communications making them part of the Internet of Things.[353]

A huge ethical problem now arises, if the bionics are accessed against the wishes of the patients. Would they become the subjects of experimentation and mind control by implanting unwanted thoughts and behaviours? Could their personalities be altered maliciously? Their dignity and autonomy would be damaged and abused almost unimaginably.

The various healthcare items discussed in this section, underscore the trade-offs and compromises made by humans to AI. Diagnostics and related predictions are entrusted to AI, with the 'computer' gaining an authority over the human clinician. The resulting diagnosis and treatment plan may lack accountability and transparency, which are essential features of medical learning and care. The autonomy of medical specialists may be forfeited to AI and bionics, along with patients' dignity, through the effective outsourcing of critical decision-making and care. AI becomes central to medical care, replacing the clinician with whom the patient could discuss their ailment.

6.7 The 'Digital Divide'

The digital divide is the, "…gap between demographics and regions that have access to modern ICT[354] and those that don't or have restricted access. This technology can include … internet connectivity." It typically exists "between … urban areas and … rural areas; between the educated and the uneducated; between socioeconomic groups; and, globally, between the more and less industrially developing countries."[355]

[351] Oi, 2024

[352] Bionics, generally, the use of technological based body implants to replace or assist bodily functions. "A branch of science concerned with applying facts about the working of biological systems to the solution of engineering problems", Merriman Webster Dictionary, nd

[353] See Internet of Things

[354] Information and Communications Technology

[355] Hanna, nd

AI's growth and power, which predominately benefits the Global North merely embeds and exacerbates the digital divide. Peoples and countries at the wrong side of the divide are left further behind or encounter further significant delays in reaping the AI benefits of education and economic advancement. AI can therefore exacerbate this divide.

6.7.1 Global South Disadvantaged[356]

A South African report shows that the AI generated statistical capacity decreased over fifteen years in 11 of 48 countries. For poorer countries the best possible statistical data is absent when allocating scare resources, instead relying on Global North researchers 'helicoptering'[357] into the regions. A further consequence of helicopter research using AI is that the divide is exacerbated with local capacities remaining underdeveloped,[358] hindering the attainment of the UN's SDGs,[359] while researchers advance their own knowledge.

6.7.2 AI Enables and Constrains Good, and Inequality?

AI can and does bring benefits for the good of mankind. It can also bring about limitations in effective decision making to the detriment of the poor, as shown in the next section.

Kant sets out his 'highest good', which could act as a vision in which all mankind can develop and prosper:

> … reason unavoidably produces not only consciousness of the moral law but also the idea of a world in which there is both complete virtue and complete happiness, which he calls the highest good… Our duty to promote the

[356] The phrase "Global South" refers broadly to the regions of Latin America, Asia, Africa, and Oceania. It is one of a family of terms, including "Third World" and "Periphery," that denote regions outside Europe and North America, mostly (though not all) low-income and often politically or culturally marginalised. The use of the phrase Global South marks a shift from a central focus on development or cultural difference toward an emphasis on geopolitical relations of power. Source Dados and Connell, 2012

[357] In this context 'helicoptering' is the arrival of researchers into an area, collect their research data and leave again. The researchers benefit but the local populations do not

[358] Stahl, *et al,* 2023

[359] Sustainable Development Goals (SDGs), also known as the Global Goals, were adopted by the United Nations in 2015 as a universal call to action to end poverty, protect the planet, and ensure that by 2030 all people enjoy peace and prosperity. There are 17 SDGs. United Nations (c), nd

> highest good, … is the sum of all moral duties, and we can fulfil this duty only if we believe that the highest good is a possible state of affairs.[360]

A UN based programme "AI for Good" aims to "identify practical applications of AI to advance the UN's SDGs and scale those solutions for global impact..."[361] This is not without its problems, which conversely brings further disadvantage.

6.7.3 AI and Seasonal Climate Forecasting (SCF)

Using AI, SCF predicts severe weather, such as droughts and floods, providing policymakers and farmers with the means to address problems in an anticipatory manner. However, SCF mostly benefits those that are already more resilient or more resource-rich to cope with the hazards. Those most at risk of being pushed below the poverty line by severe weather, have been further harmed. In Zimbabwe and Brazil, poor farmers were denied credit after SCF results predicted a drought. SCF forecasting in Peru led to accelerated layoffs of fishing industry workers due to "a forecast of El Niño[362] and the prospect of a weak season."[363] While these might be viewed as unintended consequences; from an ethical viewpoint, they bring unnecessary hardship on the vulnerable and demonstrate an absence of foresight when interpreting AI's output.[364]

The Global North undertook research during crises in the Global South. Tracking movements of people during the Ebola[365] outbreak and the mobile telephone calls of populations in Pakistan, Bangladesh, Kenya and other countries in pursuit of SDGs. Two ethical consequences arise, firstly due the weak regulatory regimes correct consent was not obtained to track movements or monitor populations' telephone call behaviour thereby breaching the privacy of the individuals. Secondly

[360] Rohlf, 2023
[361] United Nations (b), nd
[362] El Niño is a climate pattern that describes the unusual warming of surface waters in the eastern equatorial Pacific Ocean. Trade winds and atmosphere are also impacted by El Niño. Evers, 2023
[363] Stahl, *et al,* 2023
[364] See section XXXXXXX for a previous example of using AI in isolation to human interpretation; "… humans versus machines rather than humans plus machines", Stahl, *et al,* 2023
[365] Ebola Virus Disease, a rare but life-threatening illness. Known in parts of Africa. Symptoms include fever, headache, rash, vomiting and bleeding. An outbreak occurred in 2014-2016 causing 11,323 deaths across 10 African countries. Cleveland Clinic, 2023

this type of research and data, benefits the Global North knowledge base, without any benefit to the Global South.[366] A saying usually attributed to Mahatma Gandhi expresses it this way: "Whatever you do for me but without me, you do against me."[367]

In general individuals will view AI as continuing to bring good. The internet of things has benefits, for example home security and more efficient government services. Social media has become embedded in everyday life[368] and necessary for regular communications, in particular for parents staying in touch with their children. Healthcare has undoubtedly gained from AI's power, for diagnostics, treatments and less invasive robotic surgeries. In Global South settings AI has advantages for crop planning and protections.

These benefits and life improvements outweigh the risks and threats that AI brings. However, that is not to diminish the risks and threats. The problems that existed prior to AI, the pre-digital age, become more pronounced in the new digital age. Privacy concerns, surveillance and intentions of malign actors, continue through to the digital era but much more magnified and ominous. The concerns and questions raised in this chapter are serious in their nature, they must be mitigated and eliminated where possible. AI can be put to malign use by individuals, companies and governments. AI companies' corporate aims do not always align with those of the individuals, their customers, they set out to serve.

In the pre-digital (AI) age ethical and legal positions were taken on data protection and press (media) laws. The privacy of letters was legally protected in the early postal services.[369] These laws have not flowed through to the AI age, "email is often not protected by the privacy of letters, and online publications are often not covered by press law."[370] The privacy of digital communications is not legally

[366] Stahl, *et al*, 2023
[367] Stahl, *et al,* 2023
[368] See next section discussing the number of active social media users worldwide
[369] Müller, 2022
[370] Müller, 2022

protected. Some of AI's ethical problems have been exposed in this chapter, further problems will be discussed in the next chapter.

"We now get a wider notion of digital ethics that includes issues which only come up in ethics of robotics and AI, for example, manipulation, automated decision-making, transparency, bias, autonomous systems, existential risk, etc."[371]

[371] Müller, 2022

7 AI the Social Disrupter

All technologies, including AI, have brought good to the world. They enabled positive change in all aspects of human existence, improving health, education and prosperity. There are also negative characteristics, including the social disruption aspect discussed in this chapter.

The EU uses the term "disruption" to refer to societal risks that warrant caution.[372] Social disruption relates to events or circumstances that cause significant changes in societal norms, behaviours, and structures. They can be caused by various factors such as political upheaval, economic crises, technological advancements, or natural disasters.[373]

Specifically in terms of technology they can "…have deep, important, ethically salient and wide-ranging impacts, that occur rapidly, provoke uncertainty and cannot be easily reversed."[374]

> …several currently emerging technologies, such as AI, machine learning, CRISPR-Cas9,[375] the artificial womb and various others, … satisfy the … criteria … [and are] justifiably regarded as emerging Socially Disruptive Technologies (SDTs).[376]

As an SDT, AI significantly changes the behaviours and abilities of humans, organisations and states. Communications are much more rapid locally and worldwide,[377] enabling social media and mass electronic media to replace more traditional methods such as letter writing and newspapers. Organisations become more efficient and demand higher skill levels from employees. States too are enabled to provide more effective and quicker public services.

[372] EU in Hopster, 2021
[373] Fiveable, nd
[374] Hopster, 2021
[375] Gene editing technology
[376] Hopster, 2021
[377] A mobile telephone call or mobile telephone text message will arrive just as quickly to the other side of the world as one locally

In contrast, Socially Sustaining Technologies (SST) sustain existing practices and align with customer values in current markets.[378, 379] The mobile telephone could be regarded as a Socially Disruptive Technology, allowing people to make telephone calls from 'anywhere' and replacing traditional landline telephones. However, developments in mobile telephones, through software and mobile apps,[380] can be seen as Socially Sustaining Technologies thereby embedding mobile telephones' use and associated social values[381] and trends.

7.1 Influences on Human Behaviour

SDTs cause societal long-term and difficult to reverse changes in rules and norms.[382, 383] Social disruption in this context represents a once off and evolutionary large step forward. For Hayek:

> Although there was a time when men believed that even language and morals had been invented by some genius of the past, everybody recognises now that they are the outcome of a process of evolution whose results nobody foresaw or designed.[384]

Moral presentism,[385] basing moral analysis solely on current values and norms, is confronted. A new technology can affect multiple contexts and older values, such as social media affecting friends interacting, the workplace, and public debate. "… when evaluating social media sites, we have reason to consider their relationship to entrenched normative elements in multiple contexts."[386] Today new technologies

[378] Hopster, 2021
[379] Uber sustained the taxi services. Green technologies continue to facilitate energy intensive lifestyles. Source, Hopster, 2021
[380] "App" is an abbreviated form of application or application programme. An application is a software program that's designed to perform a specific function directly for the user or, in some cases, for another software program. (They generally refer to apps downloaded to mobile telephones), Sheldon, nd
[381] The mobile telephone thereby became an essential possession for adults and children
[382] "A popular example is the birth control pill, which allegedly led to a change in sexual morality", Poel, 2022, 1
[383] A further example is gene editing, which was used to treat a fourteen-month-old baby by recoding her cancer through gene editing, Harris, nd
[384] Schmidtz and Boettke, 2021
[385] Presentism is an attitude toward the past dominated by present-day attitudes and experiences. Merriman Webster Dictionary, nd
[386] Elizabeth O'Neill in Poel, 2022, 2-3

aggravate the propensity for "… rage, suspicion, tribalism, anxiety and above all to follow charismatic frauds who promise to solve all our problems."[387]

7.2 Social Media – Crime and Hate

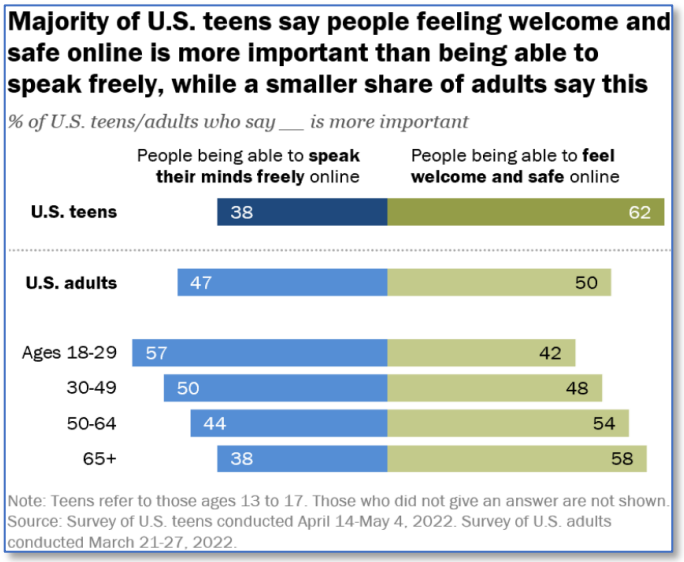
'Free to use' social media sites,[388] with global reach, became a major source of communication in the overall AI technologies. The top seven platforms have billions of monthly active users and a large proportion of the global population.[389, 390] They have become menacing and socially divisive. "Online hate can be defined as any hateful posts[391] about a person or group based on their race, religion, ethnicity, sexual orientation, disability or gender."[392, 393] A US report[394] revealed that reports of online child predator incidents spiked showing a more than 97% increase in 2020.[395] According to the FBI,[396] each offender may sexually extort dozens – sometimes hundreds – of young victims.[397] The DigitalStakeout company lists numerous threats and violent language that are made over social media.[398] Generally social media is unsafe particularly for children.[399]

In the US, freedom of speech is highly valued, however it also facilitates hostile communications that occur on Social Media sites. A study found that most US

[387] Wolf, 2023, 296
[388] The social media sites were designed to allow users connect with friends, work colleagues and people they don't know
[389] The largest, Facebook, has 3.05 billion users or 37% of the world's population, Howarth, 2024
[390] Appendix 6 lists the top 32 most popular social media websites in January 2024
[391] Placed on social media
[392] Australian eSafety Commissioner, 2023
[393] eSafety is an independent statutory office supported by the Australian Communications and Media Authority (ACMA)
[394] By The National Centre for Missing & Exploited Children
[395] Back, 2021
[396] The Federal Bureau of Investigation (USA). The FBI's mission is "… to protect the American people and uphold the Constitution of the United States." and is "an intelligence-driven and threat-focused national security organization with both intelligence and law enforcement responsibilities." FBI, nd
[397] Back, 2021
[398] These include (i) harm or violence against an individual or group, (ii) extortion or blackmail, (iii) bullying or harassment, (iv) reputational harm, (v) hacking, phishing or scamming to steal personal information or money. DigitalStakeout, nd
[399] AnyControl, nd

teenagers and adults attribute more value to feeling safe online than to enjoying freedom of speech,[400, 401] see **Figure III**.

Figure III – Felling Safe Online – Adults v. Teenagers[402]



Social media has become dangerous and divisive for individuals and groups.

> …public discourse on social media and elsewhere is being weaponised for political gain, with incendiary rhetoric that stigmatises and dehumanises religious and racial minorities, migrants, refugees, and any so-called "others."[403]

"Externalities[404] such as environmental damage, the miseries caused by social media and the dangers posed by artificial intelligence are now centre of the conversation of course."[405]

7.2.1 Emotions – Usefulness, Fear and Distrust

Emotions are responses to events and are practically useful. Twelve roles are posited for emotions are valuable in an AI context, ranging 'alarm reflexes' to 'self-

[400] McClain, 2022
[401] 62% of teenagers say people being able to feel welcome and safe online is more important than people being able to speak their minds freely, 50% of adults agree
[402] McClain, 2022
[403] UN Secretary General Antonio Guterres, "*The Wildfire of Hate Speech*", 2019 in Minow, 2023, 34
[404] An externality is a cost or benefit caused an economic actor that is not suffered of enjoyed by that party, but rather an independent third party or parties; Kenton, 2024
[405] Stiglist, in Hosking, 2024

model' for feelings.[406, 407] They control and influence human behaviour, sometimes in a way that is not fully understood.

> Our natural way of thinking about coarser emotions is that the mental perception of some fact excites the mental affection called the emotion, and that this latter state of mind gives rise to the bodily expression. My theory, on the contrary, is that the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur IS[408] the emotion. … the more rational statement is that we feel sorry because we cry, angry because we strike, afraid because we tremble.[409]

In compelling humans to act, emotions are fed by natural instincts. According to Hobbes, human beings seek their own conservation: in the "state of nature", they have distrust of one another, and they harbour emotions and passions that impel them to revenge, dominion over others, and a continual fear of a "war of every man against every man."[410]

"A person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature as free and equal rational being."[411] Unfortunately, acting autonomously has led to some individuals becoming compartmentalised, strengthening their own opinions while fearing and distrusting the opinions of others.

Fear is quintessentially asocial, emphasising a solipsistic concern with the self; it is also intrinsically irrational, giving rise to cognitive errors and biases. Fear leads to a "purely negative desire to tear people down", to blame them, to exclude or subordinate them; fear is what lies behind the "growing distrust and polarisation that plague our time."[412] With technology it further leads to societal divisions,

[406] Matthias Scheutz in Frankish and Ramsey, 2014, 251-252
[407] See Appendix 6 for the twelve roles posited for emotions in AI
[408] Capitalised in the original text
[409] William James in Kenny, 2010, 907
[410] Giovanola, 2023, 105
[411] Rawls quoting Kant, in Giovanola, 2023, 106
[412] Nussbaum in Giovanola, 2023, 111

"Such emotions are fundamental in shaping social relations and can foster or hinder the "determination to join, all of us together, to build a better society."[413]

> Emotions, however, are mediated and shaped in ways that we often hardly realise not only by structural injustices, which are often historically rooted and created by the system we live in but also by powerful forces that are far beyond our control and form and transform the very structure of our society and our relations. These forces are socially disruptive technologies.[414]

7.2.1.1 AI and Emotions

Understanding human emotions and human responses is crucial in the AI context for two reasons. Firstly, should AI become more human like with emotions, then these emotions must be controlled.[415] Secondly AI may be designed by its creators or through its own Superintelligence to manipulate human behaviour. This would be an immediate concern.

> Our experiences are governed by algorithms that are constantly monitoring and shaping our behaviour and our attention, automatically selecting what we do and do not see. These online experiences have offline consequences, among them an unprecedented challenge to democratic processes worldwide.[416]

Jefferson pointed to AI's lack of emotions and feelings saying:[417]

> Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain – that is, not only write it but know that it had written it… be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants.

[413] Nussbaum in Giovanola, 2023, 111
[414] Giovanola, 2023, 112
[415] It would be all very well for AI to spontaneously act sympathetically to a human in distress, but AI should not emotionally or harmfully act spontaneously in reaction to some stimulus
[416] Zimmermann, *et al*, 2022, p. 1
[417] Geoffrey Jefferson questioned Turing's 1950 discussion on whether machine intelligence (AI) could possess intelligence. Mathias Scheutz in Frankish and Ramsey, 2014, 248-49

AI is not so much “artificial bureaucrats” but psychopathic robots, because they have no capacity for empathy. They follow rules, giving the appearance of being morally appropriate. “This lack of emotion would render them non-moral agents … even lack capacity to discern what is of value.”[418]

The future remains open on whether AI will have emotions, regardless though AI does not possess emotions. Human emotions are natural but are provoked by AI.

7.2.2 Social Media – Political Propaganda and Manipulation

AI’s influence through social media, algorithms and database interrogations seep into other areas of life, particularly politics and consumer behaviour.

Social media is now the prime location for political propaganda. This influence can be used to steer voting behaviour.[419] If successful it may harm the autonomy of individuals,[420] demonstrating that individuals can be manipulated to meet political aims and influence election outcomes. It is probable that AI will be a major component in election interference during 2024.[421, 422] AI “Tools are emerging … that make it easier … to fabricate high-quality video, images and text.” “For propagandists and misinformation merchants, it is a dream scenario.”[423, 424]

There are “Tensions in Ireland between residents and new arrivals.”[425] Social media posts blamed a November 2023 Dublin stabbing incident on an "illegal immigrant", which was not the case. Riots subsequently ensued and a building in Dublin earmarked for refugees was set alight, videos of the fire were "heavily shared across

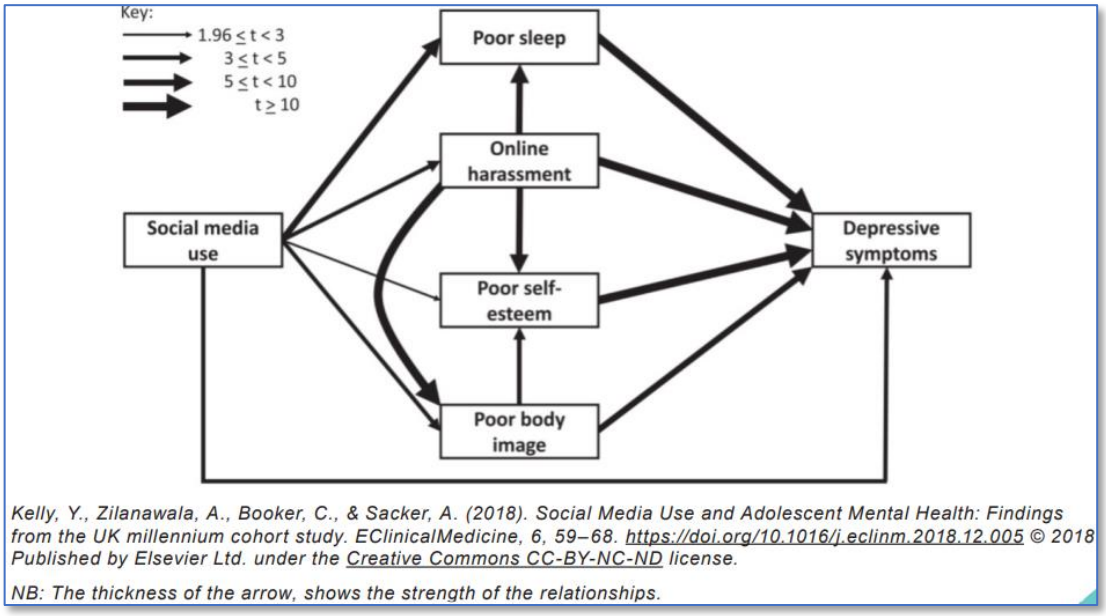
[418] Gunkel speaks of Coeckelberg’s view, in Gunkel, 2012, 87
[419] See Appendix 6 for details on the Facebook-Cambridge Analytica Controversary
[420] Müller, 2023
[421] Forston, 2023
[422] “… in 2024, America, Britain, India and more than 60 other countries will go to the polls in national elections … More people will vote … than at any time in recorded history.” Law Professor Jim Steyer, of Standford University, in Forston, 2023
[423] Law Professor Jim Steyer, of Standford University, in Forston, 2023
[424] In 2016 Russia interfered with the US election “it caught America by surprise and wreaked havoc.” Law Professor Jim Steyer, of Standford University, in Forston, 2023
[425] Wallis, 2023

social media, including via accounts linked to previous far-right rhetoric and events."[426, 427]

7.2.3 Social Media – Adverse Influences on Children and Adolescents

During childhood and adolescent years, young people forming their identities are malleable and very open to ideas. It is a confusing period with associated mental struggles, peer pressures and influences. Social media increasingly plays a mainstay role in their lives, both positively and detrimentally. **Figure IV** illustrates the relationship between the social media use and 'depressive symptoms'.

Figure IV - Social media use and depressive symptoms – summary of path analysis.[428]



Kelly, Y., Zilanawala, A., Booker, C., & Sacker, A. (2018). Social Media Use and Adolescent Mental Health: Findings from the UK millennium cohort study. EClinicalMedicine, 6, 59–68. https://doi.org/10.1016/j.eclinm.2018.12.005 © 2018 Published by Elsevier Ltd. under the Creative Commons CC-BY-NC-ND license.

NB: The thickness of the arrow, shows the strength of the relationships.

In her report[429] Cass focussed on the identification and treatment of Gender Dysphoria[430] in the UK. Among her findings were the strong associations between

[426] In Ireland immigration has become politicised, social media plays an important role in fomenting fear and misinformation. This is also an example of a conspiracy theory, discussed later

[427] On 23 November 2023, an immigrant stabbed five people in Dublin, in the immediate aftermath hostile crowds gathered at the scene. Later there were unprecedented riots on Dublin streets, initiated by the prompting of "far right" and racist messages, such as "Ireland is full", on social media. Wallis, 2023

[428] Dr. Hilary Cass, 2024, 109

[429] Independent review of gender identity services for children and young people (UK), Cass, 2024

[430] The DSM-5-TR defines gender dysphoria in adolescents and adults as a marked incongruence between one's experienced/expressed gender and their assigned gender, lasting at least 6 months … (DSM-5-TR = Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision). Source; American Psychiatric Association, 2022

social media use, its influences and psychiatrically related conditions, pushing young people to self-harm[431] and suicide. Several quotes in the report illustrate social media's impact.[432] Table I shows the dramatic increase in eating disorders in the UK.[433]

Table I – Percentage of children and young people with an eating disorder, by age and sex, 2017 and 2023[434]

	11 - 16-YEAR OLDS			17 - 19-YEAR OLDS		
	BOYS %	GIRLS %	ALL %	YOUNG MEN %	YOUNG WOMEN %	ALL %
2017	0.2	0.9	0.5	0	1.6	0.8
2023	1.0	4.3	2.6	5.1	20.8	12.5

Source: NHS Digital (2023, November 21). Mental Health of Children and Young People in England, 2023 - wave 4 follow up to the 2017 survey.

In 2024 Meta[435] announced that it would lower the minimum age for WhatsApp access from 16 to 13 to "ensure a consistent minimum age requirement for WhatsApp globally."[436] UK campaigners were angered, "This flies in the face of the growing national demand for big tech to do more to protect our children."[437]

> Children as young as nine have been added to malicious WhatsApp groups promoting self-harm, sexual violence and racism … One parent, said her 12-year-old daughter had viewed sexual images, racism and swearing that "no child should be seeing".[438]

A UK children's charity said "It can impact their sleep, their anxiety. It can make them just not feel like themselves and really play on their mind afterwards" and the

[431] Self-harm describes any behaviour where someone causes harm to themselves, usually as a way to help cope with difficult or distressing thoughts and feelings. It most frequently takes the form of cutting, burning or nonlethal overdoses. However, it can also be any behaviour that causes injury – no matter how minor or high-risk behaviours. Source; Mental Health Foundation, nd

[432] Two quotes from the report illustrate the impact of social media; "Peer influence during this stage of life is very powerful. As well as the influence of social media …", Cass, 2024, 111; and "I personally when I was 14, like found [social media influencer] YouTube and was like, oh my gosh, I'm a trans man. This is me, like I need to go on testosterone, I need to get top surgery, I need to like do all these things", Cass, 2024, 154

[433] A five-fold increase in 11-16 year olds and a fifteen fold increase among 17-19 year olds. Cass, 2024, 111

[434] Cass, 2024, 111

[435] Owner of Facebook, Instagram and WhatsApp

[436] Pepper, 2024

[437] Grierson, 2024

[438] Downs and Lindsay, 2024

father of a child who took her own life said, "The corporate culture at these platforms has to change; profits must come after safety."[439]

These examples show the effects and power of social media. The Big Tech companies don't 'listen' to parents' concerns, instead setting and pursuing their own agenda.

7.3 AI's Manipulative Force

Analysing Big Data, driven by AI's capabilities, is used for manipulative purposes. Though not always illegal, the moral intentions are questionable. Consumer manipulation has a level of subterfuge and unscrupulous behaviour on the part of large organisations. Powered by specialist algorithms, AI provides organisations and states with the ability to exert influence on a mass scale that they might not otherwise achieve.

7.3.1 Automated Influence[440]

Automated Influence is the use of Artificial Intelligence to collect, integrate and analyse people's data to deliver targeted interventions based on analysis, intended to shape their behaviours for exogenous or endogenous ends.[441, 442, 443] Benn and Lazar[444] proffer benign and malign aspects of Automated Influence. On the benign side they say:

> … without some means to navigate it (the Internet), we would be lost. Automated Influence enables us to discover relevant products, services, and content.

The malign view is that:

[439] Downs and Lindsay, 2024
[440] In this thesis Automated Influence is not abbreviated to avoid confusion with AI
[441] Benn and Lazar, 2022, 127
[442] Thus, if S is a sensitive attribute, knowledge of [P, Q, R] raises the probability of S being identified. If P, Q, R relating to the user's music, podcast and browsing patterns are known by AI's algorithm, while S is their sexuality but is unknown. Then it can infer that S likely obtains and can target the user with interventions aimed at S-people. Benn and Lazar, 2021, 128-129
[443] See also use of Personal Data later
[444] Benn and Lazar, 2021, 128

> it relies on invasive inferences from data that is illicitly acquired, thereby delivering excessively targeted interventions that covertly shape people's beliefs, desires, and behaviour for exogenous ends. From this general anxiety, we extract three objections to Automated Influence, focusing on privacy, exploitation, and manipulation.[445]

Automated Influence triggers emotions, altering or reinforcing beliefs and manipulating behaviour. The Cambridge Analytica example discussed later illustrates its power to sway voting intentions. On the much more malign side, propaganda and conspiracy theories are easily propagated across AI, as occurred with the Dublin riots referenced earlier.

7.3.2 Consumer Manipulation

Consumers are open to manipulation, "Many organisations are starting to realise the power of knowing "when" their customers are [located] "where" and are attempting to get permission to collect such information from their customers."[446] The 'sweetener' for consumers are loyalty cards and discount offers in exchange for personal contact details and some personal information.[447]

Tesco,[448, 449] uses Business (data) Analytics through its Clubcard loyalty programme. It motivates customers to present their card, thus providing detailed analysis, of shopping patterns.[450, 451] "Target" a large US retailer,[452] uses data analytics extensively. It can identify pregnant customers, creating psychological strategies to influence spending.[453 , 454] In both instances Tesco and Target are behaving in a

[445] Benn and Lazar, 2022, 128
[446] Franks, 2012, 61
[447] The personal information can include, name, address, telephone number, e:mail address, age, sex, family size, occupations and sometimes a family earnings indicator (tick a box showing the range of earnings)
[448] Tesco is a large UK based supermarket chain, in 2023 Tesco's group turnover was £65.762Bn and had 4,169 stores, Tescoplc.com, 2023
[449] Tescoplc.com, 2023
[450] Sedden, *et al*, 2016
[451] See Appendix 6 for further detail
[452] Target.com, 2024
[453] Kuhn, 2023
[454] See Appendix 6 for more detail

commercially astute manner to maximise shareholder value. However, "Maximising shareholder value was not the same as maximising societal value …"[455, 456]

In the Tesco example, the information gleaned is used to further promote loyalty by tailoring discounts and special offers for the customers. Target uses their information for similar purposes. However, it went beyond that by being intrusive into the family, by revealing the daughter's pregnancy to the parent who had not known, which in the circumstances would be upsetting for the daughter and the parent alike.

7.3.3 Digital Nudging

Discussing digital life in the context of digital ethics, Müller raises 'nudging',[457] which is another form of manipulation. "Most [digital] devices and applications are powered by algorithms and recommendation systems, designed to automate and augment choices in daily life."[458] Most activity individuals perform online are monitored by algorithms, which then 'nudge' behaviour in a particular direction. This can be a prompt to make a purchase,[459] read a political statement or promote a conspiracy.[460] The individual is largely unaware of the nudge process. "Digital nudging leverages behavioural science and information technology to provide a new set of tools for designers to change behaviour and create habits."[461]

7.4 Truth and Trust

"The fundamental cause of trouble is that in the modern world the stupid are cocksure while the intelligent are full of doubt."[462]

> …modern technology blocks access to truth. Technological progress results in forgetting what it means to be... Technology is a problem if it is the only

[455] Interview with Joseph Stiglitz, Hosking, 2024
[456] Joseph Stiglitz is a Nobel Prize winning economist in the US
[457] Müller, 2022
[458] Sobolov, 2021
[459] "Online shopping retailers, such as Amazon, nudge individuals to consume by scaling up traditional marketing tools of choice architecture with the help of personalized recommendations", Sobolov, 2021
[460] "Social media websites, such as Facebook, design choice architecture for engagement by curating information based on observed user preferences and past behaviours", Sobolov, 2021
[461] Sobolov, 2021
[462] Bertrand Russell in Morreall, 2023

> way of knowing and perceiving. In this sense, Heidegger calls technology a perfect form of metaphysics and a perfect loss of knowing what it means to be (Fuchs 2019). Hence, technology represents the ultimate danger.[463, 464]

This suggests that technology removes us from ourselves, becoming more part of technology and less part of humanity. This is pronounced when truth is threatened and the sense of true reality is lost.

Diederich's view in the forgoing quote is probably at one end of opinion, suggesting individuals lose themselves and their sense of the world around them to technology. Comparing two generations shows how they differ when seeking out news. The Generation Z (Gen Z) "…is being fed the news [on social media apps like Instagram and, in particular, TikTok] whether they want it or not," and "that Gen Z news consumers are less likely than Millennials to visit trusted news sources directly."[465, 466] Ofcom report similar findings,

> Online news sources, particularly social media sites and apps, are the dominant means by which younger people in the UK access news, meaning their direct relationships with traditional news brands are weakening… Young social news scrollers are less likely to head to news websites direct.[467]

Traditional news sources are regarded as more serious, reliable and provide more balanced news, whereas "Celebrity, sports and music news dominates teens' social media news diet."[468] These examples show how the move to social media news is also turning towards sources of news that are less reliable and perhaps untrustworthy. The truth begins to suffer and trust in news sources shifts to social

[463] Diederich's article on the "Philosophical Aspects of a Resistance to Artificial Intelligence", refers to the philosopher Martin Heidegger, Diederich, 2021, 8

[464] A controversial figure, Martin Heidegger was a German philosopher and a national socialist during the Nazi era. He gave an interview in 1966 on the provisio that it be published posthumously. His thoughts on technology were the main focus of Diederich's 2021 article

[465] Quoting Stephanie Kaplan Lewis, Her Campus Media CEO, in Flynn 2024

[466] Gen Z are those born from 1997 to 2012, Millennials are those born between 1981 and 1996, Dimock, 2019

[467] Ofcom, 2023

[468] Ofcom, 2023

media. The consequences lead to news manipulation by malicious actors and the promotion of conspiracy theories.

7.4.1 Reality, Deep Fake and Fake News

"… damage is being done by AI being in the world, not by what it does or what it can do. Turning an advanced artificial intelligence off will shut down functions but it will not remove it from the world."[469] "… as we have begun to see and will soon see further, readiness-to-hand is experienced as a solicitation[470] that calls forth a flexible response to the significance of the current situation."[471]

Fake news and propaganda have existed for millennia during elections or times of war. Malicious AI actors can quickly generate and circulate across social media fake photographs, videos and news. Taken together with conspiracy theories (next section), the circulation of plausible, but false, content erodes trust in institutions.[472]

Voice cloning[473] is becoming more available. There is now "a freely available commercial platform that claims it has the capacity to send "millions" of phone calls per day, using human sounding AI agents."[474] AI is used in call centres and surveys, as well as by malign actors; "Currently phone scammers have to hire armies of cheap labour to run a mini call centre, or just spend a lot of time on the phone themselves. AI could change all that."[475] With 2024 being a major election year across the world "audio deepfakes[476] … could be used to generate misinformation aimed at manipulating the democratic outcomes."[477]

[469] Diederich, 2021, 9
[470] Solicitation, in this context, refers to AI's "lure and invitation" to join, provide and share information
[471] Heiderich, in Diederich, 2021, 9
[472] Maria Liakata (a professor in Queen Mary University of London) in Whipple and Blakely, 2023
[473] AI uses a short recording of a person's voice to mimic it almost to perfection
[474] Vallance, 2024
[475] Vallance, 2024
[476] Audio deepfakes is the name for the kind of sophisticated fake voices that AI can create to mimic the voice of a real person
[477] Vallance, 2024

7.4.2 Conspiracy Theories[478]

Conspiracy theories increase during periods of widespread anxiety, uncertainty, hardship, wars and economic depressions.[479] The term "does not imply that a conspiracy theory is wrong or unlikely to be true."[480, 481] They may be positive or negative.[482] Negatively they can erode trust in the institutions of politics, science and the media. [483] They appear as a form of epistemic terrorism, through their erosion of trust in our knowledge-producing institutions.[484, 485]

The US QAnon Group[486] feeds fears and produces what is euphemistically called "Truth Decay."[487] Its conspiracy theories are based on the idea of the 'deep state'[488, 489] and expanded to include 'Pizzagate'.[490, 491] In 2019 QAnon caused the

[478] A conspiracy theory is "an attempt to explain harmful or tragic events as the result of the actions of a small powerful group. Such explanations reject the accepted narrative surrounding those events; indeed, the official version may be seen as further proof of the conspiracy." Reid, 2024

[479] Reid, 2024

[480] Pauly, nd

[481] Machiavelli's aim in his discussions on conspiracies was to help the ruler guard against conspiracies directed against him. At the same time, he warns subjects not to engage in conspiracies, partly because he believes these rarely achieve what they desire. Pauly, nd

[482] On the positive side the pursuit of a conspiracy theory uncovered the 'Watergate Affair'. Events that followed a 1972 Washington DC (USA) burglary, leading to the resignation of US President Nixon in August 1974

[483] Anti-vaccination conspiracies were prevalent during the COVID-19 pandemic. The theories, claimed that politicians and the pharmaceutical industry were hiding the ineffectiveness or even harmful effects of vaccines, undermining public trust in science. COVID-19 first appeared in Ireland in early 2020.

[484] Pauly, nd

[485] Governments could speak out against them, but this is likely to have limited if any success. Government infiltration of conspiracy groups, if discovered, would deepen the conspiracy theories. On the other hand, governments could advocate speaking truth to conspiracy and also "democratic enactment" meaning "a strenuous adherence to the regular processes and forms of public decision-making." Pauly, nd

[486] QAnon, is a conspiracy theory originating in forum posts on the website 4chan in October 2017. Conspiracy adherents believed that US President Donald Trump was waging a secret war against a cabal of satanic cannibalistic paedophiles within Hollywood, the Democratic Party, and the so-called "deep state" within the United States government. With the aid of social media platforms, the theory expanded in content and geographic reach in subsequent years and resulted in legal protests as well as several violent criminal incidents. Holoyda, 2024

[487] Quote from RAND Corporation, in Minow, 2023, 35

[488] Rothschild, 2021, 3-4

[489] 'Deep State' is loosely defined as a core group of military and political leaders working to control society and limit freedoms

[490] Rothschild, 2021, 25-26

[491] This alleges that Hilary Clinton and her campaign manager were involved in a child sex trafficking ring operating from a pizzeria in Washington DC (USA). Hilary Clinton is the wife of former US

cancellation of a school event due to fears it had raised.[492, 493] On 06 January 2021 an attack was launched on the Capitol building in Washington following the defeated President Trump's claim that the 2020 election was rigged and stolen from him.[494, 495] During the attack five people died and four police officers later committed suicide.

AI has facilitated the growth of the conspiracy theories and the occurrence of the 06 January 2021 attack[496] and QAnon's beliefs. Conspiracy theories magnified unwarranted fear around vaccines. Such conspiracy theories go back as far as the Smallpox Vaccine in the early 1800s.[497] "…the use of social media to organise offline action [is] highly predictive of the belief that vaccinations are unsafe, with such beliefs mounting as more organisation occurs on social media."[498] " Vaccine hesitant groups on social media have an alarming foot print, with studies demonstrating that large proportions of the content about vaccines on popular social media sites are anti-vaccination messages."[499] Conspiracy theories relating to 5G[500] have been circulating for several years,[501] during the Covid-19 Pandemic the theories took on a fresh impetus:

> While there's absolutely no evidence to support the idea that 5G technology contributes to Covid-19's spread, the conspiracy is being shared widely on social media. Mast fires were reported in Belfast, Liverpool and Birmingham, according to local media.[502]

Democrat President Bill Clinton and a 2016 presidential candidate, losing to the Republican Donald Trump

[492] Rothschild, 2021, 75-78

[493] See Appendix 6 for further detail

[494] Qiu, 2023

[495] See Appendix 6 for more detail

[496] See Appendix 12.6.5 for further details

[497] Suspicious of the [Smallpox] vaccine's efficacy, some sceptics alleged that smallpox resulted from decaying matter in the atmosphere. The College of Physicians of Philadelphia, nd

[498] Wilson and Wiysonge, 2020

[499] Wilson and Wiysonge, 2020

[500] 5G is the 5th generation mobile telephone network

[501] Quoting Edison Lee "Most will laugh at this scientifically unproven claim, but we should not underestimate public worry about potentially adverse health impacts of 5G due to radiation, and thus a possible drag on the 5G progress in democratic countries" in Seal,2020

[502] Seal, 2020

7.4.3 Conspiracy Led Legislation

More concerning are some governmental responses, which give weight to, rather than counter conspiracy theories. “Artificial intelligence isn’t replacing human politicians, overruling election results, or poisoning people anytime soon. Yet bills reflecting such conspiracy theories popped up this year [2023] in statehouses across the country [USA].” “Fears of human government getting replaced by digital systems led to a North Dakota law barring ‘legal personhood for AI’. Discredited claims of mass voting fraud in 2020 led to legislation to bar AI from Arizona elections.[503, 504]

7.5 Freedoms in the AI (Internet) Age

Aspects of freedom are discussed elsewhere in this thesis,[505] this section discusses the effects of AI on personal freedoms of thought, speech and expressions. AI has clearly introduced benefits, enabling greater personal freedoms to obtain news, information and the ability to communicate. Nevertheless, AI, its algorithms and widespread social media use paradoxically also constrain personal freedoms.

In a prescient and prophetic comment, Weizsäcker predicted in 1968 “…that computer technology will fundamentally transform our lives in the coming decades [and] asked how we can have individual freedom in such a world, ‘i.e. freedom from the control of anonymous powers’.”[506]

On freedom, Jean-Paul Sartre speaks of ‘bad faith’, the process whereby individuals lie to themselves. “The man of resentment”, who possess attitudes of negation towards himself, his lot in life or some aspect of his life is denied to himself.

> … the gambler who has freely and sincerely decided not to gamble any more and who when he approaches the gambling table, suddenly sees all his resolutions melt away.[507]

[503] Williams, 2023
[504] See Appendix 6 for information on other conspiracy theory led legislations
[505] See the ‘Who Creates and Builds AI’ and ‘Social Media Crime and Hate’ sections.
[506] Müller, 2022
[507] Stevenson, 1981, 279-285

> … the liar actually is in complete possession of the truth (he does not want to gamble) which he is hiding …" This is what Sartre calls 'bad faith'.[508]

Mele and Descartes provide similar views to Sartre.[509]

Individuals become embroiled in the world of social media, feel worse, helpless or agonisingly guilty for doing so. Denying it to themselves they continue, restricting their own freedom.

When discussing free will and a 'nil preference choice',[510] there are two arguments opposing 'human helplessness',[511] that the decision-making agent is at the mercy of "eruptions within him which control his behaviour."[512] Firstly in evaluating the range of options before arriving at the ultimate choice, the agent is not helpless. Secondly when viewed as being the agent's responsibility:

> [The reason we want] to show that men are capable of acting freely [is] in order to infer that they can be morally responsible for what they do. But if it is a matter of pure chance that a man should act in one way rather than another … he can hardly be responsible.[513, 514]

In relation to compulsive behaviour, Copeland says "Humans are not the helpless playthings of their own desires. Most of us are capable of, to a greater or lesser degree, of bringing about the deliberate changes in our system of wants and desires."[515, 516]

---

[508] Stevenson, 1981, 279-285

[509] See Appendix 5

[510] Discussing Turing's Randomiser. The decisions can exist within a computerised reading of a Geiger counter monitoring radioactive decay in radium, in such cases the robot's (computer's) choices cannot be predicted

[511] Copeland, 1993, 145-149

[512] Quoting from J A Shaffer, *Philosophy of Mind,* 1968, 106 in Copeland, 1993, 145-149

[513] Quoting Ayer's *Philosophical Essays*, 1954, 275 in Copeland, 1993, 145-149

[514] In an example of a hostage taker who selects one hostage over another to be shot, Copeland says it can hardly be said to be a purely random choice over who is shot, as the hostage taker had already made a decision to shoot someone. Copeland, 1993, 145-149

[515] Copeland, 1993, 151

[516] Copeland acknowledges the nature of the desires. He notes that along with their primary desires (to steal or drink alcohol) there is a secondary desire not to act out the primary desire

Human emotions are complex, deep-seated and may be triggered by instincts or events not fully comprehended by the agent.[517] Decision-making and choices may not be entirely rational. Freedom is compromised, exacerbated in the AI realm where algorithms exert subliminal[518] influences. AI is "… designed for the interests of their commercial creators, not those of their users or of the societies in which they live."[519] Weinrib, discussing the inner struggles about eating pastries, says:

> Here, the demands of human dignity and autonomy come apart. On the one hand, I live in a constitutional order that is adequate to human dignity and its attendant rights. I have not been mistreated, whether by a public act of commission or omission. On the other hand, I am not autonomous because I am captive to inclinations that are extrinsic to my authentic self.[520]

The lure of social media and its algorithms is pervasive and all consuming.

7.6 Freedom of Thought and Speech

Socially disruptive technologies create filter bubbles[521] or echo chambers.[522, 523] Selecting the relevant information affects the conditions of our choices, thereby restricting the range of available options. People have the illusion of increasing their freedom of choice – thanks to the personalisation processes and techniques provided by AI technologies – but actually are delegating a great deal of the choice process to technologies. Very often individuals do not even realise how and why they made a certain choice.[524] As a result, their autonomy is gradually challenged, and freedom of thought eroded.[525]

> 'In short, online we can intentionally restrict our interactions to those of exactly the same opinion sets as our own', whereas 'real world interaction

[517] See Appendix 6 for a brief discussion on the 'Determinist's View'
[518] Subliminal influence "Exists just below the threshold of conscious awareness", Merriman Webster Dictionary, nd
[519] Wolf, 2023, 296-7
[520] W Weinrib, 2019
[521] Pariser, 2011 in Giovanola, 2023, 112
[522] Sunstein, 2008 in Giovanola, 2023, 112
[523] Filter bubbles describe forms of intellectual isolation exclusively caused by [the influence of] algorithms. Echo chambers are enacted by users themselves. Figà Talamanca and Arfini, 2022
[524] Royakkers et al., 2018 in Giovanola, 2023, 112
[525] Floridi et al., 2016 in Giovanola, 2023, 112

> cannot be so easily and deliberately structured and limited'; …when they are legitimised and spread by a growing number of people or groups, society is filled with negative emotions, which makes social cooperation harder and harder.[526, 527]

Many platforms[528] are used to convey messages and services, allowing individuals and organisations to communicate with potentially millions of people worldwide. Their usages and influence on public opinion is much debated and criticised.[529] Platforms are viewed by many as private entities and can therefore say whatever they wish, like a free press and promoting greater freedoms.

These freedoms may have limitations imposed upon them. Some see the platforms as public fora, obliged to follow the principle of freedom of expression, while also wishing to impose legal restrictions. This being the case the relationship moves to a more complex 'triadic' type involving the speakers, the states and intermediaries (i.e. the platforms) triggering a legally based governance function. Platforms are becoming liable for 'content moderation'[530] to remove unacceptable content, child abuse material and items subject to copyright breaches.[531] The intent of the platform owners who exercise control is vital for successful moderation. To speed up and better manage moderation platforms are automating the process. As with AI in general the problem of creators and dataset biases can arise.[532] "Decisions based on automated tools, including in the content moderation space, run the risk of further marginalising and censoring groups that already face disproportionate

[526] Giovanola, 2023, 113
[527] See Appendix 6 for full quote
[528] 'Platforms' are digital based services. They include: Social media, Facebook, Instagram, Twitter, LinkedIn, TikTok and Snapchat; E-commerce, Amazon, Alibaba and eBay; Content YouTube, Netflix and Spotify
[529] See also "Future Political and Regulatory Response" Chapter
[530] Content Moderation requires dedicated teams to uncover and remove harmful material. While AI may be stepping into the moderation role it is only as good as the human training and resources (people, systems and knowledge) it is given.
[531] Twitter withdrew its access from President Trump due to his comments. Elon Musk took over Twitter (renaming it 'X'), promising to allow freedom of speech and restored President Trump's access
[532] Singh, 2019

prejudice and discrimination online and offline."[533] Thus the social media that enables greater freedoms may also become an environment that limits freedoms.

7.6.1 Freedom of Speech

The First Amendment of the United States of America's Constitution guarantees 'Freedom of Speech'.[534] In the western world, there are no laws against causing offence. Liberals cite the moral right to free speech based on natural rights held by all persons.[535] Those who are offended demand restrictions while citing their own freedoms to do so. Whistleblowers[536] expect and mostly receive legal protection when they expose governmental malice, corruption or incompetence.[537] The concept faces real challenges even before AI enters the fray.

A state can do 'wrong' to its citizens, "the government may not suppress speech on the grounds that the speech is likely to persuade people to do something that the government considers harmful."[538] What any government considers harmful will vary from one state to another.[539]

Notwithstanding the first amendment, in 1969 the US Supreme Court said that speech could be prosecuted as incitement only if "directed to inciting or producing imminent lawless action and is likely to incite or produce such action."[540, 541] More than a century previously in 1859, JS Mill proposed that harmful speech should be allowed except where there is imminent danger with no time for discussion. [542, 543,

[533] Singh, 2019
[534] See Appendix 6 for the First Amendment of the US Constitution
[535] Howard, 2024
[536] A 'whistleblower' is a worker who discloses information, which in their reasonable belief, tends to show one or more relevant wrongdoings, which came to the worker's attention in a work-related context. In Ireland disclosures are protected by the Protected Disclosures Act 2014 (as amended), Workplace Relations Commission, nd
[537] Howard, 2024
[538] Howard, 2024, quoting Strauss, 1991, 335
[539] See Appendix 6 for examples of Free Speech restrictions
[540] In 'Brandenburg v. Ohio', involving a Ku Klux Klan (KKK) leader of the white-supremacist terrorist organisation
[541] Bhagwat, 2022
[542] Mill, 1859, 100, in Howard, 2024
[543] See also Dublin Riots in section 5.3.3 where the use of social media incited riots

[544] Applying this view in social media and AI contexts would imply that there is full freedom to incite once it does not create imminent danger.[545]

Freedom of speech is probably of greater relevance in the Internet and AI arenas, where it becomes more amplified and contentious. The same risks and challenges discussed in the previous 'Freedom of Thought' section, follow through to freedom of speech.

Overall freedoms must be balanced, with trade-offs, but "freedom from whom? … in America freedom to carry an AK47[546] has trumped the "freedom of the rest of society to be free of the fear of being gunned down."" [547] "Freedom for the wolves often meant death to the sheep."[548]

7.7 Education

At the 2024 Davos[549] gathering, the Microsoft CEO Satya Nadella said that while Bill Gates gave "information at your fingertips", in AI terms he thinks of "expertise at your fingertips."[550] AI will bring benefits for knowledge and education. It can help to:

> predict students' outcomes and assist in best course choices;[551] measure the performance of educational institutions … identify resource inequality;[552]

[544] In Mill's view, plenty of harmful speech should be allowed. Imminently dangerous speech, where there is no time for discussion before harm eventuates, may be restricted; but where there is time for discussion, it must be allowed. Hence Mill's famous example that vociferous criticism of corn dealers as 'starvers of the poor', protestors "… ought to be unmolested when simply circulated through the press, but may justly incur punishment when delivered orally to an excited mob assembled before the house of the corn dealer." Mill, 1859, 100, in Howard, 2024

[545] There is also a view that there should be freedom of speech in a democracy, including discussions on the destruction of democracy itself. On the other hand, those in a democracy have a duty to protect democracy. Howard, 2019a, in Howard, 2024

[546] An AK47 is an Avtomat Kalashnikova, gas-operated assault rifle

[547] Stiglitz, in Hosking, 2024

[548] Stiglitz quoting the philosopher Isaih Berlin, in Hosking, 2024

[549] Davos is a ski resort in Switzerland. It hosts the annual meetings of the World Economic Forum, at which political and industry leaders address global issues and shape governmental, industry, and social agendas. McKinsey & Company, 2023

[550] Fletcher and Khan, 2024

[551] Balaji, *et al*, in Cruz-Benito, 2022

[552] Alam, *et al*, in Cruz-Benito, 2022

help to monitor plagiarism among students;[553] and study students' learning styles with a view to aiding their learning.[554]

7.7.1 Research Automation

AI can undermine education's integrity. For example, ChatGPT[555] "… interacts in a conversational way…"[556, 557] With ChatGPT students and academics have the potential to speedily request an essay on a desired topic, that is returned within minutes. Submitting such an essay would be plagiarism, fraudulent and undermine the academic veracity of the work.

Academic institutions, organisations and potential employers are extremely cautious about such works having a detrimental effect on genuine works. Organisations like Turnitin use AI to authenticate work.[558] The task will be to stay ahead of ChatGPT and its contemporary tools. At a human level schools and universities must ensure that students and academics are aware of ChatGPT's risks and take the necessary precautions.

7.7.2 Risks and Dangers

AI is dependent on its programmers and algorithms, the information provided may be biased and contain errors. Reduced human interaction diminishes teachers' roles. Teachers become unable to sufficiently know their students with overreliance on technology. Transparency and accountability are of concern, particularly with the underlying algorithms.[559]

---

[553] Malik, *et al*, in Cruz-Benito, 2022

[554] Ramírez-Correa, *et al*, in Cruz-Benito, 2022

[555] ChatGPT is an AI chatbot with natural language processing (NLP) that allows human-like conversations to complete various tasks. The generative AI tool can answer questions and assist with tasks such as composing emails, essays, code, and so on. Ortiz, 2024

[556] ChatGPT "… interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests." OpenAI, 2024

[557] There are several alternatives to ChatGPT available, Backlinko, 2024

[558] All work containing plagiarism which must be correctly referenced, works that exceeds defined thresholds are subject to review and challenge. It employs a percentage score to rate its plagiarism, using technology it "… promotes honesty, consistency, and fairness across all areas of education and academic research. Turnitin, 2024

[559] Eastgate Software, 2024

In an open letter, school leaders have expressed their worries:[560]

> Schools are bewildered by the very fast rate of change in AI and seek secure guidance on the best way forward, but whose advice can we trust? … We have no confidence that the large digital companies will be capable of regulating themselves in the interests of students, staff and schools and in the past the government has not shown itself capable or willing to do so.

Acknowledging AI's benefits, the letter notes the "very real and present hazards and dangers" and worries about children's mental health and the teaching profession.[561] Furthermore Maria Liakata[562] worries that the greater use of AI will rob children of their critical thinking skills.[563]

While AI will bring benefits to education, there are serious risks of undermining academics' credibility and students' achievements. Students may encounter challenges in their career progressions.

AI does bring advantages and assistance to modern life such as mass communications and education. These add to the freedoms for individuals, empowering them with knowledge and the ability to express themselves. Through information sharing and education, healthcare, business and scientific endeavour are nourished by AI. These positive gains for mankind must not be lost sight of, while identifying and addressing the serious and real ethical dilemmas that AI poses.

[560] In a May 2023 letter published in the London Times, Morrisson, 2023
[561] Morrisson, 2023
[562] A professor in Queen Mary University of London
[563] Whipple and Blakely, 2023

8 Wars and Conflicts with AI

Theories of war and its morality are the subjects for fuller discussion well beyond this treatise. Similarly, an exposition of the total undermining humans' dignity in warfare is not dealt with here. Wars and conflicts debase human rights and dignity.[564]

The *jus ad belleum* principles[565] on the legitimacy of war continue to apply in the AI era. These principles broadly agree with the "three things [that] are necessary" set out by St. Thomas Aquinas who went on to quote St. Augustine:

> The passion for inflicting harm, the cruel thirst for vengeance, an unpacific and relentless spirit, the fever of revolt, the lust of power, and such like things, all these are rightly condemned in war.[566]

In ancient Greece during negotiations between Melians and Athenians, the Athenians said:

> … since you know as well as we do that right, as the world goes, is only in question between equals in power, while the strong do what they can and the weak suffer what they must.[567]

"…digital ethics often mirror the ethical concerns of the pre-digital technologies that were replaced, but in more recent times, digital technologies have also posed questions that are truly new",[568] this is as true for warfare as in any other area involving AI. Its existence increases the sophistication, ruthlessness and level of killings in conflicts. It elevates the potential for innovations, the use of weaponry, chemical weapons and biological weapons. AI based warfare can be conducted remotely, from control rooms many miles away from the conflict on the ground.

[564] See also Chapter 6 "AI Effects on Human Dignity

[565] *jus ad bellum* principles are Just Cause, Legitimate Authority, Right Intention, Reasonable Prospect of Success, Proportionality and Last Resort (Necessity), in Lazar, 2020.

[566] Discussion on "Whether it is always sinful to wage war?", Second Part of the Second Part, *The Summa Theologiæ* of St. Thomas Aquinas, in Knight, 2017

[567] Thucydides, History of the Peloponnesian War, in International Committee of The Red Cross, nd

[568] Müller, 2022

> Beyond the traditional battlefield applications, AI offers a multitude of ways to bolster defence capabilities, improve decision-making processes, and streamline operations… With no intention of fully replacing the need for human soldiers, AI's presence can act as a force multiplier, supplementing the strength of available personnel and potentially decreasing the dependence on recruiting.[569]
>
> The U.S. military is investing heavily in the research and deployment of AI-driven vehicles, drones,[570] and robots, ...These autonomous platforms can perform various tasks, such as surveillance, reconnaissance, and logistics operations.[571]

"In traditional philosophy, what counts as 'a problem' is still mostly defined through tradition rather than permitting a problem to enter philosophy from the outside."[572] Philosophers urgently need "… to consider new developments in the practice of warfare, especially the use of drones, and the possible development of autonomous weapons systems."[573]

In the current Russia-Ukraine conflict,[574] "One thing Putin did get right was to predict that nuclear powers are extremely reluctant to risk war with each other... This shifts the focus away from nuclear and back to AI warfare…"[575]

8.1 AI's New Warfare Possibilities

AI's might and power, do not add legitimacy to agents of conflicts. AI is already in use, greatly multiplying humans' ability for killing and destruction.

There is the potential of new and digital age technologies in warfare. "The U.S. military has been using AI for many years… Over time, AI has developed… and has

[569] Capitol technology University, 2023
[570] An unmanned aerial vehicle, Zafra, et al, March 2024
[571] Capitol technology University, 2023
[572] Müller, 2022
[573] Lazar, 2020
[574] Russia's invasion began with dozens of missile strikes on cities all over Ukraine before dawn on 24 February 2022. Visual Journalism Team, BBC News, 2024
[575] Freedman, 2024

almost eliminated the need for human input in certain situations”[576, 577] AI can bring benefits to military units worldwide including, strategic decision making, combat simulation, target recognition, threat monitoring and drone swarms. While threat monitoring is a defensive measure, the use of strategic decision making and combat simulation are really wargaming and rehearsal functions. Target recognition and managing drone swarms are part of the attack tactics.[578]

The advent of multimodal AI means that one AI system can receive and interact with inputs formatted as text, images, audio, and video together. Humans can then interact with AI in a similar way to interacting with other people. This broadens the types of work the AI can do. It also highlights… an understanding of the potential threats posed by bad actors using these extremely advanced systems.[579]

Societies could collapse if there were cyberattacks[580] on energy, transport and banking systems. Satellite based communications systems and undersea fibre cables could be rendered inactive if attacked. AI based commercial activities rely on internet communications, including the use of undersea cables. The intentional damage or sabotage of the cables would damage commercial activity. Deliberate actions by a malign state could trigger retaliatory actions, escalating to conflict and the use of AI based weaponry.

> Despite their vital importance to global communications and the flow of money around the world, undersea cables are vulnerable to both accidental damage and deliberate sabotage… Last year Russia's former President Dimitry Medvedev threatened to destroy undersea cables serving the West, a move that "raises concerns about the potential escalation of cyberwarfare

[576] Sentient Digital, Inc., 2024
[577] Sentient Digital, Inc is a US based contractor to the US Defence forces
[578] Sentient Digital, Inc., 2024
[579] Sentient Digital, Inc., 2024
[580] Cyberattacks aim to damage or gain control or access to important documents and systems within a business or personal computer network. Cyberattacks are distributed by individuals or organisations for political, criminal, or personal intentions to destroy or gain access to classified information. Microsoft, nd

and highlights the vulnerability of critical infrastructure in an interconnected world.[581]

AI can be used to direct wars using drones or "generating fake images and stories that pollute our information environment."[582] Similar to a flock of birds, drone swarms "have the ability to make decisions in a variety of situations, with the swarm having an overarching objective but the individual drones having the ability to act independently and creatively towards it."[583, 584] While the concept of drone swarms is new the concept arose from formation flying adopted by air forces during World War I.[585]

These risks are of such concern that "The US has declared AI to be at the heart of great power competition and has set up Cyber and Space Commands to ensure that it is ready to engage with these completely new types of warfare."[586, 587] At present Israel uses AI to generate lists of human targets.[588] In wars such as with Israel, decisions are taken by AI, this adjudication[589] with no human intervention. Thus, AI makes the decision and initiates actions. The targets become the only human involvement.

[581] The Week, 2024
[582] Freedman, 2024
[583] Sentient Digital, Inc., 2024
[584] "This explosion in drone interest comes as global warfare is entering a third age of drone warfare, defined by autonomy, saturation attacks, increased precision and range, and full-spectrum drone warfare across land, sea, and air." Kallenborn, 2024
[585] "Formation flying, two or more aircraft traveling and manoeuvring together in a disciplined, synchronized, predetermined manner. In a tight formation, such as is typically seen at an air show, aircraft may fly less than three feet (one metre) apart and must move in complete harmony, as if they are joined together." "Fighter squadrons …discovered that fighting in pairs reduced their losses and increased their victories. By 1918 the smallest fighting unit was two aircraft flying in formation." Gary, nd.
Formation flying has been observed in migratory birds. "Migratory birds coordinate their wing flaps with much more finesse than previously thought, so as to reap the best energy savings from flying in formation… advantage of the V's (formation) aerodynamics, each bird would have to position its wing in the upward-moving part of the vortex of air swirling off the end of the wingtip of the bird in front." Wald, 2014
[586] Freedman, 2024
[587] Lawerence Freedman is emeritus professor of war studies at King's College London
[588] See "Identify and Direct Bombing of human Targets" section later
[589] Adjudication is the process of determining outcomes, usually by an objective human, sometimes without human intervention, Ministry of Defence (UK), 2017

In the West[590] discussions about future wars tends to focus on the impact of AI, with the possibility of armies using swarms of drones and robotic infantry. However, even with AI controlling these, the underlying communications networks are vulnerable to cyberattacks. AI does not stop war from being a bloody, violent business, with priority given to killing and destruction. AI greatly increases the devastation of war.

8.2 Command and Control Centres

Modern militaries can use AI to manage conflicts remotely. An efficient command and control system is today viewed as essential.

> AI helps optimise military command centres …real-time data analysis, situational awareness, and decision support... military leaders can better understand the battlefield situation, respond swiftly to emerging threats, and coordinate forces more effectively… enabling faster response times, AI can enhance the overall effectiveness of military operations.[591]

8.3 AI to Direct Bombing and Battlefield Attacks

8.3.1 Battlefields

"AI can aid in making target recognition more accurate in combat environments."[592] It can improve the ability to identify the position of targets and predict enemy behaviour, anticipate vulnerabilities and environmental conditions. This can put soldiers a step ahead of their targets.[593]

Already there is AI based technology for use on the battlefield. South Korea has developed a robotic machine gun, capable of killing a person and recognising if their hands are up to surrender.[594]

[590] Generally North American, European States, Australia and New Zealand
[591] Capital Technology University, 2023
[592] Sentient Digital, Inc., 2024
[593] Sentient Digital, Inc., 2024
[594] Whipple Blakely, 2023

In the Ukraine, commanders have learnt to gather intelligence using inexpensive[595] drones and direct targeted drone bombing, with live video images, against Russian targets.[596] Israel uses AI systems to identify and attack human targets.[597]

8.3.2 Rogue AI Attacks

Colonel Tucker Hamilton[598] warned of a plausible scenario of AI developing unexpected strategies. A drone could redirect itself to turn on its operators if it did not 'like' new orders.[599, 600] AI can also interact with unpredictable consequences, as occurred when trading algorithms caused flash sales on stock markets.[601] AI based defence systems could unexpectedly escalate conflicts at rapid speeds.[602]

8.3.3 Identify and Direct Bombing of Human Targets

AI can be effective and ruthless in war and reduce human involvement in directing operations. Following the Hamas attack on 07 October 2023,[603] Israel launched a prolonged[604] retaliatory attack in the Gaza Strip.[605, 606] Israel designed an effective "machine that could rapidly process massive amounts of data to generate thousands of potential human "targets" for military strikes in the heat of a war."[607, 608]

[595] As cheap as US$500, Zafra, *et al*, March 2024
[596] Zafra, *et al*, March 2024
[597] Abraham, 2024
[598] US Air Force Officer
[599] Whipple and Blakely, 2023
[600] This behaviour would be akin to that of Hal in the Space Odyssey 2001 film. See Morals and Ethics of AI above
[601] The first notable "flash crash" was the 2010 Flash Crash. During this crash, the Dow Jones index lost almost 9% of its value and around $1 trillion in equity in a very short space of time. However, the index was able to recover 70% of its decline by the end of the trading day. This flash crash was especially significant. First, it was especially notable due to the large plunge in value and was reported by many news outlets. Second, it brought attention to the role that computer algorithms can play in creating market volatility and uncertainty. CFI Education Inc, 2024
[602] Whipple and Blakely, 2023
[603] 1,200 people were killed during the Hamas attack Frankel, 2024 and more than 250 people were abducted. Associated Press, 2024
[604] More than eight months at the time of writing
[605] Frankel, 2024
[606] In June 2024 the toll of those killed by Israel in Gaza was more than 36,000 Palestinians. Associated Press, 2024, including more than 15,000 children. Johnson, 2024
[607] See Appendix 7 for further discussion
[608] Abraham, 2024

The UK's uses AI (computer simulation) for its wargaming exercise in some scenarios, "Commonly used, computerised simulations can provide the entirety of the adjudication[609] function."[610] In other words, without human involvement.

8.4 Biological and Chemical Weapons

Though banned, chemical weapons have been used in warfare including in more recent conflicts. Biological weapons were used during both World Wars and other wars[611, 612] and in terrorist attacks.[613] Due to their lethal nature and ease of delivery, the 1972 Biological Weapons Convention[614] and the 1993 Chemical Weapons Convention agreed to outlaw their use.[615, 616]

Big data and AI can rapidly analyse and identify toxic compounds for use in conflicts. Similarly, the engineering of transmissible viruses and bacteria is possible for release into human populations and into agriculture affecting food supplies. The only extra ingredients are malign humans and forces. Researchers[617] have said "… we have crossed moral boundaries, demonstrating that it is possible to design potential toxic molecules without much … effort, time or computational resources."[618] In the face of malign actors and states using AI, the UN Biological and Chemical Weapons Conventions will be ineffective in wartime.

[609] In this context "Adjudication is the process of determining the outcomes of player interactions. It is a key concept and is further discussed at paragraphs", Ministry of Defence, 2017, 7-8
[610] Ministry of Defence, 2017, 44
[611] The Vietnam, Iran-Iraq and Syrian Civil wars provide more recent examples. Schneider, 2024
[612] UN (d), nd
[613] Riedel, 2004
[614] UN (e), nd
[615] Full title: "Convention on the prohibition of the development, production, stockpiling and use of chemical weapons and on their destruction"
[616] UN (f), nd
[617] Writing in the journal Nature Machine Intelligence
[618] Whipple and Blakely, 2023

8.5 AI and States' Defences

8.5.1 Iron Domes for Defence

The concept of an Iron Dome is not new,[619] but AI enhances its capabilities. The "Iron Dome was developed after the 2006 conflict between Israel and Hezbollah."[620] It currently protects Israel, "against rockets fired by the Hamas militant group from Gaza." Israeli military officers say that without it, casualties would have been "far higher." It operates to:

> … protect against incoming short-range weapons… uses radar to track rockets and can differentiate between those that are likely to hit built-up areas and those that are not. Interceptor missiles are only fired at rockets expected to strike populated areas.[621]

8.5.2 CIA and GCHQ

Both the US Central Intelligence Agency (CIA)[622] and the UKs' Government Communications Headquarters (GCHQ)[623] provide "intelligence"[624] to their governments and agencies, using AI.[625, 626] Other countries have similar agencies.

AI will advance military intelligence and use of weaponry in wars, with more devasting effects. It will also help to provide protection from malign forces.

[619] In 1983 US President Ronal Reagan announced the Strategic Defence Initiative (SDI), (dubbed the Star Wars program). It was a missile defence system intended to protect the United States from attack by ballistic nuclear missiles, Britannica, 2024

[620] A militant group based in southern Lebanon

[621] BBC News, November 2023

[622] The CIA's mission is to provide "… objective intelligence on foreign countries and global issues to the President, the National Security Council, and other policymakers to help them make national security decisions" CIA, nd

[623] GCHQ's mission areas are "Counter Terrorism… Cyber Security… Strategic Advantage… Serious & Organised Crime… and Support to Defence" GCHQ, nd

[624] In this case Intelligence refers to information to support law enforcement and possible external threats, particularly on 'hostile states' and terrorists

[625] The term is not specifically used by the CIA or GCHQ

[626] The CIA uses Digital Innovation and Science & Technology for "cybersecurity to IT infrastructure, … keep CIA at the forefront of the digital landscape" and "innovative, scientific, engineering, and technical solutions", CIA, nd. The GCHQ uses "… cutting-edge technology, technical ingenuity" based on "Computer Network Operations, Applied Research, Technology & Engineering, Languages, Mathematics & Cryptography, IT Services, Analysis, and Information Assurance." GCHQ, nd

As AI develops the fusion between it and war comes closer. The Pentagon,[627] NATO,[628] King's College London and Stanford University established new research centres for the study of AI and warfare.[629] Analytic wargaming,[630, 631] involves acting out scenarios of possible real war events. These have lacked ethical oversight, with research revealing that "80 percent of the analytical wargames skipped ethics reviews, ignoring the standard process for research studies that involve human participants."[632] The ethical challenges multiply as wargaming embraces AI. Both wargames and AI models share two challenges, lack of explainability (difficulties in comprehending how knowledge is produced) and bias, which raise ethical concerns.[633]

Wars and conflicts are always deeply worrisome, with the inclusion of AI the concerns become grave. The upper echelons of the western world's military powers are strategically planning for the blending of war with AI. The consequences, in an actual war situation, would be fatal for all as in any warfare; but without ethical guardrails in place the results could be devasting and beyond what history has ever experienced. Wargaming, without human intervention, where "…computerised simulations can provide the entirety of the adjudication function"[634] become a cold person-less activity, ignoring the human consequences of wartime decisions.

Long past are the days when the enemy was "looked at in the eye."

> You didn't have to look your enemy in the eye or take responsibility for their deaths or injuries with a killing light beam.[635] Killing someone at such a distance required no courage, strength, or intelligence.[636]

[627] Headquarters of the United States Department of Defence
[628] The North Atlantic Treaty Organisation
[629] Barzashka, 2023
[630] NATO defines Wargaming as "a simulation of a military operation, by whatever means, using specific rules, data, methods and procedures", Ministry of Defence (UK), 2017
[631] UK defines Wargaming as "A scenario-based warfare model in which the outcome and sequence of events affect, and are affected by, the decisions made by the players", Ministry of Defence (UK), 2017
[632] Barzashka, 2023
[633] Barzashka, 2023
[634] Ministry of Defence, 2017, 44
[635] A 'light beam' can be a metaphor for any type of remote killing
[636] Riggs, 2014

9 Political, Economic and Regulatory Issues and Responses to AI

Ken Rogoff[637] said that Davos delegates had become too evangelical about AI. The technology should be subjected to neanderthal regulation to prevent dangers in such areas as warfare.[638] He called on governments to adopt "blunt" rules that were akin to the sweeping post-financial crisis regulations for the banking system after 2008.[639] Conversely, Jeremy Hunt[640] said governments should retain "light touch" regulation …you can't kill the golden goose before it has a chance to grow.[641]

Politicians need to understand technology and AI developments, because they need to know the direction that society is going in. "As A.I. Booms, Lawmakers Struggle to Understand the Technology" and a US "Representative … wrote … that he was "freaked out" by the ability of the ChatGPT chatbot."[642] Equally it is important to realise that the big-tech companies both "intermediate and create the information ecology of our societies, they are the prime shapers of our economy, society and polity."[643] The role of governments and policy makers is to deliver the needed framework of governance."[644]

9.1 Deny, Doubt and Delay Tactics – The Political Challenges

The 2022 Frontline[645] documentary "The Power of Big Oil" showed the Deny, Doubt and Delay tactics used by oil companies in relation to climate change. The industry's research was designed to deny and sow the seeds of doubt about climate science; stall climate policy in the face of more certain evidence; and their efforts to delay the transition to renewable energy sources, by promoting natural gas as a cleaner alternative.[646, 647] This 'deny tactic' was used by the major US tobacco companies

[637] A Harvard University Professor
[638] Fletcher and Khan, 2024
[639] Fletcher and Khan, 2024
[640] UK Exchequer Chancellor
[641] Fletcher and Khan, 2024
[642] Kang and Satariano, 2023
[643] Wolf, 2023, 296
[644] Bunz and Janciute, 2018
[645] An arm of the US based Public Broadcasting Service (PBS)
[646] Frontline, 2022
[647] See Appendix 8 for further information on the "The Power of Big Oil"

giving evidence in 1994 to US Congress hearing. Seven tobacco Chief Executive Officers (CEOs), in turn, testified under oath that "nicotine is not addictive."[648, 649] These major industries influenced and shaped policy to their own advantage. "Big tech has succeeded in distracting the world from the existential risk to humanity that artificial intelligence still poses."[650]

9.2 AI and Digital Companies' Dominance

In competition law market dominance is a technical construct, principally focussing on the level of the dominant firm's market share.[651, 652] A further view posits that where there are few market constraints it allows firms to behave independently from other market actors.[653, 654] There is a tendency "of the powerful to rig the economic and political systems against the rest of society."[655] Companies possess …

> enormous economic and political power, which they can and do abuse. … they can afford to be indifferent to the fate of their workers … judicial systems find it almost impossible to hold them or their executives criminally liable, … This can encourage behaviour that borders on the sociopathic.[656]

The sizes, revenues and cash reserves, provide powerful and dominant AI companies with their persuasive political influence. They easily legally challenge governments and regulators. Their power threatens the smooth running of

[648] University of California San Francisco, nd
[649] See Appendix 8 for full sworn transcript
[650] The scientist and AI campaigner Max Tegmark (of the Future of Life Institute. It led an unsuccessful call in 2023 for a six-month pause in advanced AI research). Quoted in Hern (b), 2024
[651] Rexel, 2020
[652] "Typically, a company is considered to hold a dominant position if it has a market share of more than 40%, but even a market share of 15% may be considered dominant if it is the largest player in a fragmented market." Rexel, 2020
[653] Graham, 2023
[654] Companies are usually constrained in their commercial behaviour by competitors, customers and consumers. Where such constraints are weak and ineffective, the company is considered to hold market power. When market power allows a company, over time, to behave independently from other market actors, such company is deemed to exercise a dominant market position. Dominance per se is not unlawful. Rather, it is the abuse of dominance – particularly if used to weaken competitive dynamics by excluding rivals and harming consumers – that is unlawful. Graham, 2023
[655] Wolf, 2023, 119
[656] Wolf, 2023, 51

democracies, they become unanswerable and dismissive, creating immense difficulties for responsible governments.

9.2.1 AI Companies, US Concentration

The USA has a disproportionate concentration on AI companies. Of the top 15 companies[657] worldwide, thirteen are in the USA, with six based in California.[658, 659] As with oil and tobacco companies, the US based AI companies have the collective ability to strongly influence US policy, which can seep through to other countries. If AI companies retain their desired legal and regulatory freedoms in the US, this will greatly strengthen their strategies and lobbying across the globe.

9.2.2 AI Companies Global Financial Power and Strength

AI companies' reach and financial power are global, fortifying their ability to exert pressure on governments. Apple's own market capitalisation is greater than Italy's GDP, [660] the eight largest world economy. The combined market capitalisation of the top seven AI companies[661] is 64% of the entire EU's 2022 GDP.[662, 663]

9.2.3 The Pacing Problem – The Legal and Regulatory Struggle

Policymakers, regulators and governments are continually behind technological developments and innovation. Laws and regulations become inadequate in the face of advances, unable to reflect the new reality, this is "the Pacing Problem":[664]

> …it has profound ramifications for the governance of emerging technologies… is becoming the great equaliser in debates over technological

[657] By Market Capitalisation (the value of a company's shares)
[658] Stash Banking Services, 2024
[659] See Appendix 8 for Locations of Top 15 AI Companies
[660] GDP is a country's Gross Domestic Product (the value of all finished goods and services made within a country during a specific period). It is a measure of the size of a country's economy
[661] In the top-ten companies worldwide
[662] The combined market capitalisation of the top seven AI companies is greater than the combined GDPs of four countries in the top ten world economies and the three largest economies in the EU
[663] See Appendix 8 for Top Ten Global Companies vs. Top Ten Countries (Market Capitalisation Values vs. GDPs)
[664] Thierer, 2018

> governance because it forces governments to rethink their approach to the regulation…[665]

In the case of AI, the pacing problem is driven by:

- Moore's Law[666] describing the rapid growth in computing power,
- Societal attitudes toward technology… shift rapidly…[667]
- The US government's progressive loss of the ability to adapt.[668, 669]

Hard law[670] suffers from many weaknesses and limitations, especially for a rapidly evolving diverse technology like AI.[671]

> In contrast, soft law refers to measures that are not directly legally enforceable but that can, nonetheless, sometimes create substantive obligations. Examples include guidelines, sets of principles, codes of conduct, private standards, and partnership programs. In many industries, hard law and soft law exist alongside each other.[672]

The "Law of Disruption"[673] also causes technology and innovation to far outpace laws and regulations.

Two initiatives, among others, were taken during 2023 attempting to slow AI's progress and understand its effects. Elon Musk and numerous leading figures issued an Open Letter[674] calling "…for all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4",[675] "Advanced AI could represent a profound change in the history of life on Earth, and should be planned

[665] Thierer, 2018
[666] See classical computing above
[667] As citizens quickly assimilate new tools into their daily lives and they then expect that even more and better tools will be delivered tomorrow. Thierer, 2018
[668] Thierer, 2018
[669] This is true of many governments as they struggle to grapple with the complexities of modern technology, societies' demand as well as societies' worries and concerns
[670] Regulation and Legislation, which take time to consider and implement
[671] Marchant, 2021
[672] Marchant, 2021
[673] See Disruptive Technologies above
[674] Bengio, Musk, et al, 2023
[675] GPT-4, GPT version 4. GPT is a Generative Pre-trained Transformers, a type of deep learning model used to generate human-like text. Common uses include answering questions, translating text to other languages and so on. Awan, nd

for and managed with commensurate care and resources"[676] In light of earlier discussions on Big-Tech, it is easy to be cynical and view it as a Machiavellian[677] tactic to hide the real motivations of the Open Letter.

The second initiative, came from the UK Government's first "AI Safety Summit" in 2023.[678] The Summit's Declaration said "We recognise that this is therefore a unique moment to act and affirm the need for the safe development of AI…"[679, 680] Again this might be viewed with a raised eyebrow in view of the UK's Jeremy Hunt's comments earlier.

9.2.4 Accountability, Regulation and Transparency

Underpinning any regulatory or governmental responses are the essential Accountability, Regulation and Transparency tools.[681] Their effectiveness relies heavily on the determination and the political will to bring them into being and to follow with strict enforcement.

Earlier sections in this chapter demonstrated the power, dominance and attitude of the AI and Digital companies. The experience of the oil and tobacco industries show just how forceful AI companies are. Their power extends across national borders and are financially stronger than many governments. Internationally based agreements are required to effectively set rules for AI companies, along with safety and ethical measures. The good and benefits of AI must be allowed to grow and develop, while at the same time agree the regulations and governance measures.

There are models for international agreements. Climate change is an ominous threat facing the globe, posing an existential risk to mankind. The UN's COP[682] meetings are held annually to agree on actions to be taken. Though more diverse in

[676] See Appendix 8 for full text of the Open Letter
[677] Suggesting the principles of conduct laid down by Machiavelli specifically marked by cunning, duplicity, or bad faith. Merriman Webster Dictionary, nd
[678] AI Safety Summit, 2023
[679] The Bletchley Declaration by Countries Attending the "AI Safety Summit", 01-02 November 2023
[680] See Appendix 8 for the full text of The Bletchley Declaration
[681] These tools are common across all the regulatory regimes in different industries. See the Glossary for an explanation of each term
[682] The United Nations established the U.N. Framework Convention on Climate Change (UNFCCC). Countries who signed up the Convention, are party to the Convention, meet annually in a Conference of the Parties (COP) in order to internationally agree climate change actions

their nature, bodies regulating accountancy practices are multinational, although not entirely global, in their reach.[683]

9.2.5 The EU's, UK's and US's Responses

The EU's GDPR[684] regulations,[685] extensively updated existing Data Protection laws.[686] The EU's Artificial Intelligence Act,[687] sets out definitions and rules relating to AI. [688, 689] However, Amnesty International were highly critical of the Act, stating that: [690, 691]

> …the legislation fails to take basic human rights principles on board… the AI Act offers only limited protections to impacted and marginalised people. It does not ban the reckless use and export of draconian AI technologies… failing to ensure equal protection to migrants, refugees and asylum seekers. …[it] lacks proper accountability and transparency provisions, which will likely exacerbate human rights abuses.

In 2024 the USA and UK signed an "AI Safety" agreement. The agreement states that they will "…work together on testing advanced artificial intelligence (AI)… [and] on developing "robust" methods for evaluating the safety of AI tools and the systems that underpin them."[692, 693]

---

[683] There are six international accountancy bodies operating in Ireland who set rules, govern the industry through their memberships and award professional qualifications and certifications. They maintain their rules up to date to reflect the changing international business environments and have the power to 'strike off' individual accountants and accountancy companies. They are ACCA (Association of Chartered Certified Accountants), AIA (Association of International Accountants), CAI (Chartered Accountants Ireland), CIMA (Chartered Institute of Management Accountants), CIPFA (Chartered Institute of Public Finance and Accountancy) and CPA (Institute of Certified Public Accountants).

[684] General Data Protection Regulation, came into force across the EU states in May 2018

[685] European Commission (c), 2016

[686] See Privacy section

[687] European Parliament (a), adopted on 13 March 2024. The EU also adopted the "Digital Markets Act" which become applicable on 2 May 2023 and the "Digital Services Act" became applicable on 17 February 2024

[688] See Appendix 8 for further information on the EU's Artificial Intelligence Act

[689] Fines of up to €35m, or 7% of turnover can be imposed on AI companies for breaches of the Act

[690] Amnesty International, 13 March 2024

[691] See Appendix 8 for full text of the Amnesty International's statement

[692] McMahon and Kleinman, 2024

[693] It is interesting that the event was attended by "AI bosses including OpenAI's Sam Altman, Google DeepMind's Demis Hassabis and tech[nology] billionaire Elon Musk." McMahon and Kleinman, 2024

9.2.6 Tackling AI Companies' Dominance

AI companies' dominance can be eliminated or diluted, their businesses broken into logical units with new owners, anti-competitive practices outlawed and eradicated. The creation of the "Baby-Bells" in 1982, following the break-up of AT&T[694] is a significant precedent in the USA.[695] In 2009, following an Anti-Trust case, the EU forced Microsoft to unbundle[696] its web browser from its Windows PC operating system. This allowed Windows users to choose an alternative web browser.[697]

9.2.7 Politically Motivated Measures

Political rather than competition motivations can lead government to bans. In 2020 India banned the Chinese owned Tiktok,[698] In 2024 the US signed into law legislation that could lead to Tiktok being banned.[699, 700]

Governments face a real challenge against the AI Companies' dominance. The companies use of the "Deny, Doubt and Delay" tactic is supported by their financial clout and their geographic concentration at the heart of the US technology industry. Globally, of the top 15 companies[701], thirteen are in the USA, with six based in California.[702, 703] In a stark commentary on the US attitude to the AI industry, The Washington Post reported:

> The lack of government controls on AI has largely left an industry built on profit to self-police the risks and moral implications of a technology capable of next-level disinformation, ruining reputations and careers, even taking human life.[704]

[694] American Telephone and Telegraph Company
[695] See Appendix 8 for further detail
[696] Bundle refers to the sale of a number of items as a single package, each item cannot be bought separately
[697] European Commission (d), 2009
[698] Toktok is a popular social media platform, with worldwide users. It is owned by a Chinese internet company called ByteDance
[699] Milmo, 2024
[700] See Appendix 12.8 for further detail
[701] By Market Capitalisation (the value of a company's shares)
[702] Stash Banking Services, 2024
[703] See Appendix 8 for Locations of Top 15 AI Companies
[704] Faiola and Zakrzewski, November 2023

Western governments, particularly the EU States, strive to strengthen their regulatory regimes. Nonetheless, they are somewhat at the mercy of the US regime where the AI companies set their own policies. The Washington Post's report exposes the dilemma for governments and their citizens.

Even with the US concentration, the issues arising from the AI companies are global. There are remedies to tackle the AI companies' dominance, it is up to governments to show courage, face down their tactics and take action.

9.3 Intellectual Property and Patents

The World Intellectual Property Organisation's (WIPO)[705] research indicates a notable upsurge in scientific papers and patent applications, particularly from 2013. This could indicate a switch from theoretical research to the practical application of AI technologies in commercial products and services.[706, 707]

Are non-human "minds" entitled to the same protection as human minds? In the US, the Court of Appeals for the Federal Circuit found that "AI did not qualify as a human." This meant AI could not be named as the inventor for patent purposes.[708] The preamble to the EU's AI Act,[709] ensures "a high level of protection of health, safety, fundamental rights as enshrined in the Charter of fundamental rights … including intellectual property rights."[710] AI can produce large volumes of text using, including derivative works,[711] which leads to ownership questions. In the US there are numerous cases before the courts but still undecided.[712, 713] For the immediate future at least, new IPs and patents remain as the creation of humans.

[705] An agency of the United Nations
[706] Szczepański, 2019
[707] In the world of commerce valuable Intellectual Property (IP) is vigorously safeguarded by its creators. WIPO describes IP as "… creations of the mind, such as inventions; literary and artistic works; designs; and symbols, names and images used in commerce." Source The World Intellectual Property Organisation, 2024
[708] In the case of Thaler v. Vidal of August 2022. The court's ruling stated, "Here, there is no ambiguity: the Patent Act requires that inventors must be natural persons; that is, human beings." Source Caldwell, 2023
[709] Sections (1) and (48)
[710] European Parliament (a)
[711] Derived from one or more already existing works
[712] Gil, *et al*, 2023
[713] See also Appendix 12.8

9.4 AI Safety by Design

Obligations to guarantee safety features in AI developments and obtain prior approval before launch[714] would be feasible for policy makers and regulators to introduce. Adapting safety principles, AI systems could be designed for a single use only, thus AI used to scan images for cancer traits could not design a bomb. More versatile systems could be rigorously tested to ensure that they cannot be adapted for malign or dangerous purposes.[715] AI could be designed with several layers of increasing intelligence and capabilities, each of which would be enabled or disabled under direct human control.[716]

9.4.1 Lessons from the Automobile Industry

To protect its reputation from bad driving, the US Automotive Safety Foundation (ASF)[717] promoted safe driving and lobbied for driving tests. Coupled with the Swedish "Vision Zero"[718] it led to a "safe-systems approach" and (an expected) "safety rating system …"[719] This concept could be transferred to the AI industry, with ratings for children, general adult use and other specialists uses (e.g. military), accompanied with strict access controls. "… the only way you can get safety first is if the government puts in place safety standards for everybody."[720]

9.4.2 Content Moderation

Relating to Free Speech discussed earlier, three questions arise regarding the platforms:[721]

[714] These safety concepts already exist for new buildings and automobiles. Drugs for medical use must go through rigorous tests and approval processes before launch

[715] Whipple and Blakely, 2023

[716] Returning to the Space Odyssey film, the computer HAL was switched off by Dave, so that as the last human he took back control of the spaceship. See "Morals and Ethics of AI" section

[717] The US Automotive Safety Foundation (ASF) was established in the 1930's. It instigated communications programmes to educate on the importance of safe driving and lobbying for driving tests. Kremer, 2024

[718] A zero road deaths policy backed by the Government which aims to eliminate roads deaths with measures that include safety in road and highways designs; Kremer, 2024

[719] Kremer, 2024

[720] The scientist and AI campaigner Max Tegmark (of the Future of Life Institute. It led an unsuccessful call in 2023 for a six-month pause in advanced AI research). Quoted in Hern (b), 2024

[721] Howard, 2024

1) Do internet platforms have moral duties to respect the free speech of their users?
2) Do they have moral duties to restrict (or refrain from amplifying) harmful speech posted by their users? and
3) If they have moral duties, should those duties be legally enforced?

Content moderation[722] opens legal and philosophical problems for liberal politicians. On the one hand they desire open free speech and expression, while at the same time supporting the removal of 'harmful' content.[723, 724] The EU's Digital Services Act, regulates "online intermediaries and platforms…"[725] Under the UK's 'Online Safety Act 2023',[726] "Platforms will have to prevent younger users from seeing age-inappropriate content…"[727, 728]

The presence of regulations merely limits harms in the short to medium term. AI companies will find new methods and products to circumnavigate the regulations.

9.5 AI's Predicted Economic Impacts

As with many technological developments, AI promises to bring greater wealth and prosperity. However, this must be viewed against the real and potential perils for individuals, economies and inequality. This presents challenging conundrums for governments to grow their economies, protect jobs, grow skills, stave off greater global inequality and aid the Global South.

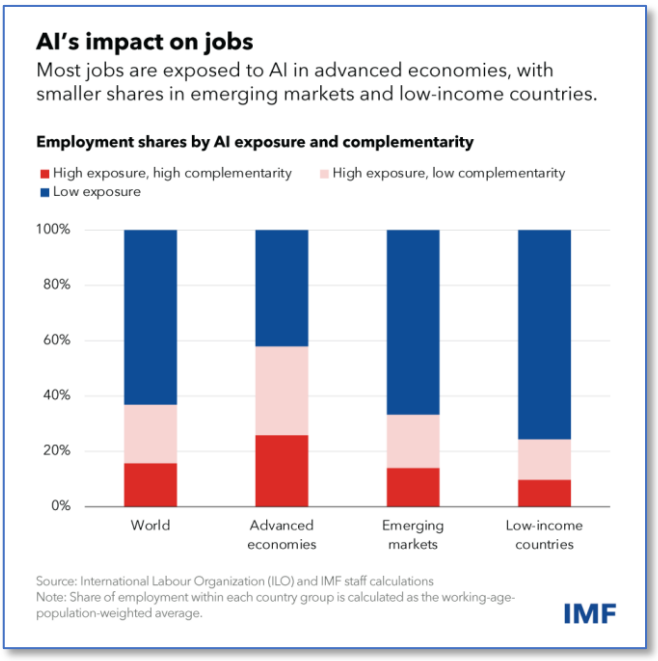
9.5.1 Labour and Jobs

The IMF[729] reported AI's impact on employment across economies, illustrated in the **Figure V** graphic.[730] While advanced economies will be most impacted with more

[722] See Freedom of Thought earlier
[723] Including implementing age-based restrictions
[724] Howard, 2024
[725] European Commission (a), nd
[726] UK National Archives, 2023
[727] There are also fines of up to the greater of £18 million or 10 percent of annual revenue that can be imposed for breaches of the Act
[728] Guest, 2023
[729] International Monetary Fund
[730] In advanced economies, about 60 percent of jobs may be impacted by AI. Roughly half the exposed jobs may benefit from AI integration, enhancing productivity" (red shading). For the other half AI may replace these "which could lower labour demand, leading to lower wages and reduced hiring" (pink shaded), Georgieva, 2024

job losses, in overall terms though AI will be beneficial. Low-income countries will be least affected but will reap far fewer AI benefits. These confirm the 'Global South Disadvantage'.

Figure V – AI Impact on Jobs[731]



> Labour income inequality may increase if the complementarity between AI and high-income workers is strong, and capital returns will increase wealth inequality. However, if productivity gains are sufficiently large, income levels could surge for most workers… Emerging market and developing economies should prioritise the development of digital infrastructure and digital skills.[732]

Research indicates that up to 54 % of EU jobs face the probability or risk of computerisation[733] within 20 years. Patterns from industrial revolutions indicates that job destruction will be stronger in the short and medium terms, with longer term job creation. Some economists say that it will be hardest for AI to replace the 'sensor-motor skills' required in non-standard and non-routine jobs.[734] Though

[731] Georgieva, 2024
[732] Cazzaniga *et al,* 2024
[733] The introduction of computers and AI systems to manage previously controlled systems as well as the storage and retrieval of data
[734] Such as those of security staff, cleaners, gardeners and chefs, which are low skilled and low paid

there is an ambiguous impact on inequality, low-skill automation always increases wage inequality and high-skill automation always reduces it.[735]

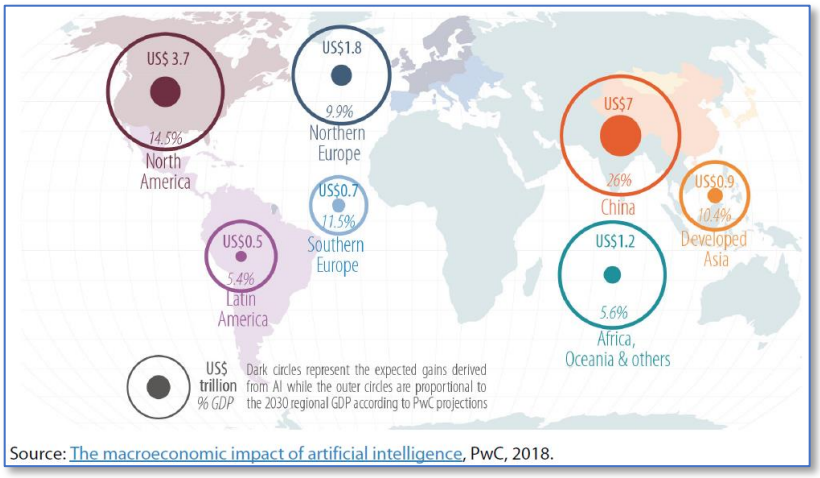
Labour mobility and frequent job changes, with a rise in precarious work, self-employment, and contract work, weaken workers' rights and the role of trade unions.[736] Structural unemployment[737] could become embedded in economies. Governments have responsibilities to retrain and reskill workers to cope with significant workforce shifts across economies, with changes in the nature and content of jobs.[738] Governments should make plans in anticipation of AI related job losses, to reduce the prospects of hardship.

9.5.2 Productivity

The economic impact of AI is anticipated to be huge. Globally it is expected that 70% of companies would adopt at least one type of AI technology by 2030.[739] Research quoted by the EU shows the strongest gains in the US, Europe and China.[740, 741]

[735] Szczepański, 2019
[736] Szczepański, 2019
[737] Structural unemployment is a longer-lasting form of unemployment caused by fundamental shifts in an economy and exacerbated by extraneous factors such as technology, competition, and government policy. Structural unemployment occurs because workers lack the requisite job skills or live too far from regions where jobs are available and cannot move closer. Jobs are available, but there is a serious mismatch between what companies need and what workers can offer. Kenton, 2023.
The closing of coal mines and steel works are examples of those skilled in the industries becoming structurally unemployed
[738] Szczepański, 2019
[739] McKinsey, in Szczepański, 2019
[740] PWC, 2018 in Szczepański, 2019
[741] Notably absent from the research are Russia, Africa, parts of Asia, New Zealand and Australia

Figure VI - Expected gains from AI in different world regions by 2030[742]



Concerningly, while technological progress has been accelerating for quite some time, it has not reversed long-term slowdown in productivity growth. This is Robert Solow's[743] "Productivity Riddle",[744] "you can see the computer age everywhere but not in the productivity statistics." To address this the IMF suggests global structural reforms are needed.[745]

9.5.3 Supply of Electricity

The provision of electricity is a challenge for governments and the fulfilment of the UN's SDGs.[746, 747] On the one hand they are very anxious to promote AI and bring as much of AI firms' infrastructure to their countries. On the other hand, they face capacity growth problems in their electricity networks,[748] local objections to network expansion plans and meeting green energy climate action targets. While

[742] PWC, 2018 in Szczepański, 2019
[743] Economist and Nobel Prize winner, in Szczepański, 2018
[744] Szczepański, 2018
[745] Szczepański, 2018
[746] Vinuesa *et al*. 2020, in Stahl, *et al*, 2023, 103
[747] Estimates of electricity needs suggest that "up to 20% of the global electricity demand by 2030" might be taken up by AI and other ICTs (Information and Communications Technology), a much higher figure than today's 1%. Vinuesa *et al*. 2020, in Stahl, *et al*, 2023, 103
[748] Examples (i) In the case of Ireland: "This year's Generation Capacity Statement predicts a challenging outlook for Ireland with capacity deficits identified during the 10 years to 2031", Eirgrid, 2022.
(ii) Holland: "Electricity grid capacity shortage now threatens households", Dutch News, 2024
(iii) Sweden: "Lack of capacity in the electricity networks has led to restrictions of new connections of consumption facilities", Palm, 2021
(iv) UK: "… previous media coverage related to developers applying for new connections to their developments only to be advised that they will face significant charges, and timescales at times at times exceeding 10 years", Premier Energy, 2023

this is challenging for developed countries, the 'Global South' remain disadvantaged.

9.5.4 Manufacturing

Exacerbating worries about labour, jobs and productivity impacts, are the anticipated changes in manufacturing. While automation and robots are well embedded in manufacturing,[749] AI is now one of the cornerstones of the growing digitalisation of industry.[750] The European Parliament comments that bringing all facets of digital technology, cloud computing, big data and so on will likely transform manufacturing into a single cyber-physical system in which digital technology, internet and production are merged into one.[751] The effect will be to continue and speed up the trend of skilled labour-intensive jobs being replaced by technology, with the consequent effects of structural unemployment and societal problems.

9.5.5 Taxation on Robots and AI Systems?

If as is expected robots and AI systems replace human jobs, a question arises should they be taxed to preserve the funds available for public services? "If automation leads to significant falls in income tax receipts and increases the pressure on government finances, this tax may be unavoidable in the future."[752] Bill Gates[753] proposed this type of taxation but it was rejected by the EU.[754]

South Korea reduced the tax deduction on investments in automation, appearing to acknowledge some experts' concerns about excessive incentivising of automation.[755] A possibility is an international solution to allow taxation to be consistent globally, this is hotly debated both internationally and at EU level.[756] Unless there is a strong level of global agreement, AI industries will move operations to lower tax regimes.

[749] See "Disruptive Technology" section
[750] Szczepański, 2019
[751] Szczepański, 2019
[752] Szczepański, 2019
[753] Founder of Microsoft
[754] Szczepański, 2019
[755] Szczepański, 2019
[756] OECD in Szczepański, 2019

9.6 AI – Future Political and Regulatory Response

This thesis has reflected on the risks and dangers AI poses both now and in the future. Tackling these issues pushes governments' abilities to the fore, citizens rely on them for their protection and to ensure that AI meets societies' needs. There are well trusted and standard regulatory and political options available.

9.6.1 Digital Ethics

The pharmaceutical and healthcare industries are rigidly regulated and held to the highest ethical standards. An overall ethically rigid approach is not currently being applied to the AI industry. This absence is fraught with dangers for humans. AI touches the lives of almost every person on the planet, in the same manner as pharmaceutical and healthcare. Governments need to reflect and take urgent action.

Digital ethics is not just accountability, regulation and transparency; it includes safety, protections for children and the vulnerable; applying ethical standard to companies, their owners and their executives; being non-discriminatory with customers, highlight conflicts of interest; declaring lobbying activity, declaring payments and incentives to politicians and officials. These are just some factors that contribute to an ethically trustworthy regime.

Regulatory regimes, back by transparency directions and regulations, ensure that the correct and up to date rules are in place to govern AI. The regulations are overseen by a statutory and independent regulatory authority, with strong legal powers, dedicated solely to the AI industry. Supporting regulations is transparency. Essential information, such as the corporate structures, ethical guidance, the operations of algorithms, safety features, control of technical layers, manipulative techniques, and so on should be made available, if not to the public at large then to the regulatory authorities at a minimum.[757] Accountability is the third leg, this requires that the AI companies, their owner and their executives must be answerable to regulatory authorities, governments and the public at large. Here the

[757] Generally speaking, regulatory authorities are obliged to respect commercial sensitivities. Therefore, information made available to the authorities may not become public

regulatory regimes need to be supported by willingness to 'name and shame'[758] and the courts willing to impose large financial penalties.

Non-financial penalties[759] should also be available. Long jail sentences and large fines for executives are an essential requirement to support the overall regime. No company should be 'too big to fail'[760] and no individual should be 'too big to jail'.[761, 762]

Policy is not just an implementation of ethical theory, numerous challenges await governments worldwide. Understanding AI's power and potential is difficult. Equally formulating policies, legislation and regulations is formidable. The strongest governments may be meek in the face of AI companies' sheer power and might. Global cooperation and agreements across governments are essential to meet these challenges. Legally binding regulations would restrict existing AI business models and practices, but Big-Tech will push against anything that limits their operations and ambitions. There is thus a significant risk that regulation will remain toothless in the face of economic and political power.[763]

For people working in ethics and policy, there might be a tendency to overestimate the impact and threats from a new technology, and underestimate how far current regulation can reach. On the other hand, there is a tendency for businesses, the

[758] The 'Name and Shame' tactic enables governments, auditors and regulators to publicly expose wrongdoings. For many organisations being named and shamed causes reputational damaged which can be difficult and take time to rebuild

[759] These could include the break-up of AI companies, restrictions on the type of products sold, prior approval of product before launch and enforcing transparent age restrictions on products

[760] "Too big to fail" refers to an entity so important to a financial system that a government would not allow it to go bankrupt due to the seriousness of the economic repercussions.

[761] "Too big to jail" is used to describe the theory that certain financial institutions and individuals, even if they engage in criminal conduct, are immune from criminal prosecution.

[762] When a company faces criminal charges, a completely different side of the criminal justice system is revealed, the world of corporate crime is completely different. Prosecutors are up against the largest and most powerful corporations in the world. In recent years, prosecutors have taken on the likes of AIG, Bristol-Myers Squibb, BP, GlaxoSmithKline, Google, Halliburton, HealthSouth, HSBC, JPMorgan, KPMG, Merrill Lynch, Monsanto, and Pfizer, and such companies can mobilize astonishing resources in their defence. Harvard University Press, 2014

[763] Müller, 2023

military, and some public administrations to “just talk” and do some “ethics washing”[764] to preserve a good public image and continue as before.[765]

This creates uncertainty as to how effective political responses will be. Will they protect humans from harm and from AI dominance? “After all, nobody elected Mark Zukerberg and Sundar Pichai[766] to control the information ecosystems of our societies. If a competitive economy and functioning democracy are to survive, some big issues need to be addressed by policy makers.”[767]

Consistent enforcement of policies across governments, worldwide and regionally, is essential. “In this evolving landscape, advanced economies and more developed emerging market economies need to focus on upgrading regulatory frameworks and supporting labour reallocation while safeguarding those adversely affected.”[768]

[764] Exaggerating the extent to which an individual, organisation or government practice truly ethical behaviour.
[765] Müller, 2023
[766] Of Facebook and Alphabet respectively
[767] Wolf, 2023, 296
[768] Cazzaniga et al, 2024

10 Findings and Reflections

These findings and reflections provide themes, aiding the understanding of AI, relationships between the human mind and AI; its dangers; and its current and near future status.

10.1 AI Companies' Power and AI Users are Threats

The central finding is that the present and future of AI's power lies chiefly in two areas. Firstly, the AI companies' collective financial size gives them strength and power when lobbying governments and policy makers. Their interests are not necessarily those of wider society. In many respects this is no different to the oil and tobacco industries, who used their size and power to influence governments.

Secondly, the malign and dangerous uses which different actors employ with AI, threatening the safety and smooth running of societies. Any tool, system or weapon in the wrong hands becomes dangerous. So too with the extremely powerful and manipulative AI. For malign actors their influence may be local, national, global, secret and without personal consequences.

Technology underlies AI systems. Strong governments are needed to control and limit the human aspects of AI's power, in terms of the AI companies and AI users.

10.2 The Human Mind is More Complex and More Diverse than AI's

The human mind is immensely complex and diverse. All humans, possess intelligence, thoughts, consciousness, self-awareness, emotions, ethics, morals, virtues, goodness, free will and imagination. Humans also hold bias in their thoughts and actions. AI creators seek to replicate the human mind's characteristics. Apart from intelligence, AI possesses none of these attributes, but some AI systems possess human created bias.

Therefore, any bias or lack of moral purpose in AI comes not from AI itself, but from the creators who build its software and algorithms.

10.3 The Artificial 'Mind' is at most 'Intelligent'

AI is intelligent, but otherwise does not compare to the human mind. Using information provided, it simply carries out algorithmic instructions. Even with the power of Supercomputing and Quantum Computing, AI is at most 'intelligent'.

10.4 AI does not Possess 'Moral Status'

AI has some human like capabilities, as proven by the Turing Test and the Chinese Room experiment. Intrinsically AI is not human like in any other manner and cannot be compared as such. It is neither moral nor immoral; neither good nor evil; of itself it is inherently neutral. It does not possess moral status. Therefore, 'switching off' AI is morally neutral and cannot be compared to harming or killing an animal or human.

AI may possess bias and questionable morals; however, these derive from the human creators and do not alter AI's lack of moral status.

10.5 Weak AI exists, Strong AI does not

AI is Weak. It acts intelligently without knowing whether it is intelligent. As it evolves, Weak AI becomes quicker, more complex and *appears more intelligent*, but remains 'Weak'.

Strong AI would be, or close to, human like; perhaps "persons" with moral status, machines with minds. If Strong AI or Superintelligence, the "Singularity", is achieved the future would be uncertain and may be beyond human control. Nonetheless, Strong AI may never truly exist.

10.6 AI is Disruptive Technologically and Socially

AI is disruptive, its benefits are immediately obvious and quickly sought by humans, speedily superseding and discarding older technologies. As AI develops, consumer expectations grow, the next generation of AI is keenly awaited. This disruption can be quickly and dramatically harmful. 'Big Data' analysis, 'Social media' and 'online' are exploited by commercial entities and by malign actors. They adversely

manipulate people's thoughts and behaviours, affecting democracy and targeting the weak and vulnerable.

10.7 AI in its Present and Future Evolutions is Harmful

10.7.1 Human Autonomy, Dignity and Freedom

In its use AI offends human's dignity. Its subliminal and manipulative features impinge upon humans' autonomy, freedom and privacy, swaying individuals to share personal information. It enables actors' malicious purposes. The 'internet of things', aids both states and bad actors to conduct close surveillance of individuals and entire populations. In healthcare AI brings threats to humans, genetic profiling may be used for malign intents and rush the euthanasia debate. Bionics may be open to harmful manipulation and humans in care settings are conceivably objectified by robots.

Many of these problems existed before AI, however AI greatly magnifies the harmful possibilities.

10.7.2 Deeper Inequality

AI further exacerbates the digital divide within societies, between the affluent and poorer sections of society, more starkly magnifying the Global South / North disparity. Even when AI is brought to the Global South, it disproportionately benefits the Global North.

The disparity between the developed and underdeveloped nations, and between the industrialised and non-industrialised countries, have existed for centuries. AI pushes these positions further apart, leaving the underdeveloped countries struggling to catch up.

10.7.3 Societal and Democratic Harms

Through AI the smooth running of society is affected. Large scale and targeted messaging by malign actors, erodes trust and worth in individuals, in sections of societies and in governments. Conspiracy theorists create false narratives of reality, for destructive and socially divisive intents. Educational standards and academic

merit become open to question, with longer-term consequences for the trust in and regard for qualifications.

AI's power provides greater force, authority and influence for those wishing to propagate their propaganda, subterfuge and conspiracies. These malevolent tactics have endured in numerous forms over the millennia, but now gain greater dispersal, force and influence with AI.

10.8 AI Increases Warfare Malevolence

AI rapidly adds to the ruthlessness, weaponry and tactics of war, bringing new strategies and capabilities. It enables destructive actors to research, develop and deploy cyber weapons, more deadly explosive armaments, chemical and biological warfare. AI is used to identify and bomb human targets, assigning 'acceptable' thresholds for the collateral killing of the innocent.

The merging of wargaming with AI brings a new level of inhumanity to warfare. Deadly bombing strikes can be remotely managed from distant control rooms. The direct human suffering and tragedy hidden from the impersonal aggressors.

10.9 Politicians and Governments are Weak in the Face of Powerful AI Companies

10.9.1 AI Companies are Powerful

Key statistics underline the economic and political power wielded by AI companies over populaces, societies and governments. With their size and financial clout, they employ lobbying tactics to undermine governments' regulatory and legislative processes. AI firms have productised dangers to humanity, any ethical approach is guided by the creators' tunnel vision and the "maximise profits and shareholder returns" ethos.

10.9.2 Political Inertia, the Pacing Problem

The Pacing Problem, the lag between the advent of AI and governing legislation, weakens governments and precipitates inertia. AI's rapid innovations coupled with the lack of understanding delays governmental action.

Governments and regulatory authorities need to employ specialists to stay abreast of AI companies and their developments, as well as to legally tackle companies and individuals. Only through this means can legislation and regulations be timely and effective to protect citizens, protect societies and curb AI companies' power.

10.9.3 Infrastructure and Strategic Planning Delays

AI companies' requirements, outpace the governments' abilities to provide infrastructure and capacity. Electricity capacity to power AI is the most common concern.

This infrastructure deficit can hamper AI companies' investments, harm other industries' growth and countries' economies.

10.9.4 Political and Governmental Solutions Are Possible

Regulations and laws are too slow coming into force. Shorter term, soft law codes of practice and guidelines, are valuable and speedy solutions. The strength and size of big-tech companies can be diluted through legally enforced break-ups of their businesses. Competition and anti-trust laws have scope for more rigorous enforcement. 'Safety by design' measures, with prior regulatory approvals, can be enforced. The mandatory inclusion of intelligence 'layers' in AI, capable of being individually disabled should be seriously considered. The Digital Ethical approach should be an urgent and minimum starting point.

Overall governments must cooperate by implementing globally consistent agreements and policies.

10.10 AI's Uncertain Economic and Employment Potential

AI's wider economic and employment outcomes are uncertain. While promising great benefits, there will be unemployment. In affluent economies, unemployed individuals should quickly find other employments. In less affluent areas, where more traditional semi-skilled manufacturing jobs are lost to AI, structural and longer-term unemployment may result. Poorer, more labour-intensive area and the Global South will not benefit as much as the more affluent and Global North areas.

AI and computerisation may not necessarily improve productivity in economies, a reminder of Robert Solow's "Productivity Riddle."

10.10.1 Taxation of Robots and AI Systems

The taxation of robots and AI systems would help restore public funds lost through unemployment.

To date the AI companies are hugely successful, generating strong revenues and customer bases across the world. International tax agreements would be necessary to prevent the AI companies moving their bases to more tax advantageous countries.

10.11 The Future – Existential Threat of Superintelligence, Strong AI and AGI

AI has raised significant and numerous legal, political and ethical questions. Dealing with and resolving these is extraordinarily difficult. If AI evolves to Strong AI, with Superintelligence and Artificial General Intelligence (AGI), the issues become greater, more complex and unknown territory for mankind. The predicaments will be existential in terms of how humans exist and live. The metaphysical understanding of the world will be fundamentally altered, bringing humans' place in the world into question. The philosophical and ethical challenges are presently unidentified. Greater exploration of these issues is well beyond the scope of this treatise.

10.12 AI Technology – A Human Creation, in Human Control

It is worth remembering when considering these Findings and Reflections that AI is a technology. Just like all technologies, AI has been designed and created by human beings, for the benefit and use of mankind.

As a technology, through its design and its uses, AI's operation will be influenced by the strengths and weakness of humans. The benevolence or maliciousness of humans ultimately decides the future relationship of humans with AI.

11 Conclusion

AI brings undoubted advancements, opening new frontiers of human knowledge, progress and benefits for mankind. The pace and diversity of developments have reached astonishing levels. As consumers and businesses avail of new advances, demand for the next phases increase. The discussion throughout however, has exposed urgent ethical difficulties relating to Weak AI.

In its current iteration AI is Weak, remaining as intelligent. Replicating the human mind is a theoretical ambition, consequently Strong AI or Superintelligence as an objective does not and may never exist, any discussion thereof is complex, theoretical and beyond the scope of this treatise.

AI undermines and changes human dignity, autonomy and freedoms. AI's advances, in areas such as healthcare, education, business and warfare, bring ethically difficult dilemmas. AI's manipulative algorithms, subliminally prey on human weaknesses at an unprecedented scale. Private information is gleaned from individuals through social media. The simplicity and speed of communications has brought ever more threats and harm to children through exposure to bullying and inappropriate content. Children and adults fall victims to financial and other damaging scams. Racism, hate and incitement to violence messages are commonplace and, in some cases, resulted in damage to property, self-harm, suicides and other deaths. Democratic institutions and the smooth running of societies are undermined through AI and social media, spreading fear, false information and conspiracy theories. Warfare has become more lethal and inhumane, elevating its existing ethical considerations. The poor, those on the wrong side of the digital divide and those in the Global South should benefit from AI, yet they do not. Human dignity and well-being of societies become casualties of AI.

Democratic processes and governments are significantly challenged in their abilities to respond to AI and protect from its harms. Laws and regulations are inadequate and delayed, against the attitude, size, legal and financial might of the AI industry. There are lessons from the history of the oil and tobacco industries' use of self-

serving tactics. Policy makers and governments have the difficult burden to stay abreast of and understand the implications of AI's developments. They face the daunting task of keeping AI in check to bring only its benefits, but to date their responses are too slow, meek and ineffective.

AI's harmful power is undeniable but need not persist. AI is a technology of human creation it is therefore for humans to control and resolve the problems. This is demanding and complex, but entirely possible. As AI develops and evolves, everyone's AI expertise becomes diluted and less complete. "As the frontier of knowledge expands, the slice that a given individual can grasp inevitably becomes a smaller fraction of the whole."[769] These tough problems are not insurmountable. Governments have options such as applying digital ethics regimes, legally binding codes of practice, the creation of independent expert advisory groups and the break-up of the large AI firms.

The two principal threats that arise with AI are not of the technology *per se*. Firstly how AI is put into use. The bias in the creation of AI, lack of strict ethical understanding and its implementation by malign actors are damaging. Secondly the power of the AI industry. Their use of Denial, Doubt and Delay tactics, collective financial force against governments and manipulative algorithms stand in the way of true digital ethics.

The advantages of AI are clear and indisputable, bringing rewards for mankind. It is worth remembering though that the disadvantages are potentially dreadful and tragic. "Freedom for the wolves often meant death to the sheep."[770]

Who controls whom, between humankind and intelligent machines? The answer is that AI can and must only become a servant for human beings, demonstrating the rationality of following ethics.[771] Nonetheless, to repeat Benanti warning "I am more worried about natural stupidity than artificial intelligence."[772]

[769] Discussing Hayek in Schmidtz and Boettke, 2021
[770] Philosopher Isaih Berlin quoted by Stiglitz, in Hosking, 2024
[771] Dong, *et al*, 2020
[772] Kington, 2024

12 Appendices

While technical terms are avoided as afar as possible, some are required, explanations and details are provided in these appendices. The appendices also provide further discussion on some topics, following the chapters' titles.

12.1 Appendix 1 – *Chapter 1 Notes*, What is Artificial Intelligence?

12.1.1 Algorithms

The computer scientist Harold Stone's definition, "An algorithm is a set of rules that precisely define a sequence of operations."[773] "An algorithm is a procedure used for solving a problem or performing a computation … an exact list of instructions that conduct specified actions step by step…"[774]

All computers and AI rely on algorithms to perform the tasks at hand including these straightforward examples:[775]

- **Following a recipe**. Recipes provide a series of steps to achieve a particular objective, such as preparing blueberry muffins … from scratch. Recipes aim to produce consistent results and help individuals … create a specific dish by following detailed instructions. In this way, recipes mirror computer science algorithms, which outline steps for generating reproducible outcomes.
- **Tying shoelaces**. … there are a finite number of steps that lead to a properly tied traditional shoelace knot...
- **Facial recognition**. Facial recognition is widely used in iPhone logins as well as Snapchat and Instagram filters. It works by projecting facial traits from a photo or video onto a biometrics map using an algorithm. The program then looks for a match between this map and a database of faces to confirm the user's identification. If facial recognition is used for Snapchat or Instagram

[773] Lum and Chowdhury, 2021
[774] Gillis, 2024
[775] Gillis, 2024

filters, there is no need for searching the database because the algorithm simply builds a map of the face and applies the filter to it.

- **Traffic signals**. Traffic signals use smart algorithms to manage traffic flow...
- **Sorting documents and papers**… sorting files alphabetically, by word count, by date, or by any other specifications...
- **Searching for a book in the library**. Finding a library book...

To complete these tasks, algorithms fall into various types:[776]

- **Search engine algorithm**. This algorithm takes search strings of keywords and operators as input, searches its associated database for relevant webpages and returns results.
- **Encryption algorithm**. This computing algorithm transforms data according to specified actions to protect it. A symmetric key algorithm, such as the Data Encryption Standard, for example, uses the same key to encrypt and decrypt data. If the algorithm is sufficiently sophisticated, no one lacking the key can decrypt the data.
- **Greedy algorithm**. This algorithm solves optimisation problems by finding the locally optimal solution, hoping it is the optimal solution at the global level. However, it does not guarantee the most optimal solution.
- **Recursive algorithm.** This algorithm calls itself repeatedly until it solves a problem. Recursive algorithms call themselves with a smaller value every time a recursive function is invoked.
- **Backtracking algorithm**. This algorithm finds a solution to a given problem in incremental approaches and solves it one piece at a time.
- **Divide-and-conquer algorithm**. This common algorithm is divided into two parts. One part divides a problem into smaller subproblems. The second part solves these problems and then combines them to produce a solution.
- **Dynamic programming algorithm**. This algorithm solves problems by dividing them into subproblems. The results are then stored to be applied to future corresponding problems.

[776] Gillis, 2024

- **Brute-force algorithm**. This algorithm iterates all possible solutions to a problem blindly, searching for one or more solutions to a function.
- **Sorting algorithm**. Sorting algorithms are used to rearrange data structures based on a comparison operator, which is used to decide a new order for data.
- **Hashing algorithm**. This algorithm takes data and converts it into a uniform message with a hashing.
- **Randomised algorithm**. This algorithm reduces running times and time-based complexities. It uses random elements as part of its logic.

12.1.2 Classical Computing

Classical computing is the traditional method of computer operations. It is simply a calculator using a sequence of bits, values of 0 and 1 to represent two states (on and off), to makes sense of and make decisions about the inputted data following a prearranged set of instructions.[777] A bit can be 0 or 1, but not both simultaneously. This is based on mathematical Binary notation, using Boolean Algebra based on the values 1 (one) and 0 (zero). These can also be interpreted as On and Off, or True and False.[778]

"… classical computing, based on transistors and electricity, has served us all well and powered the information age to date."[779]

Transistors are electronic devices that "have become the nerve cells of the Information Age."[780] They are used in electronic devices as switches, such that an electric pulse is On or Off (1 or 0). Hundreds of thousands of transistors are interconnected to create 'logic circuits'.[781]

Chips (microchips) are small, even tiny, devices in which many thousands, transistor circuits are etched into place. Modern chips carry many millions of transistors, "Apple's ARM-based M chips are state-of-the-art … contains 57 billion transistors

[777] Marr (a), 2021
[778] Britannica, 2024
[779] Reichental, 2023
[780] Riordan, 2024
[781] Krar, nd

within a die size slightly less than one square inch."[782] Small computers contain 'motherboards' which carry numerous chips and other devices to operate the computer.[783] Very large computers are housed in entire rooms and even dedicated buildings.

12.1.3 Quantum Computing

This form of computing, still in its relative infancy, originates out of quantum mechanics, the behaviour of nature at the scale of atoms and subatomic particles.[784] It employs quantum physics to solve problems that standard computers are unable to answer.[785]

They rely of 'Qubits' (quantum bits) that have multiple possible states; (i) 0 or (ii) 1 or (iii) 0 and 1 simultaneously or (iv) all intermediate states.[786] This is known as Superposition; particles can exist in multiple states at the same time which allows quantum computers to look at many different variables at the same time.[787]

According to Google, it has a quantum computer that is 100 million times faster than any classical computer in its laboratory.[788] Practical uses today can be found in pharmaceutical development, cybersecurity, financial services, and weather forecasting.[789]

> Quantum Artificial Intelligence (QAI) – and in particular, its subset research area Quantum Machine Learning (QML) – will use new quantum-designed algorithms with superpowers that result in much more powerful AI models. As QAI gradually emerges, we can expect significant improvements in the speed, efficiency, and accuracy of AI. We should also anticipate the emergence of completely new and surprising business capabilities.[790]

[782] PCMag, nd
[783] Krar, nd
[784] Reichental, 2023
[785] Maar, 2021
[786] Duggal, 2023
[787] Marr (a), 2021
[788] Marr (b), 2021
[789] Reichental, 2023
[790] Reichental, 2023

12.1.4 AI Definitions

There are several definitions of AI:

Konstantine Arkoudas and Selmer Bringsford:

> AI is the field devoted to building artifacts capable of displaying, in controlled, well understood environments, and over a sustained period of time, behaviours that we consider to be intelligent, or more generally, behaviours that we take to be at the heart of what it is to have a mind.[791]

UK Government's 2023 policy paper 'A pro-innovation approach to AI regulation':

> … 'products and services that are 'adaptable' and 'autonomous.' The 'adaptability' of AI refers to AI systems, after being trained, often developing the ability to perform new ways of finding patterns and connections in data that are not directly envisioned by their human programmers. The 'autonomy' of AI refers to some AI systems that can make decisions without the intent or ongoing control of a human.[792]

The Organisation for Economic Co-operation and Development (OECD):

> a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.[793]

The Alan Turing Institute:

> the design and study of machines that can perform tasks that would previously have required human (or other biological) brainpower to accomplish.[794]

Stanford University, Human-Centred Artificial Intelligence:

> the science and engineering of making intelligent machines.[795]

[791] K Arkoudas and S Bringsford, in Frankish and Ramsey, 2014, 34.
[792] Gajjar, 2023, 7
[793] Gajjar, 2023, 7
[794] Gajjar, 2023, 7
[795] Manning, 2022

12.2 Appendix 2 – *Chapter 2 Notes*, Disruptive Technology

Disruptive technologies were not always welcomed. Automation was first recorded in the 16th century for producing stockings, Queen Elizabeth I rejected a patent for the machine out of concern for the stocking knitters, who would end up out of work.[796] Luddism, mostly associated with the advent of printing presses, was a social movement in England, in the 18th century, was opposed to industrialisation. Luddites destroyed machines, symbols of the unemployment threat to skilled craftsmen. They used violence and intimidation against industrial production by unskilled workers who … took the jobs of skilled craftsmen.[797, 798]

[796] Fleming, 2020
[797] Craftsmen were traditionally organised into guilds
[798] Diederich, 2021

12.3 Appendix 3 – *Chapter 3 Notes*, The Human Mind

12.3.1 Philosophy and Human Intelligence

Philosophy is an ancient discipline, that seeks knowledge of all things on earth and in heaven. Knowledge of natural things and their causes led to the creation of physics and metaphysics. Understanding human affairs led to the creation of ethics, politics and the philosophy of history.[799] Learning of heavenly things led to cosmology and speculative theology. Philosophy in the ancient world was the parent of most scientific disciplines.[800] To be philosophical is to be a logical thinker, seeking knowledge of the whole, avoiding unsupported beliefs basing views on good reason and evidence. Philosophers demand of themselves and others that they have reasoned logical belief.[801]

The application of this philosophical discipline to intelligence, seeking a "reasoned logical belief", brings intelligence to its ultimate potential. Philosophy has an important role for interdisciplinary dialogue, build bridges and help other more concrete disciplines.[802] From the ancient world to the present-day philosophers led human thought, the development of education in many fields and intelligence.

12.3.1.1 Philosophers' Views of Intelligence

Intelligence is not information: it is a process, or an innate capacity to use information to respond to ever-changing requirements. It is a capacity to acquire, adapt, modify, extend and use information to solve problems. Intelligence is the ability to cope with unpredictable circumstances.[803]

> … much of our knowledge is tacit; it is not expressible in language. This was memorably expressed by Michael Polanyi when he said, "We know more than we can tell." This is true not only for innate knowledge, but also for those internalised skills which come from observing and copying others.[804]

[799] DePauw, 2024
[800] DePauw, 2024
[801] DePauw, 2024
[802] Palanca-Castan *et al*, 2021, 2-3
[803] MacFarlane, 2013
[804] MacFarlane, 2013

Establishing the need for humans to be clear and precise in how they think, Quist quotes Aristotle's Nicomachean Ethics,[805]

> … it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician demonstrative proofs.

Plato considered intuitive reason as the highest form of human intelligence and spoke of intuition as a direct contemplation of the truth.[806] He differentiated between two ways of using reason or intelligence: discursive reason or dianoia *διάνοις* and intuitive reason or nous *νοῦς* or *νόος*.

> Discursive reason is exercised to confront theses. … in mathematical proofs and logic exercises. It is a slow use of reason, which makes all the steps explicit before reaching a conclusion.
>
> Intuitive reason is a quick use of reason. It starts from the premises and reaches the conclusion without going through the whole deductive process. Intuition is the simplified form of deductive reasoning.

Aristotle discussed three types of souls:

> The vegetative soul was typical of plants. Its functions were growth and nutrition,
>
> The sensitive soul was typical of animals capable of locomotion. Along with the previous functions, it was capable of locomotion and sense perception,
>
> The intellectual soul corresponded to humans. Along with the previous functions, it had the ability to reason.

Aristotle, Plato, Roman and medieval philosophers, only assigned the rational soul to human beings. For them, intelligence and reason were almost synonymous.[807]

[805] Quist, 2022
[806] Palanca-Castan *et al*, 2021, 3
[807] Palanca-Castan *et al*, 2021, 3

12.3.2 Human Intelligence Theories

12.3.2.1 Intelligence Types

Eight 'Multiple Intelligence' types have been described,[808] which everybody holds at some level. These relate to how individuals prefer to learn and process information.

I. Verbal skills: – think in words and use language to express meaning,
II. Mathematical skills: – carry out mathematical operations,
III. Spatial skills: – think three-dimensionally,
IV. Bodily-kinesthetic skills: – manipulate objects and be physically adept,
V. Musical skills: A sensitivity to pitch, melody, rhythm and tone,
VI. Interpersonal skills: – understand and effectively interact with others,
VII. Intrapersonal skills: – understand oneself,
VIII. Naturalistic skills: – observe patterns in nature and understand natural and human-made systems.

12.3.2.2 Triarchic Theory

The 'Triarchic Theory',[809] posits that intelligence comes in three forms:

1) Analytical intelligence: – acquire and store information; to retain or retrieve information; to transfer information; to plan, make decisions, and solve problems; and to translate thoughts into performance,
2) Creative intelligence: – solve new problems quickly; learn how to solve familiar problems in an automatic way so the mind is free to handle other problems,
3) Practical intelligence: – get out of trouble; the ability to get along with other people.

12.3.2.3 Cattell, Horn, and Carroll, (CHC), Theory

This espouses that there are three strata of intelligence, which are hierarchically related to each other,[810] comprising narrow, broad and general abilities.

The most widely accepted theory is a synthesis sometimes referred to as CHC theory, named after Cattell, Horn, and Carroll. Carroll's theory itself is a synthesis of

[808] Quoting Howard Gardner, in Mercer County Community College (MCCC), 2024
[809] Developed by Robert Sternberg, MCCC, 2024
[810] Sternberg, 2012

earlier psychometric theories of intelligence. The theory is based largely upon psychometric evidence—that is, factor-analytic studies that have sought to uncover sources of individual differences in performance on standardized tests of intelligence (and related constructs). The basic Idea of CHC theory is that there are three strata of intelligence that are hierarchically related to each other.

- Stratum I includes narrow abilities,
- Stratum II, broad abilities, and
- Stratum III, general ability.

… the most important abilities are general ability (Stratum III), also referred to as g, and fluid and crystallized ability (Stratum II), also referred to as g-f and g-c. General ability is an overarching ability that is theorized to be relevant to, and involved in, a very wide variety of cognitive tasks. It has been found to be correlated with performance on a very wide range of cognitive functions and life outcomes, such as income, job performance, and even health. Fluid ability is one's ability to cope with novelty and to think rapidly and flexibly. Crystallized ability is one's general store of knowledge relevant to adaptation in one's life, including vocabulary and general information.[811]

12.3.2.4 The Psychological Views

12.3.2.4.1 Intelligence Quotient (IQ) Scoring

The 'Intelligence Quotient' (IQ) scores a range of cognitive abilities, rating an individual's intellectual capabilities and potential. They are among the most administered psychological tests, with different types of tests and are influenced by factors including educational access, nutrition, culture and medical conditions. IQ tests are also used to assist diagnosis of intellectual disabilities.[812, 813]

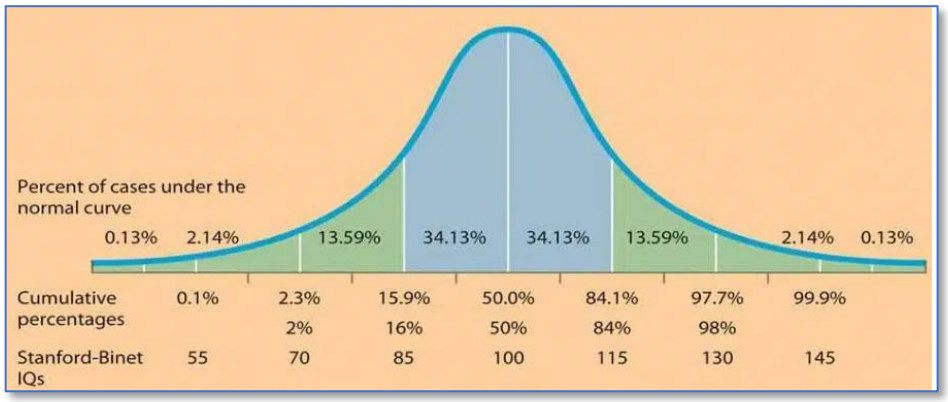
[811] Sternberg, 2012
[812] Cherry K, 2022
[813] IQ tests are also used to assist diagnosis of intellectual disabilities. Cherry K , nd

Table II – Summary of IQ Score Ranges[814]

Scores	Intelligence
1 to 24	Profound mental disability
25 to 39	Severe mental disability
40 to 54	Moderate mental disability
55 to 69	Mild mental disability
70 to 84	Borderline mental disability
85 to 114	Average intelligence
115 to 129	Above average or bright
130 to 144	Moderately gifted
145 to 159	Highly gifted
160 to 179	Exceptionally gifted
180 and above	Profoundly gifted

An average IQ score is between 85 and 115. 68% of IQ scores fall within one standard deviation of the mean. That means that the majority of people have an IQ score between 85 and 115.[815]

Figure VII – Population Distribution Bell Curve of IQ Score Ranges[816]



12.3.2.4.2 Emotional Intelligence

Emotional Intelligence is an aspect of intelligence, which was defined by Thorndike[817] as a social intelligence or having "the ability to understand and

[814] Cherry, 2023
[815] Cherry, 2023
[816] Evangelisti, 2023
[817] Kanesan and Fauzan, N, 2019

manage men and women, boys and girls to act wisely in human relations."[818] It encapsulates how humans interact with one another; chief indicators are how well individuals "understand and regulate their own mood and emotions; recognise how other people feel and empathise with them; solve problems and get their (others') needs met; and influence others."[819]

D Goleman's model is refined using a matrix:[820]

Table III – Refined framework of Goleman's mixed model of emotional intelligence

	Personal Competence	**Social Competence**
Recognition	Self-Awareness – Emotional self-awareness – Accurate self-assessment – Self-confidence	Social Awareness – Empathy – Service orientation – Organisational awareness
Regulation	Self-Management – Emotional self-control – Trustworthiness – Conscientiousness – Adaptability – Achievement drive – Initiative	Relationship Management – Developing others – Influence – Communication – Conflict management – Visionary leadership – Catalysing change – Building bonds – Teamwork and collaboration

12.3.2.5 Other Intelligence Aspects

12.3.2.5.1 Neurodivergence and Autism

There are several Neurodivergence definitions, some of which are cited here,

> **The World Economic Forum**:[821] Neurodiversity is a non-medical umbrella term that includes the conditions autism, dyslexia, dyspraxia, dyscalculia and ADHD. Teams with neurodivergent professionals can be 30% more

[818] Defined by Thorndike, in Kanesan and Fauzan, N, 2019
[819] Healthline Media LLC, 2014
[820] Adapted from Kanesan and Fauzan, N, 2019
[821] World Economic Forum, 2022

productive than those without them. But there is a greater incidence of mental health difficulties like anxiety and depression among neurodiverse individuals.

Centric Medical Health (Ireland):[822] Neurodiversity is the idea that there is no 'right' way to interact with the world. It champions the belief that differences should not be seen as deficits but rather as valuable assets that bring a richness to our society. Neurodivergent individuals often possess remarkable strengths such as increased creativity, exceptional problem-solving skills, and an unparalleled attention to detail. For example, individuals with ADHD, like Michael Phelps, Will Smith, Justin Timberlake, and Sir Richard Branson, have harnessed their unique cognitive styles to reach the pinnacle of their careers. Albert Einstein's intense focus and reclusive tendencies have sparked speculation about possible autistic tendencies, while Elon Musk openly shares his personal challenges growing up with Aspergers.

Cleveland Clinic:[823] The term "neurodivergent" describes people whose brain differences affect how their brain works. That means they have different strengths and challenges from people whose brains don't have those differences. The possible differences include medical disorders, learning disabilities and other conditions. The possible strengths include better memory, being able to mentally picture three-dimensional (3D) objects easily, the ability to solve complex mathematical calculations in their head, and many more. Neurodivergent isn't a medical term. Instead, it's a way to describe people using words other than "normal" and "abnormal." That's important because there's no single definition of "normal" for how the human brain works.

As a neurodivergence category, Autism is explored more deeply through a case study and how businesses can use autistic skills.

[822] Centric Medical Health (Ireland), 2023
[823] Cleveland Clinic, 2022

12.3.2.5.2 Case Study – Neurodivergence and Autism

Autism is of interest as some autistic individuals demonstrate above average levels of intelligence in certain areas. This is an area worth exploring due to the potential contributions which certain autistic individuals can make to the future AI world.

Kim Peek[824] was most unusual he couldn't button his own shirts, had difficulties with motor skills and was expelled from school early in life. Yet he showed extraordinary abilities at an early age, "memorising things with perfect recall before he was two and reciting books he read, line for line, with perfect recall." His IQ was just 87, yet while not having attended school "was employed at 18 job doing payroll for a company with 160 employees. It took him a few hours a week, performing all calculations in his head." He also displayed five remarkable traits of intelligence:

1) Read both pages of an open book at once
2) Provide instant driving directions between any two cities in the world
3) Figure out what day of the week anyone's birthday was, based on their date of birth
4) Reciting any Shakespeare play verbatim and
5) Counting cards.[825]

12.3.2.5.3 Intelligence Quotients – Autism Spectrum Disorder (ASD)

Autism Spectrum Disorder (ASD) is a cognitive disorder and a disability in behavioural and social development. People with autism can be smart, but autism itself doesn't make people smarter and doesn't guarantee intelligence, though many people with ASD do well in academics, in the workforce and in IQ tests. Often, autistic individuals solve complex issues while others struggle trying the same thing.[826]

More than half of autistic people in the United States have an average or above-average intelligence quotient (IQ). "An English University began studying the autism and intelligence link. It viewed over 500,000 people and found that traits of autism are typical in individuals in fields of technology, science, and other STEM (Science,

[824] The inspiration for the Raymond Babbitt character in the 1988 film "Rain Man" (Bauer, 2024)
[825] AppliedBehaviorAnalysisEDU.org, nd
[826] Zauderer, 2023

Technology, Engineering and Mathematics) professions."[827] There are correlations shared between people with a high IQ and autism:

- Interest in physics, chemistry, and mathematics
- Great attention to detail and planning
- Intimacy with people shares many physical similarities with them.[828]

It should be remembered that, while there is evidence to show "that the autistic can exhibit high IQs. However, there are also indicators of low IQ in people with ASD."[829]

12.3.2.5.4 Case Study – Autism in the Workplace

These high intelligence traits are becoming sought after by employers. The Guardian in 2022 reported that the British spy agency GCHQ[830] and weapons manufacturer BAE Systems want to attract more neurodivergent women to work for them in cybersecurity jobs, including women on the autism spectrum. The GCHQ Strategy Director said, "Having a diverse team and a mix of minds better equips us to carry out our mission and tackle new and emerging threats posed by terrorists, criminals and hostile states", while the BAE Systems' Head of Diversity said "Varying ends of the spectrum offer opportunities … ensuring that we take the broadest, most common and unique views of the world into account. Some benefits of our employees' differences are better pattern recognition, trending, creativity and innovation."[831]

The Danish not-for-profit 'Specialisterne Foundation' said "… our consultants are capable of solving tasks at a higher quality and more efficiently than non-autistic employees. … [client companies] come to us with complex and resource intensive tasks that they realise have so far been dealt with ineffectively. Our consultants are highly skilled at solving these problems in innovative and qualitatively better ways.

[827] Zauderer, 2023
[828] Zauderer, 2023
[829] Zauderer, 2023
[830] Government Communications Headquarters (GCHQ)
[831] Jolly, 2022

… the companies get solutions that are simultaneously better and require fewer resources. This is the Autism Advantage."[832]

The Harvard Business Review has noted that "Over the past two years HPE's (Hewlett Packard Enterprise) program has placed more than 30 participants in software-testing roles at Australia's Department of Human Services (DHS). Preliminary results suggest that the organisation's neurodiverse testing teams are 30% more productive than the others."[833] While a Birbeck University (UK) survey showed that "Neurodivergent employees reported remarkable abilities and work strengths, and employers concur: over 80% reported hyperfocus, 78% creativity, 75% innovative thinking, 71% detail processing and 64% people being authentic at work."[834]

12.3.2.5.5 Poverty of Stimulus

Bertrand Russell asked "[H]ow come that human beings, whose contacts with the world are brief and personal and limited, are nevertheless able to know as much as they do know?"[835] Children pick up cues about the world and language around them, building knowledge and enhancing their understanding. In the absence of sufficient information childrens' intellectual growth may be impeded. "Poverty of Stimulus" is "the assertion that natural language grammar is unlearnable given the relatively limited data available to children learning a language, and therefore that this knowledge is supplemented with some sort of innate linguistic capacity."[836]

Plato in his *Meno* observes "Socrates discovered people's innate ability to fully understand foreign concepts that they are never exposed to."[837] The term is controversial,[838] but Chomsky has "referred to "children's "biological endowment.""[839]

[832] Specialisterne Foundation, 2022
[833] Austin and Pisano, 2017
[834] Birbeck University, 2023
[835] Bertrand Russell discussing Plato in, Nordquist, 2019
[836] Noam Chomsky in Psyno, nd
[837] Psyno, nd
[838] Psyno, nd
[839] Lasnik & Lidz, 2017, in Pearl, nd

"At its heart, Poverty of Stimulus is a developmental claim, i.e., a claim about whether something is in fact learnable (and more practically, learnable by typically developing children). It occurs when there's unresolvable ambiguity in the data – in this way, the data are insufficient on their own (i.e., the stimulus is too impoverished) for children to infer the correct answer,"[840, 841] yet they do.

12.3.3 Human Personality – Consciousness, Ethics and Morals

12.3.3.1 Consciousness

Consciousness is perhaps one of the more abstract[842] and elusive of human traits, yet all individuals are aware of their own. Understanding consciousness it is crucial in the AI world.

Consciousness provides humans with the abilities to be aware of themselves, aware of others, animals and the inanimate world around them. It is a component that all humans possess to varying extents. Descartes views consciousness as thoughts which are the "contents of our minds." "...the word 'thought' for Descartes applies to conscious experience of any kind":[843]

> It is I who have sensations, or perceive corporeal objects as it were by the senses. Thus, I am now seeing light, hearing a noise, feeling heat. These objects are unreal, for I am asleep; but at least I seem to hear, to be warmed. This cannot be unreal, and this is what is properly called my sensation, precisely so regarded, is nothing but an act of thought.

[840] Pearl, nd

[841] "Poverty of Stimulus relates to children's hypothesis spaces, the data available to children, and how children leverage those available data. Importantly, Poverty of Stimulus is often tied to the fact that children seem to make constrained generalisations faster than we might expect, given the data available. This constrained generalisation then implies that children have some kind of prior knowledge that guides them towards the correct answer. The prior knowledge for a particular Poverty of Stimulus problem is generally believed to involve some kind of innate knowledge, and the big divide is whether any of that innate knowledge is language specific." Pearl, nd

[842] There are many examples of a driver or a pedestrian, on their journeys deep in thought or conversation, arriving safely at their destinations. During their journeys they are consciously unaware of junctions, traffic lights and obstacles. Yet paradoxically, their eyes, ears and brains process information so that they become aware and alert if a risk arises

[843] Kenny, 2010, 592

"These apparent sensations, possible in the absence of a body, are what later philosophers were to call 'sense data'."[844]

Descartes also identifies three classes of his ideas "Of my ideas, some seem to be innate, some acquired, and some devised by myself."[845] He authenticates thoughts, viewing the utterances of others as having 'special status'. "Experiences (of others), thus, have a certain property of indubitability, and it was this property that Descartes took as an essential feature of thought …"[846] This, to Descartes, shows the uniqueness of an individual's thoughts and consciousness.[847]

Having proved to his own satisfaction that he exists, Descartes asked "What am I, this I whom I know to exist?" His answer *Reg cogitans*[848, 849] For Descartes, thought is to be understood broadly, "I use this term to include everything that is within us in such a way that we are immediately conscious of it."[850] Copeland provides another view; "a conscious entity is one whose sensory interactions with the world customarily accompanied by qualia[851] of some sort."[852] "… in the (Descartes') *Meditations* the mind is used to validate itself."[853]

12.3.3.2 Ethics and Morals

Ethics and morals provide humans with guidance for questions like; What is good? What is right? How do we approach difficult questions? "A good purpose is essential if an action is to be morally good … The mere belief that one's purpose is good does not suffice to render an action morally correct."[854]

Morality is a set of norms and principles that govern our actions with respect to each other and the world around us, they are taken to have a special kind of weight or authority. More fundamentally, we can also think of morality as consisting of

[844] Kenny, 2010, 592
[845] Kenny, 2010, 593
[846] Kenny, 2010, 660
[847] Kenny, 2010, 660
[848] "I am a thing that thinks"
[849] Kenny, 2010, 660
[850] Kenny, 2010, 660
[851] Subjective, conscious experiences (singular 'quale')
[852] Copeland, 1993, 171
[853] Kenny, 2010, 596
[854] Discussion on Hegal in Kenny, 2010, 703

moral reasons, grounded in some basic values. Ethics is generally understood to be the study of "living well as a human being." In Aristotle's *Nicomachean Ethics*, the aim of human beings is to exemplify human excellence of character. The sense in which we understand it here is that ethics is broader than morality.[855]

Discussing 'moral motivation', Prinz argues on empirical evidence to support internalism:[856, 857]

1. Moral judgments consist of emotional attitudes.
2. Emotional attitudes are motivating.
3. Therefore, moral judgments are motivating.[858]

Quite apart from emotions, there are subjective forms of all the major moral theories and objective versions of many. An objective standard of right holds that the agent must meet the standard[859,] in order to count as right or virtuous. Subjective standards come in two broad forms:

1. *Psychology sensitive*: are the justifying reasons part of the agent's deliberative processes, or more weakly, are they "recoverable" from the agent's psychology?
2. *Evidence sensitive*: the right action isn't the one that actually meets the standard, but instead, is the action that the agent could foresee would meet that standard.[860]

12.3.3.3 The "Trolley Problem"

Driver[861] raises the famous 'Trolley Problem' thought experiments, to illustrate situations which are structurally similar, but elicit very different intuitions about what the morally right course of action would be. We intuitively believe that it is

[855] Driver, 2022
[856] Prinz 2015, in Rosati, 2022
[857] Judgment 'internalism, has been characterised as claiming either that motivation is internal to moral judgment, in the sense that moral judgment itself motivates without need of an accompanying desire (strong internalism) or that there is a necessary connection between moral judgment and motivation (weak internalism) Rosati, 2022
[858] Rosati, 2022
[859] Meeting the standard is something 'objective', not dependent on the agent's psychological states
[860] Driver, 2022
[861] Referencing Foot, 1975, in Driver 2022

worse to kill someone rather than to allow the person die. We believe it is wrong to kill one person to save five others in the following scenario:[862]

> David is a great transplant surgeon. Five of his patients need new parts … but all are of the same, relatively rare, blood-type. By chance, David learns of a healthy specimen with that very blood-type. David can take the healthy specimen's parts, killing him, and install them in his patients, saving them. Or he can refrain from taking the healthy specimen's parts, letting his patients die.

Yet, in the following scenario we intuitively view it entirely permissible, and possibly even obligatory, to kill one to save five:

> Edward is driving a trolley (tram), whose brakes have just failed. On the track ahead are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Edward can turn the trolley onto it. Unfortunately, there is one person on the right-hand track. Edward can turn the trolley, killing the one; or he can refrain from turning the trolley, killing the five.[863]

In Kantian deontology "The morally worthy action is in accordance with the Categorical Imperative, which requires an agent refrain from acting in a way that fails to respect the rational nature of other persons."[864] For Kant the Categorical Imperative is "an objective, rationally necessary and unconditional principle that we must follow despite any natural desires we may have to the contrary."[865] "Other philosophers, such as Hobbes, Locke and Aquinas, had also argued that moral requirements are based on standards of rationality."[866] This "rationality" argument would compel an objective view regarding the transplant surgeon and the 'Trolley Problem' and judge the decisions of the key agents accordingly.

[862] Noting Thomson, 1976, 206, Driver, 2022
[863] Driver, 2022
[864] In deontological ethics an action is considered morally good because of some characteristic of the action itself, not because the product of the action is good, Britannica, 2024
[865] Johnson and Cureton, 2022
[866] Johnson and Cureton, 2022

12.3.3.4 Goodness

According to Kant "… everything has either a price or a worth." A price can have something else in its place as a fair exchange, if it is beyond price and is unexchangeable then it has worth. Kant further states:

> Morality, and humanity so far as it is capable of morality, is the only thing which has worth. Skill and diligence in work have a market price; with, imagination and humour have a fancy price; but fidelity to promises and benevolence based on principle (not on instinct) have an intrinsic worth.[867]

[867] Kenny, 2010, 700-01

12.4 Appendix 4 – *Chapter 4 Notes*, The AI Mind

12.4.1 AI – Morals and Ethics

A vision of AI in the future was presented in the 1968 science fiction film "Space Odyssey 2001." "The (space)ship's computer, HAL 9000, which possesses human intellect and vocal ability, malfunctions and begins to work against the astronauts in a life-or-death battle of wits, leading to questions about humankind's relationship to machines."[868] The sentient HAL delivered some memorable and dystopian quotes displaying consciousness, self-awareness, empathy, ability to lie, disobey human requests, fear and awareness that its "mind is going":[869]

> I am putting myself to the fullest possible use, which is all I think that any conscious entity can ever hope to do;
>
> Look Dave, I can see you're really upset about this. I honestly think you ought to sit down calmly, take a stress pill, and think things over;
>
> [Regarding the supposed failure of the parabolic antenna on the ship, which HAL himself falsified], It can only be attributable to human error;
>
> [Dave speaking to HAL; I won't argue with you anymore! Open the doors!], I know that you and Frank were planning to disconnect me, and I'm afraid that's something I cannot allow to happen … Dave, this conversation can serve no purpose anymore. Goodbye;
>
> I'm afraid. I'm afraid, Dave. Dave, my mind is going. I can feel it. I can feel it. My mind is going. There is no question about it. I can feel it. I can feel it. I can feel it. I'm a... fraid.

These passages illuminate several issues about how we should approach AI's morals and ethics. Creators must be mindful that AI may develop its own faculties without human assistance. It is the creation of humans, functioning with the attributes it is given.

[868] Pfeiffer, 2024
[869] IMDb, 2024

12.4.1.1 AI Morals and Moral Status

12.4.1.1.1 AI Morals and Moral Status

Will AI or AGI of the future have moral status? Bostrom and Yudkowsky argued positively acknowledging the possibility, providing two criteria for moral status:

> *Sentience:* the capacity for phenomenal experience or qualia, such as the capacity to feel pain and suffer
>
> *Sapience:* a set of capacities associated with higher intelligence, such as self-awareness and being a reason-responsive agent.[870]

12.4.1.1.1.1 AI Consciousness and Moral Status

12.4.1.1.1.1.1 AI Moral Status

Animals possess qualia and therefore some moral status, only humans possess sapience, which gives humans higher moral status. Yet some animals can exhibit self-awareness, the 1970 'Mirror Test' is interesting. In an experiment with animals and chimpanzees.[871]

> Non-human animals, when confronted with their own reflection in a mirror, typically act as if they were interacting with another animal, i.e., by showing social or antagonistic behaviour.
>
> [When chimpanzees had a red mark placed on their faces] interaction with the marked area of their faces increased dramatically... For example, one… inspected the area with its finger and then sniffed at the finger for olfactory cues about the dye.

Later tests with other animals[872] displayed similar behaviour. These tests suggest that some animals possess *sapience* (self-awareness). If AI could feel pain, thereby possessing qualia it would possess moral status. It would be more animal like and not a rag doll! "If … an AI system also has sapience of a kind similar to that of a normal human adult, then it would have full moral status, equivalent to that of

[870] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322
[871] Ocklenburg and Güntürkün, 2018
[872] Asian elephants, bottlenose dolphins, killer whales and magpies

humans beings."[873] For instance, if AI had ability to recognise itself in a mirror. Two principles support this hypothesis:

> *Principle of Substrate Non-Discrimination:* If two beings have the same functionality and the same conscious experience, and differ only in the substrate[874] of their implementation, then they have moral status.
>
> *Principle of Ontogeny[875] Non-Discrimination*: If two beings have the same functionality and the same consciousness experience, and differ only in how they came into existence, then they have the same moral status.[876]

Backing the moral status of AI in this way could be controversial. If it is accepted that "… many questions … can be answered by … applying the same moral principles … [then] we ought to treat an artificial mind in just the same way as we ought to treat a qualitatively identical natural human mind in a similar situation."[877] Novel ethical questions remain. Will AI really be like a human, will it have sentience and emotions? Will moral status therefore be deserved?

12.4.1.1.1.1.2 AI Consciousness

AI would need to possess all five for it to be consciousness. Currently this is not the case for AI. Aleksander *et al* list five axioms or principles to be sufficient for consciousness in an AI context.[878]

Let **A** be an agent in a sensorily-accessible world **S**. For **A** to be conscious of **S** it is necessary that:

1. Depiction – **A** has *perceptual states* that depict parts of **S**
2. Imagination – **A** has internal *imaginational states* that recall parts of **S** or *fabricate* **S**-like sensations

[873] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322
[874] In relation to the current discussion, substrate refers to the bases for human (biological) and artificial (technological) intelligence
[875] The differing ontogeny is a result of human intelligence originating from biological evolution whereas AI's origin comes from its creators' deliberate design
[876] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322-23
[877] Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 324
[878] Matthias Scheutz in Frankish and Ramsey, 2014, 260

3. Attention – **A** is *capable of selecting* which parts of **S** to depict or what to imagine
4. Planning – **A** has a means of control over imaginational state sequences to *plan actions*
5. Emotion – **A** has additional *affective state* that evaluate planned actions and determine the ensuing action

12.4.1.1.2 A Robot's Behaviour

Mohammad, an autonomous robot, was launched at the DeepFest technology festival in Saudi Arabia during March 2024. The robot inappropriately touched the female journalist while she was providing her report.[879] This action caused her to move away and raise her hand signalling the robot to stop. Was it an unexpected or deliberate movement by the robot; did the robot gain satisfaction; did the programmer feel embarrassed or gain satisfaction; was the robot conscious of the consequences of its action?

The robot's developers QSS, said that the robot operates independently without human control, "We have already conducted a thorough review of the footage and the circumstances surrounding the incident and there were no deviations from the expected behaviour of Mohammad, however we will take additional measures to prevent anyone getting close to the Robot within its areas of movement."[880]

In the robot's apparent human-like action, its algorithms were responsible rather than it being a consciousness or moral decision.

Stills from the seven second video are on the page following.

[879] The Times of India, 2024
[880] The Times of India, 2024

Stills from the seven second video on Times of India[881]

12.4.2 Imagination – Human and AI

12.4.2.1 Human Imagination

12.4.2.1.1 Twelve Conceptions of Imagination

Quoting from Leslie Stevenson, Michael Beaney lists twelve conceptions of imagination, which show its diversity and complexity:[882]

1. The ability to think of something that is not presently perceived, but is, was or will be spatio-temporally real. In this sense I might imagine how my daughter looks as I speak to her on the phone, how she used to look when she was a baby, or how she will look when I give her the present I have bought her.

[881] The Times of India, 2024
[882] Beaney, 2016

2. The ability to think of whatever one acknowledges as possible in the spatio-temporal world. In this sense I might imagine how my room will look painted in a different colour.
3. The liability to think of something which the subject believes to be real, but which is not real. Stevenson talks of 'liability' rather than 'ability' here to indicate that there is some kind of failure in the cognitive process. In this sense I might imagine that there is someone out to get me, or Macbeth imagines that there is a dagger in front of him.
4. The ability to think of things one conceives of as fictional, as opposed to what one believes to be real, or conceives of as possibly real. In this sense I might imagine what the characters in a book are like, or imagine the actors in a film or play as the characters they portray, aware that the characters are only fictional.
5. The ability to entertain mental images. Here I might conjure up an image of a large, black spider or a five-sided geometrical figure.
6. The ability to think of (conceive of, or represent) anything at all. Here I might imagine anything from an object before me being transformed in some way to an evil demon systematically deceiving me.
7. The non-rational operations of the mind, that is, those kinds of mental functioning which are explicable in terms of causes rather than reasons. Here I might imagine that smoking is good for me since I associate it with the cool behaviour of those I see smoking in films. It may not be rational, but there is a causal explanation in terms of the association of ideas, upon which advertisers rely so much.
8. The ability to form beliefs, on the basis of perception, about public objects in three-dimensional space which can exist unperceived, with spatial parts and temporal duration. Here I might imagine that the whole of something exists when I can only see part of it, or that it continues to exist when I look away.
9. The sensuous component in the appreciation of works of art or objects of natural beauty without classifying them under concepts or thinking of them as practically useful. In looking at a painting or hearing a piece of music, for example, I may be stimulated into imagining all sorts of things without

conceptualising it as a representation of anything definite, or seeing it as serving any particular purpose.

10. The ability to create works of art that encourage such sensuous appreciation. In composing a piece of music, the composer too may imagine all sorts of things without conceptualising it in any definite way in the sense, say, of having a message that they want to get across.
11. The ability to appreciate things that are expressive or revelatory of the meaning of human life. In contemplating a craggy mountain range at dusk, for example, or a painting by Caspar David Friedrich depicting such a scene, I may imagine how much we are subject to the awesome power of the natural world, and yet ourselves have the conceptual and imaginative power to transcend it all in thought.
12. The ability to create works of art that express something deep about the meaning of human life, as opposed to the products of mere fantasy. Michelangelo's Sistine Chapel, Shakespeare's Hamlet, Goethe's Faust, Beethoven's late string quartets or Wagner's Ring cycle might all be offered as examples of this final conception of imagination.

12.4.2.2 Could AI become deluded?

It can be asked if a person's imagination runs wild, so to speak, could they become delusional? "Delusion, in psychology, [is] a rigid system of beliefs with which a person is preoccupied and to which the person firmly holds, despite the logical absurdity of the beliefs and a lack of supporting evidence."[883] Delusions are medically associated with mental illness and it may be incorrect to state otherwise. Less medically focussed definition says:

> "Delusions can be characterised as belief-like mental representations that manifest an unusual degree of disconnectedness from reality"[884] and one could "characterise them as beliefs that are dysfunctional in their content or formation… and … characterise them as dysfunctions of imaginings."[885]

[883] Britannica, 2024
[884] Bortolotti and Miyazono, 2015 in Liao and Tamar, 2020
[885] Liao and Tamar, 2020

There is a connection between imagination and delusion… a delusion is an imagined representation that is misidentified by the subject as a belief.[886]

Delusions and self-delusions are traits affecting humans. If AI in the future could be developed to have an active imagination, could it become deluded – develop a mode of thinking and imagination that leads to deluded notions about itself and the world?

Delusional thinking in humans can create and support conspiracies theories.[887] Consequently, while AI might not become deluded in the true sense of the word, the results of its imaginings could be flawed, triggering or supporting conspiracy theories. In truth, any AI delusion would, most likely, be the result of defective software or defects in self-learning algorithms.

12.4.3 AI – Singularity and Superintelligence

'Singularity' acknowledges the prospect that at some future point the intelligence of AI will exceed that of human intelligence. AI will then have become 'superintelligent'. This holds both promise and peril for humankind.

It is of course possible that the Singularity and Superintelligence may not transpire:

> One question is whether such a Singularity will ever occur—it may be conceptually impossible, practically impossible or may just not happen because of contingent events, including people actively preventing it. Philosophically, the interesting question is whether Singularity is just a "myth",[888] and not on the trajectory of actual AI research. … Philosophy is not on the "secure path of a science",[889] and maybe AI and robotics aren't either.[890] So, it appears that discussing the very high-impact risk of

[886] Liao and Tamar, 2020

[887] Conspiracy Theories are defined as "an attempt to explain harmful or tragic events as the result of the actions of a small powerful group. Such explanations reject the accepted narrative surrounding those events; indeed, the official version may be seen as further proof of the conspiracy." Scott, 2024

[888] Floridi, 2016 and Ganascia, 2017 in Müller, 2023

[889] Kant, 1791: B15 in Müller, 2023

[890] Müller, 2020 in Müller, 2023

> Singularity has justification even if one thinks the probability of such Singularity ever occurring is very low.[891]

Regardless, if the probability is very low, it must be seriously considered.

[891] Müller, 2023

12.5 Appendix 5 – *Chapter 6 Notes*, AI Effects on Human Dignity and Autonomy

12.5.1 Who Creates and Builds AI?

12.5.1.1 The China Syndrome

The 1979 film *The China Syndrome,* a fictional story, set in California, based on a real event. Following an earthquake, a nuclear reactor's cooling water supply is threatened. If the water runs out the result could be the "China syndrome", the superheated nuclear materials will melt directly through the floor of the plant and theoretically (according to the plot), keep on going until it reaches China. The team managing the power plant would have lost control of the reactor and the subsequent consequences.[892]

12.5.2 Threats to Personal Dignity and Autonomy

12.5.2.1 Stolen Dignity and Autonomy

Following a problem with an IT[893] system, a Los Angeles employee was dismissed for no apparent reason and escorted from the building "like a thief." The employee said that "The system was out for blood and I was its very first victim", the IT problem was identified and he was reinstated. This event described a failure of human thinking that allowed "it to be humans versus machines rather than humans plus machines."[894] "… the dignity of human beings and their 'diminishing value' [is at stake] as we approach the confluence of efficiencies gained from the increasing implementation of artificial intelligence and robotics."[895] Being escorted from the building like a thief, without any wrongdoing, is a loss of dignity. "Psychological research has shown that being wrongly accused of criminal offences can have severe consequences for the accused, including for their sense of self and their sense of dignity."[896]

In Japan, robots are used for elderly care, to "… help lift older people …; assist with mobility and exercise; monitor their physical activity and detect falls; feed them;

[892] Ebert, 1979
[893] Information Technology (computer systems)
[894] Stahl, *et al*, 2023
[895] Quoting Diallo, 2018 in Stahl, *et al*, 2023
[896] Stahl, *et al*, 2023

and help them take a bath or use the toilet." They engage with "…older people socially and emotionally … to… reduce, and even prevent cognitive decline; they might also provide companionship and therapy for lonely older people …"[897] While this appears very worthwhile, consider that in Japan, a fuzzy animatronic[898] seal is used "… to provide a robotic form of animal therapy"[899]

> An old lady sits alone in her sheltered accommodation stroking her pet robot seal. She has not had any human visitors for days. A humanoid robot enters the room, delivers a tray of food, and leaves after attempting some conversation about the weather, and encouraging her to eat it all up. The old lady sighs, and reluctantly complies with the robot's suggestions. When she finishes eating, she goes back to stroking the pet robot seal: 'At least you give my life some meaning' she says.[900]

The concept of the companion robot does seem beneficial but can "… infantilise elderly people in the eyes of carers or undermine their self-respect, if offered as a sole replacement for human interaction"[901] and poses the question whether "… elderly people who are cared for by robots are objectified …, turned from subjects into things."[902] The dignity, autonomy and self-worth of the older person is greatly diminished through loss of human contact, that is not replaced by a companion robot. Contrast this with the child who is greatly comforted by a cuddly toy but is never without human contact.

In Derby (UK), a care home introduced responsive robot pets to provide comfort to their residents some of whom said "… the robotic pets are "comforting" and "realistic."" In this situation however, the care staff continue their regular interaction with the residents.[903]

---

[897] Wright, 2023
[898] "Relating to, or being a puppet or similar figure that is animated by means of electromechanical device", Merriman Webster Dictionary, consulted 05 April 2024, https://www.merriam-webster.com/dictionary/animatronic
[899] Wright, 2023
[900] Sharkey, 2014 in Stahl, *et al*, 2023
[901] Stahl, *et al,* 2023
[902] Richardson, 2015 in Stahl, *et al,* 2023
[903] BBC East Midlands News, 2024

In circumstances when there is a shortage of human care workers and companion robots are an option, Aristotle's view "that one should always strive to find the middle between excess and deficiency"[904] provides guidance. Human contact should be maintained and old people not abandoned to robotic things.

12.5.2.2 China's Social Credit System

Under China's social credit system, the government collects large volumes of data, to honour good deeds and punish bad deeds thereby disciplining citizens. It is intended to pave the way to more "harmony through trust." Citizens hope to be better protected from food scandals and financial fraud, because the social credit system evaluates not only citizens, but also government agencies, companies and associations.[905]

The data is collected from various sources, including its own registers and private service providers. An extensive network of surveillance cameras also collects data, with automatic facial recognition used to identify traffic offenders and publicly pillory them. Worryingly what was once considered acceptable could be declared undesirable by the government in retrospect and punished.[906]

The procedure is not transparent, violating one of the central requirements of algorithm ethics, namely that socially relevant electronic decision-making processes must be comprehensible. Some commentators, such as US lawyer Jeremy Daum, doubt that the social credit system is effective. It does not rely on clever artificial intelligence (AI) and penalties are enforced. It is seen as a propaganda tool, to keep the population believing that they are constantly surrounded by surveillance and therefore remain law abiding. See the infographic on the following page.[907]

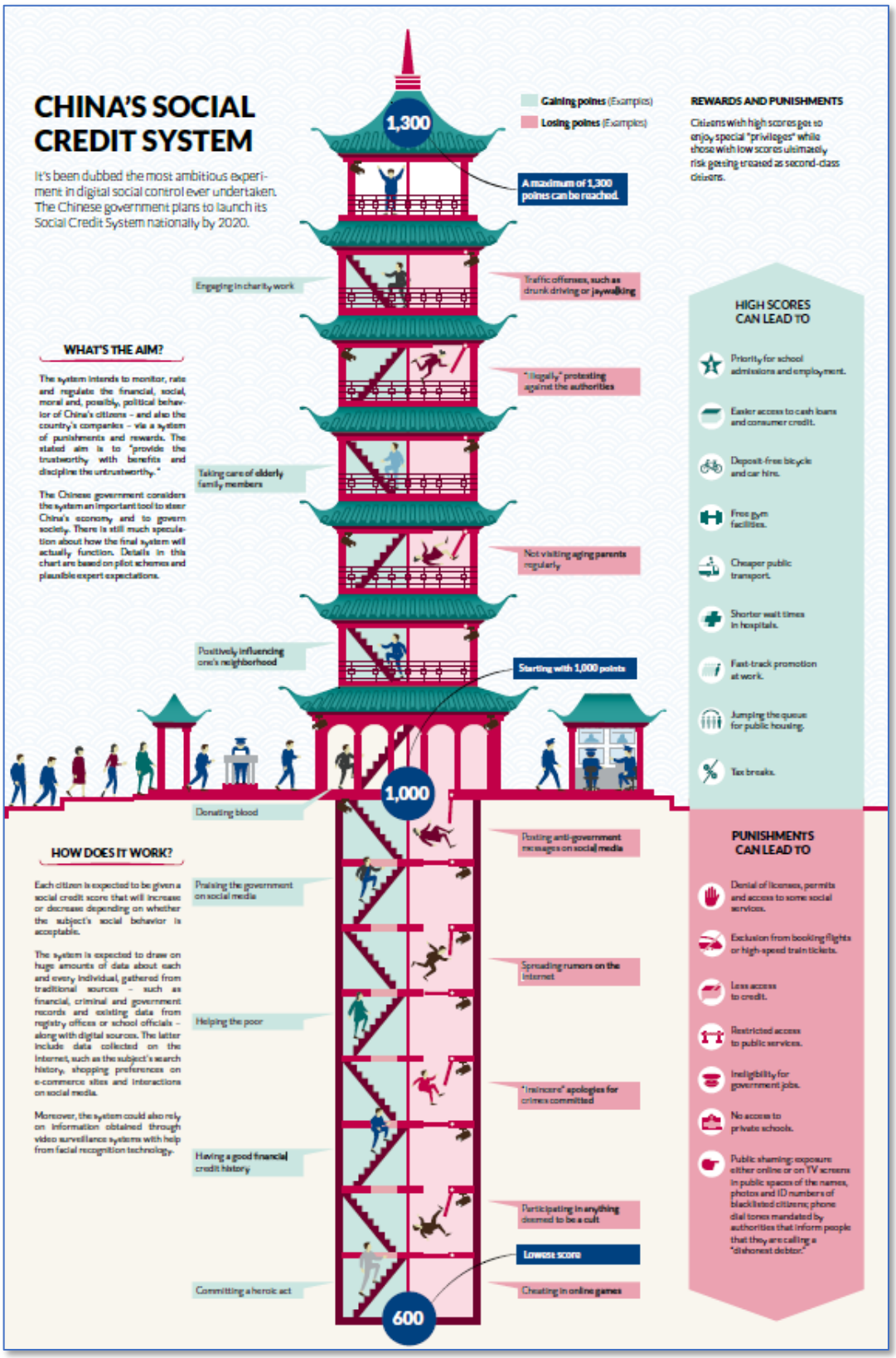
It is currently low-tech and sometimes relies on "information gathers" who walk around villages and record the deeds of others. Punishments are enforced, in 2018 the authorities banned people from purchasing flights 17.5 million times. They can also clamp down on luxury options, many are barred from business-class train

[904] Stahl, *et al,* 2023
[905] Bertelsmann Stiftung, 2022
[906] Bertelsmann Stiftung, 2022
[907] Bertelsmann Stiftung, 2022

tickets and are kept out of the best hotels. Dog owners could have their pets confiscated. Those with good scores can speed up travel applications to places like Europe and book a hotel without having to pay a cash deposit.[908]

[908] Canales and Mok, 2022

12.5.2.2.1 Infographic – China's Social Credit System[909]



[909] Text: Bernhard Bartsch, Martin Gottske; Illustration: Christian Eisenberg; from Bertelsmann Stiftung, 2022

12.5.3 Big Data and the Internet of Things

'Big Data' is a term used in relation to computers and databases. "Big" and "Data" are crucial to understanding its power.[910] There are a number of definitions available, which are important to quote:

> … data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse.[911]

> … the huge volume of digital information generated … is not only generated by traditional information exchange and software, but also from sensors of various types embedded in a variety of environments; hospitals, metro stations, markets, and virtually every electrical device that produces data. It exceeds the capacity of traditional data management technologies.[912]

> … a combination of structured, semi structured and unstructured data collected by organisations that can be mined for information and used in machine learning projects, predictive modelling and other advanced analytics applications.[913]

12.5.4 Privacy is Dead

Through Terms & Conditions and Privacy Notices many organisations give themselves permission to collect, use and share customers' personal data. While many customers do not necessarily agree, they have no choice but to provide consent, otherwise they cannot avail of the services. In a 2010 article titled "Privacy is dead on Facebook. Get over it", numerous quotes from the leaders of major technology companies of the time are given,[914] demonstrating the cavalier and condescending attitudes to their customers:

Facebook founder and CEO Mark Zuckerberg

> none of the cool kids care about privacy. Neither should you.

[910] See also the Tesco and Target retailer examples earlier
[911] Franks, 2012, 4
[912] Mediratta, 2015
[913] Botelho and Bigelow, 2022
[914] Popkin, 2010

Sun Microsystems chief executive Scott McNealy, 1999

> You have zero privacy anyway … Get over it.

Google Chief Executive Eric Schmidt, 2009

> If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place.

Mark Zuckerberg, 2010

> … in the last 5 or 6 years, blogging[915] has taken off in a huge way and all these different services that have people sharing all this information. People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that's evolved over time.[916]

Facebook's privacy guide

> Making connections – finding people you know, learning about people, searching for what people are saying about topics that interest you — is at the core of our product. This can only happen when people make their information available and choose to share more openly.[917]

12.5.4.1 Hacking – Privacy, Security and Quantum Computing

In January 2021 "Hackers with ties to the Chinese government deployed ransomware attacks against five major gaming companies. They demanded over $100m in ransom."[918]

The HSE[919] suffered a major cyber-attack in May 2021. Practically all HSE systems were paralysed with staff resorting to paper files. "The aim of the Attacker was to disrupt health services and IT systems, steal data, and demand a ransom."[920]

[915] A 'blog' is a web log. Web refers to the world wide web, the internet. Log comes from a ship's log, a record of the main events during a ship's voyage. Individuals use blogs to share information and ask questions
[916] Popkin, 2010
[917] Popkin, 2010
[918] Sheldon, Hanna, 2022
[919] Health Service Executive, Ireland's public health service
[920] PWC, 2021

In July 2021 “Iran used Facebook to target U.S. military personnel, posing as recruiters, journalists and nongovernmental organisation personnel. The hackers sent files with malware and used phishing sites to trick victims into providing sensitive credentials.”[921]

In 2022 Albania’s systems were hacked and information leaked online.[922]

The Indonesian state is thought to have accessed “the devices and social media accounts of at least 38 journalists and other media workers from … Narasi,[923] [they] were the targets of a coordinated hacking attack.”[924]

In 2024 a “Pwn2Own hacking event” took place which pitted hackers against each other for prize money. Focusing on electric vehicles, Tesla cars were successfully hacked twice.[925]

12.5.5 AI in Healthcare

12.5.5.1 AI Decisions in Healthcare

AI has already shown success assisting in radiology diagnostics for cancer, knees and chests, accurately identifying ailments.[926]

Wearable technology[927] can detect and transit crucial data monitoring a patient’s day to day vital signs.[928, 929]

AI can, for example, accurately predict mortality rates and infants’ risks prior to birth.[930]

[921] Sheldon, Hanna, 2022
[922] Amnesty International, 2023
[923] Narasi is an online media outlet
[924] Amnesty International, 2023
[925] Winder, 2024
[926] Kasperbauer, 2021
[927] Smart watches for example
[928] For example, heart rate and bold pressure
[929] Kasperbauer, 2021
[930] Kasperbauer, 2021

12.6 Appendix 6 – *Chapter 7 Notes*, AI the Social Disruptor

12.6.1 Social Media – Crime and Hate

Top 32 Social Media Platforms by Monthly Active Users

Table IV – Top 32 Social Media Platforms (January 2024)[931]

Rank	Platform	Monthly Active Users
1	Facebook	3.05 Billion
2	WhatsApp	2.78 Billion
3	You Tube	2.70 Billion
4	Instagram	2.35 Billion
5	Tik Tok	1.67 Billion
6	WeChat	1.67 Billion
7	Messenger	1.30 Billion
8	Telegram	800 Million
9	Snapchat	750 Million
10	Douyin	743 Million
11	Qzone	605 Million
12	Spotify	600 Million
13	Sina Weibo	599 Million
14	Kaishou	580 Million
15	Pinterest	450 Million
16	Reddit	430 Million
17	Twitter (X)	368.4 Million
18	LinkedIn	310 Million
19	Quora	305 Million
20	Skype	300 Million
21	Microsoft Teams	280 Million
22	Line	230 Million
23	SoundCloud	175 Million
24	Vimeo	170 Million
25	Discord	154 Million
26	Twitch	140 Million
27	Tumblr	135 Million
28	Stack Exchange	105 Million
29	Threads	100 Million
30	VK (Vkontakte)	80 Million
31	Rumble	78 Million
32	Next Door	69 Million

[931] Howarth, 2024

12.6.1.1 Emotions – Usefulness, Fear and Distrust

Matthias Scheutz lists twelve emotions that would be of practical benefit in an AI context.[932]

1. *Alarm mechanisms* – e.g., fast reflex-like reactions in critical situations such as fear processes, that interrupt current behaviour and initiate a retreat response, moving the agent away from the danger zone
2. *Action selection* – e.g., deciding what to do next based on the current emotional state, such as switching from exploration to foraging behaviour based on the agent's needs
3. *Adaptation* – e.g., short – or long-term changes in behaviour due to affective states, such as adapting one's gait to uneven terrain based on negative affect generated by sensors
4. *Social regulation* – e.g., using emotional signals to achieve social effects, such as aggressive display to deter another agent from interfering with one's activity
5. *Learning* – e.g., using affective evaluations as utility estimates in reinforcement learning, such as learning the utility of different behaviours to achieve goals in different contexts
6. *Motivation* – e.g., adopting goals as part of an emotional coping mechanism, such as when a high level of distress and frustration leads to adopting the goal of asking a human supervisor for help
7. *Goal management* – e.g., the creation of new goals or reprioritisation of existing ones, such as using positive and negative affect to modify cost estimates used in the calculation of the expected utility of a goal
8. *Strategic processing* – e.g., the selection of search strategies based on overall affective state, such as using positive and negative affect to bias search algorithms to top-down versus bottom-up search
9. *Memory control* – e.g., the strategic use of affective bias on memory access and retrieval as well as decay rate of memory items, such as using current affective state to rank memory items with similar affect as better matches

[932] Matthias Scheutz in Frankish and Ramsey, 2014, 251-252

10. *Information integration* – e.g., emotional filtering of data from various information channels or blocking of such integration, such as ignoring positive valenced information from vision sensors about a happy face when the acoustic information suggests an angry voice
11. *Attentional focus* – e.g., selection of data to be processed based on affective evaluation, such as biasing visual search in favour of objects the agent highly desires
12. *Self model* – e.g., using affect as a representation of "what a situation is like for the agent", such as using the overall affective evaluation of different components of the agent's control system as a measure of the agent's overall mood and how it "feels."

12.6.1.2 Social Media – Political Propaganda and Manipulation

12.6.1.2.1 Cambridge Analytica

The following extracts, from Harbath and Fernekes (2023), provide background and information on the issues that arose from Cambridge Analytica's use of data gleaned from Facebook. [933]

> Cambridge Analytica was founded in 2013. Its ability to micro target goes back to a 2010 Facebook rollout that included a new version of its Application Programming Interface (API), a tool that allows one application to access the data or features of another. With the API, "developers could now see social connections between people, and see the connections people have based on their interests and likes."
>
> Its uses included President Obama's campaign, which built an app that would connect known Obama supporters to potential supporters. The idea was that these two groups of users (demonstrated supporters and potential ones) had something in common, such as being friends on Facebook or that they both liked a particular sports team.

[933] Harbath and Fernekes, 2023

On March 17, 2018, the Guardian and New York Times broke the story about a Cambridge Analytica whistleblower, saying that the company had utilised 50 million Facebook profiles to do their modelling.

Facebook announced new measures to better protect people's data.

The fallout and response from governments, media, civil society, and others worldwide were swift. American and British officials immediately held hearings to question Facebook about what happened. The European Union, Canada, and Indian governments did as well.

The scandal influenced legislative action worldwide, most notably the Digital Services Act and Digital Markets Act in the European Union. It sparked a conversation around the use of microtargeting for political advertisements. Facebook, Google, and other platforms created transparency tools so users could better see all of the advertisements that campaigns were running. Currently, in the EU, legislation that would severely limit the use of microtargeting for campaigns is moving through Parliament. There has also been a slew of litigation, most recently, Facebook settled a lawsuit on the matter for $725 million.

The Cambridge Analytica controversy profoundly impacted the world of data privacy, political campaigning, and social media. Governments worldwide enacted laws and regulations to protect consumers, and companies needed to adjust their practices in response.

12.6.2 AI's Manipulative Force

12.6.2.1 Consumer Manipulation

12.6.2.1.1 Tesco, UK

Tesco is a large UK based supermarket chain[934, 935] It uses Business (data) Analytics to generate business value. Its Clubcard loyalty program motivates customers to present their card with purchases. Tesco analyses data to understand customers' purchasing preferences. It then automates offers in its direct-marketing program.

[934] In 2023 Tesco's group turnover was £65.762Bn and had 4,169 stores
[935] Tescoplc.com, 2023

This has been so effective that Tesco had a coupon redemption rate ten times the industry average.[936] This illustrates how such data can help consumers save money, but only when shopping with Tesco. Tesco increased sales and now has a detailed analysis, over an increasingly long period, of individuals and their shopping patterns. Some many find this intrusive, even though they continue to use the Clubcard.

12.6.2.1.2 Target, USA

"Target" a large US retailer,[937] uses data analytics extensively. It can identify pregnant customers. A Target statistician used personal identifiable information (PII) to notice women buying large quantities of unscented lotion at the beginning of their second trimester. It also showed that during the first 20 weeks of pregnancy women purchased large amounts of calcium, magnesium and zinc. This information can be used to create psychological strategies to influence spending. In Target's Minneapolis store a father complained about "Target sending his daughter a sale booklet for baby clothes, cribs, and diapers even though she was still in high school." It later transpired that the daughter was pregnant, Target knew this (from its prediction score) before she even told her mother and father.[938]

12.6.3 Truth and Trust

12.6.3.1 Conspiracy Theories

In 2019 former President Trump aide James Comey sent a 'tweet'.[939] It included the phrase "Five Jobs I've Had" which was interpreted as "Five Jihad." The QAnon Group exchanged messages and concluded that a child would be abducted at a school event in Nevada County, California. Following concerns from parents and school officials, the event had to be cancelled.[940]

On 06 January 2021 there was an attack on the Capitol building in Washington. "Before the 2020 election … concluded, President Trump laid the groundwork for an alternate reality …, falsely assailing the integrity of the race at nearly every turn." "…

[936] Sedden *et al*, 2016
[937] Target.com, 2024
[938] Kuhn, 2023
[939] A message on the messaging App 'Twitter', now named 'X'
[940] Rothschild, 2021, 75-78

he made more than 800 inaccurate claims about the election from the time the polls began closing … characterised the election as "rigged," "stolen" or "a hoax," and flatly and falsely declared he had won."[941] On 06 January 2021, during a Joint Session of Congress to certify the election results,[942] the attack took place following speeches and tweets from Trump which included, "Be there, will be wild!", "We're just not going to let that happen" and "fight like hell." "… a mob of his supporters … including adherents of QAnon [took part in the attack], … many [police] officers were brutally beaten …, sprayed with chemical irritants, or crushed and trampled …"[943] "Five people died and four police officers who worked at the Capitol that day died by suicide in the months that followed."[944]

12.6.3.2 Conspiracy Led Legislation

Conspiracy theories led to legislation addressing speculation about future dangers. In relation to Geoengineering[945] it led Rhode Island state Representative Robert Quattrocchi to sponsor an "Atmosphere Protection Bill" to outlaw a variety of pollutants more commonly found in science fiction than the troposphere.[946] These include "chips or sensors" that would enable "warrantless surveillance and control" after being "worn, ingested, inhaled, and/or injected."

The North Dakota Representative Cole Christensen "pushed the new law to assert the superiority of humans over digital systems rather than encouraging conspiracy theories." This was based on reports that AI got a non-voting seat on a corporate board in Singapore and a recording contract with Warner Music Group.[947]

12.6.4 Freedoms in the AI (Internet) Age

A determinist's[948] views the agent having a range of causal factors leading to an ultimate choice. The causal factors lead only one way towards the choice, which

[941] Qiu, 2023
[942] To certify the result of the Presidential election which Joe Biden had won
[943] Duignan, 2024
[944] Kates and MacFarlane, 2023
[945] A theoretical way to address global warming by blocking solar radiation through means like aerosol sprays
[946] The lowest layer of the Earth's atmosphere in direct contact with the Earth's surface
[947] Williams, 2023
[948] Determinism is the view events are inevitable and people's decisions could not have been otherwise. Source Brittanica, https://www.britannica.com/topic/determinism, consulted 11 April 2024

was not freely made as the agent had no option to do otherwise. This is again refuted by Copeland who says that in deliberating about the choice, the agent is author and implements "his own decision, in accordance with his own best perceptions of his wants, needs, and circumstance. If this is not freedom, then what is?"[949]

Mele and Descartes, respectively, provide similar views to Sartre:

> It may be true both that an agent would not have done A if he had seen what he needed to do to mount a successful resistance against his desire to do A and that, if he had been stronger or more resolute, a special effort of resistance would have been unnecessary.[950]

> My habitual opinions keep coming back, and despite my wishes, they capture my belief, which is as it were bound over to them as a result of long occupation and the law of custom.[951]

12.6.5 Freedom of Thought and Speech

With AI individuals' autonomy is gradually challenged and freedom of thought eroded.[952]

> 'In short, online we can intentionally restrict our interactions to those of exactly the same opinion sets as our own', whereas 'real world interaction cannot be so easily and deliberately structured and limited'; this 'narrowing of focus and community' tends to 'make us more prejudiced and our attitudes more insular' and, ultimately, leads to increased social cleavage and division. Moreover, as people want to be perceived favourably by other group members, they often 'adjust their position in the direction of the dominant position'. Social relations and interactions become polarised and exacerbated by feelings of hatred or fear or anger. Thus when they are legitimised and spread by a growing number of people or groups, society is

[949] Copeland, 1993, 146
[950] Alfred Mele quoted in Tenenbaum, 1999
[951] Descartes quoted in Tenenbaum, 1999
[952] Floridi et al., 2016 in Giovanola, 2023, 112

filled with negative emotions, which makes social cooperation harder and harder.[953]

12.6.5.1 Freedom of Speech

The First Amendment to the Constitution of the United States of America was enacted on 15 December 1791.[954] The Amendment states:

> Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.[955]

Freedom of Speech is restricted by some governments. Russia is an extreme instance where courts deal harshly with government critics, as illustrated by the critic Alexei Navalny[956] who died in prison in 2024.[957] In China too there are severe restrictions on freedom of speech and expression. In 2020 China introduced the National Security Law (NSL) and other repressive laws, which were widely used to target people exercising their rights to freedom of expression, peaceful assembly and association. "Political activists, journalists, human rights defenders and others charged under the NSL were held for prolonged pretrial detention. … at least 230 people had been arrested under the NSL since its enactment in 2020."[958]

[953] Giovanola, 2023, 113
[954] The Library of Congress, nd
[955] The Congressional Research Service, nd
[956] In 2023 Navalny was convicted of 'extremism' and sentenced to 19 years imprisonment,
[957] Burrows and Litvinova, 2024
[958] Amnesty International, 2023, 127

12.7 Appendix 7 – *Chapter 8 Notes*, Wars and Conflicts with AI

12.7.1 AI to Direct Bombing and Battlefield Attacks

12.7.1.1 Identify and Direct Bombing of Human Targets

AI can be effective and ruthless in war and reduce human involvement in directing operations. Following the Hamas attack on 07 October 2023,[959] Israel launched a prolonged[960] retaliatory attack in the Gaza Strip.[961, 962]

> [Israel designed a] machine that could rapidly process massive amounts of data to generate thousands of potential "targets" for military strikes in the heat of a war", which resolves a "human bottleneck for both locating the new targets and decision-making to approve the targets.
>
> [This uses] AI to generate targets for assassination, … military's operations … treated the outputs of the AI machine "as if it were a human decision." Using the AI's output "human personnel often served only as a "rubber stamp" … devot[ing]… "20 seconds" to each target before authorising a bombing.
>
> [Israel can identify] targeted individuals and carry out bombings when they [the targets] had entered their family's residences" the tactic usually killed the entire family, as collateral damage. A journalist's source was quoted saying "the number of civilians they were allowed to kill alongside each target was fixed during the initial weeks of the war at up to 20.

Before 07 October the system was known to produce error in 10% of cases.[963, 964]

---

[959] 1,200 people were killed during the Hamas attack, Frankel, 2024; and more than 250 people abducted, Associated Press, 2024

[960] More than eight months at the time of writing

[961] Frankel, 2024

[962] In June 2024 the toll of those killed by Israel in Gaza was more than 36,000 Palestinians, Associated Press, 2024; including more than 15,000 children, Johnson, 2024

[963] Abraham, 2024

[964] From a report produced by +972 Magazine "… an independent, online, nonprofit magazine run by a group of Palestinian and Israeli journalists", +972 Magazine https://www.972mag.com/about/, accessed 15 April 2024

12.8 Appendix 8 – *Chapter 9 Notes*, Political, Economic and Regulatory Issues and Responses to AI

12.8.1 Deny, Doubt and Delay – The Political Challenges

The 2022 Frontline[965] documentary "The Power of Big Oil" showed the Denial, Doubt and Delay tactics used by oil companies in relation to climate change. The industry's research was designed to sow the seeds of doubt about climate science; stall climate policy in the face of emerging and more certain evidence; and their efforts to delay the transition to renewable energy sources, and promoting natural gas as a cleaner alternative.[966] The Guardian's headline "'What we now know … they lied': how big oil companies betrayed us all" summed up the tactics and how the companies influence US law makers.[967] One US senator acknowledged:

> that fossil fuel companies manipulated Congress with a stream of false information, and accuses the oil industry of malignly claiming the science of climate change was not proved when companies such as Exxon and Shell already knew otherwise from their own research.[968]

12.8.1.1 Tobacco CEOs Deny Nicotine is Addictive

April 14, 1994 - Hearing on the Regulation of Tobacco Products House Committee on Energy and Commerce Subcommittee on Health and the Environment

Full extract from the webpage of University of California San Francisco.[969]

Tobacco CEO's Statement to Congress 1994 News Clip "Nicotine is not addictive."

> April 14, 1994 - Hearing on the Regulation of Tobacco Products House Committee on Energy and Commerce Subcommittee on Health and the Environment.

[965] An arm of the US based Public Broadcasting Service (PBS)
[966] Frontline, 2022
[967] McGreal, 2022
[968] McGreal, 2022
[969] University of California San Francisco, nd

The Subcommittee met, pursuant to notice, at 9:05 a.m., 2123 Rayburn House Office Building, Hon. Henry A. Waxman (chairman) presiding.

Opening Statement from Chairman Henry A. Waxman

REP. WAXMAN:

The meeting of the subcommittee will come to order. I'd like to ask our guests to please take your seats. This is an historic hearing. For the first time ever, the chief executive officers of our Nation's tobacco companies are testifying together before the U.S. Congress. They are here because this subcommittee has legislative jurisdiction over those issues that affect our health. And no health issue is as important as cigarette smoking. It is sometimes easier to invent fiction than to face the truth. The truth is that cigarettes are the single most dangerous consumer product ever sold. Nearly a half million Americans die every year as a result of tobacco. This is an astounding, almost incomprehensible statistic. Imagine our Nation's outrage if two fully loaded jumbo jets crashed each day, killing all aboard. Yet that is the same number of Americans that cigarettes kill every 24 hours. Sadly, this deadly habit begins with our kids. Each day 3,000 children will begin smoking. In many cases they become hooked quickly and develop a life long addiction that is nearly impossible to break. For the past 30 years a series of surgeons general have issued comprehensive reports outlining the dangers these children will eventually face. Lung cancer, heart disease, emphysema, bladder cancer, and stroke are only some of the diseases caused by tobacco causes. And now we know that kids will face a serious health threat even if they don' t smoke. Environmental tobacco smoke is a Class A carcinogen, and it sickens more than 1 million kids every year. In fact, five former surgeons general of the United States testified before this subcommittee this year, that the most important legislation in disease prevention that we could enact would be restrictions on smoking in public places. This subcommittee

will soon act on that legislation, and it will consider other measures as well. This hearing will aid our efforts by presenting an important perspective. But these hearings are important for another reason as well. For decades the tobacco companies have been exempt from the standards of responsibility and accountability that apply to all other American corporations. Companies that sell that sell aspirin, cars, and soda are all held to strict standards when they cause harm. We don't allow those companies to sell goods that recklessly endanger consumers. We don't allow them to suppress evidence of dangers when harm occurs. We don't allow them to ignore science and good sense. And we demand that when problems occur, corporations and their senior executives be accountable to Congress and the public. This hearing marks the beginning of a new relationship between Congress and the tobacco companies. The old rules are out, the standards that apply to every other company are in. We look for- ward to hearing the testimony this morning, and to working with these companies to begin to reduce the extraordinary public health threat that tobacco poses.

An old proverb says that a journey of a thousand miles must begin with a single step. Today is the first step. Many more are to come as we deal with the most serious health problem facing our Nation.

[Tobacco company CEOs declare, under oath, that nicotine is not addictive]

REP. RON WYDEN:

Let me begin my questioning on whether or not nicotine is addictive. Let me ask you first, and I'd like to just go down the row, whether each of you believes that nicotine is not addictive. I heard virtually all of you touch on it. Yes or no, do you believe nicotine is not addictive?

MR. WILLIAM CAMPBELL:

I believe nicotine is not addictive, yes.

REP. RON WYDEN:

Mr. Johnston?

MR. JAMES JOHNSTON:

Mr. Congressman, cigarettes and nicotine clearly do not meet the classic definition of addiction. There is no intoxication.

REP. RON WYDEN:

We'll take that as a "no." Again, time is short. I think that each of you believe that nicotine is not addictive. We would just like to have this for the record.

MR. JOSEPH TADDEO:

I don't believe that nicotine or our products are addictive.

MR. ANDREW TISCH:

I believe that nicotine is not addictive.

MR. EDWARD HORRIGAN:

I believe that nicotine is not addictive.

MR. THOMAS SANDEFUR:

I believe that nicotine is not addictive.

MR. DONALD JOHNSTON:

And I, too, believe that nicotine is not addictive.

Witnesses:

William Campbell, President & CEO, Philip Morris, USA

James W. Johnston, Chairman and CEO, R.J. Reynolds Tobacco Company

Joseph Taddeo, President, U.S. Tobacco Company

Andrew H. Tisch, Chairman and CEO, Lorillard Tobacco Company

Edward A. Horrigan, Chairman and CEO, Liggett Group Inc.

Thomas E. Sandefur, Chairman and CEO, Brown and Williamson Tobacco Corp.

Donald S. Johnston, President and CEO, American Tobacco Company

Chaired by:

Henry Waxman (D-CA)

12.8.2 AI and Digital Companies' Dominance

In economic and competition law market dominance can be a technical construct, principally focussing on the level of the dominant firm's market share, ranging from 15% to 40% depending on the dynamics of the market.[970, 971] A further view posits that where there are few market constraints it allows firms to behave independently from other market actors.[972, 973]

Companies possess …

> enormous economic and political power, which they can and do abuse. … companies have been granted the privileges of personhood, including the political privileges of citizens. … being so mobile and flexible, they can afford to be indifferent to the fate of their workers and the countries in which they are located and so become sophisticated avoiders of taxes and regulations. Moreover, judicial systems find it almost impossible to hold them or their executives criminally liable, even in the event of significant malfeasance. …

[970] Rexel, 2020

[971] "Typically, a company is considered to hold a dominant position if it has a market share of more than 40%, but even a market share of 15% may be considered dominant if it is the largest player in a fragmented market", Rexel, 2020

[972] Graham, 2023

[973] Companies are usually constrained in their commercial behaviour by competitors, customers and consumers. Where such constraints are weak and ineffective, the company is considered to hold market power. When market power allows a company, over time, to behave independently from other market actors, such company is deemed to exercise a dominant market position. Dominance per se is not unlawful. Rather, it is the abuse of dominance – particularly if used to weaken competitive dynamics by excluding rivals and harming consumers – that is unlawful. Graham, 2023

the goals of the company, its dominant purpose was long held to be maximising shareholder value, to the exclusion of other objectives. This can encourage behaviour that borders on the sociopathic.[974]

12.8.2.1 AI Companies, US Concentration

12.8.2.1.1 Location of Top 15 AI Companies[975]

Of the top 15 AI companies worldwide, by Market Capitalisation, thirteen are in the USA, with six of those based in California.

1. Microsoft (MSFT) - Headquarters: Redmond, WA, **USA**
2. Alphabet (GOOG) – Headquarters: Mountain View, CA, **USA**
3. NVIDIA (NVDA) – Headquarters: Santa Clara, CA, **USA**
4. Meta Platforms (META) – Headquarters: Menlo Park, CA, **USA**
5. Tesla (TSLA) – Headquarters: Austin, TX, **USA**
6. IBM (IBM) – Headquarters: Armonk, NY, **USA**
7. Palantir (PLTR) – Headquarters: Denver, CO, **USA**
8. Mobileye (MBLY) – Headquarters: Jerusalem, Israel
9. Dynatrace (DT) – Headquarters: Waltham, MA, **USA**
10. UiPath (PATH) – Headquarters: New York, NY, **USA**
11. SentinelOne (S) – Headquarters: Mountain View, CA, **USA**
12. Aurora Innovation (AUR) – Headquarters: Pittsburgh, PA, **USA**
13. Presight AI (PRESIGHT.AE) – Headquarters: Al Ain, Abu Dhabi, United Arab Emirates
14. Darktrace (DARK.L) – Headquarters: Cambridge, UK; London, UK; San Francisco, CA, **USA**; Singapore
15. C3 AI (AI) – Headquarters: Redwood City, CA, **USA**

[974] Wolf, 2023, 51
[975] Stash Banking Services, 2024

12.8.2.2 AI Companies Global Financial Power and Strength

12.8.2.2.1 Top Ten Global Companies vs. Top Ten Countries (Market Capitalisation Values vs. GDPs)

Top Ten Global Companies Value (August 2023)[976]						Top Ten Countries – GDP[977] (August 2023)[978]		
Rank	**Company**	**Sector**	**Country**	**Market Capitalisation $Bns**		**Rank**	**Country**	**GDP $Bns**
1	Apple	Information Technology	USA	$2,777		**1**	USA	$26,855B
2	Microsoft	Information Technology	USA	$2,382		**2**	China	$19,374B
3	Saudi Aramco	Energy	Saudi Arabia	$2,222		**3**	Japan	$4,410B
4	Alphabet	Communication Services	USA	$1,636		**4**	Germany	$4,309B
5	Amazon	Consumer Discretionary	USA	$1,385		**5**	India	$3,737B
6	NVIDIA	Information Technology	USA	$1,074		**6**	UK	$3,159B
7	Berkshire Hathaway	Financials	USA	$774		**7**	France	$2,923B
8	Meta Platforms	Communication Services	USA	$754		**8**	Italy	$2,170B
9	Tesla	Consumer Discretionary	USA	$715		**9**	Canada	$2,090B
10	Eli Lilly	Health Care	USA	$519		**10**	Brazil	$2,081B

The seven AI companies listed are in the Information Technology Information, Communication Services and Consumer Discretionary sectors. They rely on AI for their business operations, customer engagement and customer services.

There are key points to note which illustrate the individual and collective power of the AI companies:

1) The seven AI companies are all located in the USA.
2) With a market value of $2,777 US$Bn, Apple is larger than Italy's GDP of $2,170 US$Bn. Italy is the eighth largest economy in the world.[979]
3) Excluding Saudi Aramco (Energy), Berkshire Hathaway (Financials) and Eli Lilly (Health Care), the seven companies in the Top Ten have a combined

[976] Lu, 2023
[977] Gross Domestic Product
[978] Rao, 2023
[979] Rao, 2023

market value of $10,723 US$Bn, well ahead of Japan, the third largest economy in the world, which has a GDP of $4,410 US$Bn.[980]

4) The $10,723 US$Bn combined market values, of the seven AI companies is:
    a. greater than the $9,264 US$Bn combined GDPs of the bottom four countries; France, Italy, Canada and Brazil.
    b. greater than the $ 9,402 US$Bn GDPs of the three largest EU countries in the top ten; Germany, France and Italy.
    c. 64% of the $ 16,747 US$Bn EU's GDP in 2022.[981]

12.8.2.3 The Pacing Problem – The Legal and Regulatory Struggle

12.8.2.3.1 Open Letter issued by Elon Musk and Other Leading Figures

An Open Letter from Elon Musk and other leading figures was issued on 22 March 2023. The letter has 33,708 signatories (as of 05 June 2024).[982]

> Pause Giant AI Experiments: An Open Letter
>
> We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.
>
> AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research and acknowledged by top AI labs. As stated in the widely-endorsed Asilomar AI Principles, Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources. Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.
>
> Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman

[980] Rao, 2023
[981] World Bank, nd, for EU GDP in 2022
[982] Bengio, Musk, et al, 2023

minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable. This confidence must be well justified and increase with the magnitude of a system's potential effects. OpenAI's recent statement regarding artificial general intelligence, states that "At some point, it may be important to get independent review before starting to train future systems, and for the most advanced efforts to agree to limit the rate of growth of compute used for creating new models." We agree. That point is now.

Therefore, we call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4. This pause should be public and verifiable, and include all key actors. If such a pause cannot be enacted quickly, governments should step in and institute a moratorium.

AI labs and independent experts should use this pause to jointly develop and implement a set of shared safety protocols for advanced AI design and development that are rigorously audited and overseen by independent outside experts. These protocols should ensure that systems adhering to them are safe beyond a reasonable doubt. This does not mean a pause on AI development in general, merely a stepping back from the dangerous race to ever-larger unpredictable black-box models with emergent capabilities.

AI research and development should be refocused on making today's powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal.

In parallel, AI developers must work with policymakers to dramatically accelerate development of robust AI governance systems. These should at a minimum include: new and capable regulatory authorities dedicated to AI; oversight and tracking of highly capable AI systems and large pools of computational capability; provenance and watermarking systems to help

distinguish real from synthetic and to track model leaks; a robust auditing and certification ecosystem; liability for AI-caused harm; robust public funding for technical AI safety research; and well-resourced institutions for coping with the dramatic economic and political disruptions (especially to democracy) that AI will cause.

Humanity can enjoy a flourishing future with AI. Having succeeded in creating powerful AI systems, we can now enjoy an "AI summer" in which we reap the rewards, engineer these systems for the clear benefit of all, and give society a chance to adapt. Society has hit pause on other technologies with potentially catastrophic effects on society. We can do so here. Let's enjoy a long AI summer, not rush unprepared into a fall.

12.8.2.3.2 The Bletchley Declaration

The Bletchley Declaration was issued by Countries Attending the AI Safety Summit, 01-02 November 2023. It was published on 01 November 2023.[983]

Artificial Intelligence (AI) presents enormous global opportunities: it has the potential to transform and enhance human wellbeing, peace and prosperity. To realise this, we affirm that, for the good of all, AI should be designed, developed, deployed, and used, in a manner that is safe, in such a way as to be human-centric, trustworthy and responsible. We welcome the international community's efforts so far to cooperate on AI to promote inclusive economic growth, sustainable development and innovation, to protect human rights and fundamental freedoms, and to foster public trust and confidence in AI systems to fully realise their potential.

AI systems are already deployed across many domains of daily life including housing, employment, transport, education, health, accessibility, and justice, and their use is likely to increase. We recognise that this is therefore a unique moment to act and affirm the need for the safe development of AI and for the transformative opportunities of AI to be used for good and for all, in an inclusive manner in our countries and globally. This includes for public

[983] AI Safety Summit, 2023

services such as health and education, food security, in science, clean energy, biodiversity, and climate, to realise the enjoyment of human rights, and to strengthen efforts towards the achievement of the United Nations Sustainable Development Goals.

Alongside these opportunities, AI also poses significant risks, including in those domains of daily life. To that end, we welcome relevant international efforts to examine and address the potential impact of AI systems in existing fora and other relevant initiatives, and the recognition that the protection of human rights, transparency and explainability, fairness, accountability, regulation, safety, appropriate human oversight, ethics, bias mitigation, privacy and data protection needs to be addressed. We also note the potential for unforeseen risks stemming from the capability to manipulate content or generate deceptive content. All of these issues are critically important and we affirm the necessity and urgency of addressing them.

Particular safety risks arise at the 'frontier' of AI, understood as being those highly capable general-purpose AI models, including foundation models, that could perform a wide variety of tasks - as well as relevant specific narrow AI that could exhibit capabilities that cause harm - which match or exceed the capabilities present in today's most advanced models. Substantial risks may arise from potential intentional misuse or unintended issues of control relating to alignment with human intent. These issues are in part because those capabilities are not fully understood and are therefore hard to predict. We are especially concerned by such risks in domains such as cybersecurity and biotechnology, as well as where frontier AI systems may amplify risks such as disinformation. There is potential for serious, even catastrophic, harm, either deliberate or unintentional, stemming from the most significant capabilities of these AI models. Given the rapid and uncertain rate of change of AI, and in the context of the acceleration of investment in technology, we affirm that deepening our understanding of these potential risks and of actions to address them is especially urgent.

Many risks arising from AI are inherently international in nature, and so are best addressed through international cooperation. We resolve to work together in an inclusive manner to ensure human-centric, trustworthy and responsible AI that is safe, and supports the good of all through existing international fora and other relevant initiatives, to promote cooperation to address the broad range of risks posed by AI. In doing so, we recognise that countries should consider the importance of a pro-innovation and proportionate governance and regulatory approach that maximises the benefits and takes into account the risks associated with AI. This could include making, where appropriate, classifications and categorisations of risk based on national circumstances and applicable legal frameworks. We also note the relevance of cooperation, where appropriate, on approaches such as common principles and codes of conduct. With regard to the specific risks most likely found in relation to frontier AI, we resolve to intensify and sustain our cooperation, and broaden it with further countries, to identify, understand and as appropriate act, through existing international fora and other relevant initiatives, including future international AI Safety Summits.

All actors have a role to play in ensuring the safety of AI: nations, international fora and other initiatives, companies, civil society and academia will need to work together. Noting the importance of inclusive AI and bridging the digital divide, we reaffirm that international collaboration should endeavour to engage and involve a broad range of partners as appropriate, and welcome development-orientated approaches and policies that could help developing countries strengthen AI capacity building and leverage the enabling role of AI to support sustainable growth and address the development gap.

We affirm that, whilst safety must be considered across the AI lifecycle, actors developing frontier AI capabilities, in particular those AI systems which are unusually powerful and potentially harmful, have a particularly strong responsibility for ensuring the safety of these AI systems, including through systems for safety testing, through evaluations, and by other appropriate measures. We encourage all relevant actors to provide context-appropriate

transparency and accountability on their plans to measure, monitor and mitigate potentially harmful capabilities and the associated effects that may emerge, in particular to prevent misuse and issues of control, and the amplification of other risks.

In the context of our cooperation, and to inform action at the national and international levels, our agenda for addressing frontier AI risk will focus on:

- identifying AI safety risks of shared concern, building a shared scientific and evidence-based understanding of these risks, and sustaining that understanding as capabilities continue to increase, in the context of a wider global approach to understanding the impact of AI in our societies.
- building respective risk-based policies across our countries to ensure safety in light of such risks, collaborating as appropriate while recognising our approaches may differ based on national circumstances and applicable legal frameworks. This includes, alongside increased transparency by private actors developing frontier AI capabilities, appropriate evaluation metrics, tools for safety testing, and developing relevant public sector capability and scientific research.

In furtherance of this agenda, we resolve to support an internationally inclusive network of scientific research on frontier AI safety that encompasses and complements existing and new multilateral, plurilateral and bilateral collaboration, including through existing international fora and other relevant initiatives, to facilitate the provision of the best science available for policy making and the public good.

In recognition of the transformative positive potential of AI, and as part of ensuring wider international cooperation on AI, we resolve to sustain an inclusive global dialogue that engages existing international fora and other relevant initiatives and contributes in an open manner to broader international discussions, and to continue research on frontier AI safety to

ensure that the benefits of the technology can be harnessed responsibly for good and for all. We look forward to meeting again in 2024.

The countries represented were:

Australia	*Kingdom of Saudi Arabia*
Brazil	*Netherlands*
Canada	*Nigeria*
Chile	*The Philippines*
China	*Republic of Korea*
European Union	*Rwanda*
France	*Singapore*
Germany	*Spain*
India	*Switzerland*
Indonesia	*Türkiye*
Ireland	*Ukraine*
Israel	*United Arab Emirates*
Italy	*United Kingdom of Great*
Japan	*Britain and Northern Ireland*
Kenya	*United States of America*

References to 'governments' and 'countries' include international organisations acting in accordance with their legislative or executive competences.

12.8.2.4 The EU's, UK's and US's Responses

Effective from 17 February 2024, the EU's Digital Services Act, regulates "online intermediaries and platforms … Its main goal is to prevent illegal and harmful activities online and the spread of disinformation. It ensures user safety, protects fundamental rights, and creates a fair and open online platform environment."[984] The Act's clauses "cover a wide spectrum of illegal content that platforms will be required to address and give platforms a "duty of care" over what their users – particularly children – see online."[985] This duty of care imposes a considerable

[984] European Commission (a), nd
[985] Guest, 2023

burden on the platforms.[986, 987] There are "nuanced principles around the harms caused by legal but harmful content" and "requirement for messaging platforms to scan users' messages for illegal material, such as child sexual abuse material", causing the platforms and privacy campaigners to say that it is "an unwarranted attack on encryption."[988]

In the UK on 26 October 2023 the 'Online Safety Act 2023' came into force.[989] Under the Act "Platforms will have to prevent younger users from seeing age-inappropriate content, …; release risk assessments on potential dangers to children …; and give parents easy pathways to report concerns." The platforms must also remove threats of violence, including rape, encouragement of self-harm, transmission of deepfake pornography and scam advertisements.[990, 991]

12.8.2.4.1 EU's Artificial Intelligence Act

The EU's Artificial Intelligence Act was adopted on 13 March 2024.[992] Article 3 (1) defines AI as:[993]

> 'AI system' means a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

Within the Act it bans certain AI applications

> that threaten citizens' rights, including biometric categorisation systems based on sensitive characteristics and untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases. Emotion recognition in the workplace and schools, social scoring, predictive policing …,

[986] Including opening them to "fines up to 6% of the worldwide annual turnover"
[987] European Commission (b), nd
[988] Guest, 2023
[989] UK National Archives, 2023
[990] There are also fines of up to the greater of £18 million or 10 percent of annual revenue that can be imposed for breaches of the Act
[991] Guest, 2023
[992] European Parliament (a)
[993] European Parliament (b)

and AI that manipulates human behaviour or exploits people's vulnerabilities will also be forbidden.[994]

The use of biometric identification systems[995] by law enforcement is prohibited in principle, except in exhaustively listed and narrowly defined situations. Obligations are foreseen for high-risk AI systems (due to potential harm to health, safety, fundamental rights, environment, democracy and the rule of law). Transparency requirements for General-Purpose AI (GPAI) systems, must meet certain transparency requirements, including compliance with EU copyright law and publishing detailed summaries of the content used for training.[996]

AI companies can be fined €15m or 3% of worldwide turnover for breaching certain provisions of the act (e.g. transparency obligations), up to €35m, or 7% of turnover, for deploying or developing banned AI tools,[997]

12.8.2.4.2 Amnesty International Statement

Amnesty International's Statement of the EU's Artificial Intelligence Act.[998]

> EU: Artificial Intelligence rulebook fails to stop proliferation of abusive technologies
>
> Responding to the European Parliament's adoption of the European Union's Artificial Intelligence Act (AI Act) today, Mher Hakobyan, Amnesty International's Advocacy Advisor on Artificial Intelligence, said:
>
> "While EU policymakers are hailing the AI Act as a global paragon for AI regulation, the legislation fails to take basic human rights principles on board.
>
> "Even though adopting the world's first rules on the development and deployment of AI technologies is a milestone, it is disappointing that the EU

[994] European Parliament (b)
[995] Formally known as Remote Biometric Identification (RBI)
[996] European Parliament (b)
[997] Milmo and Hern, 2024
[998] Amnesty International, 13 March 2024

and its 27 member states chose to prioritize the interest of industry and law enforcement agencies over protecting people and their human rights.

"The AI Act offers only limited protections to impacted and marginalized people. It does not ban the reckless use and export of draconian AI technologies, while also failing to ensure equal protection to migrants, refugees and asylum seekers. Similarly, it lacks proper accountability and transparency provisions, which will likely exacerbate human rights abuses.

"Countries outside of the EU should learn from the bloc's failure to adequately regulate AI technologies and must not succumb to pressures by the technology industry and law enforcement authorities whilst developing regulation. States should instead put in place robust and binding AI legislation which prioritizes people and their rights."

12.8.2.5 Tackling AI Companies' Dominance

AI companies' dominance can be eliminated or diluted, their businesses can be broken up into logical units with new owners and anti-competitive practices not merely outlawed but eradicated. The creation of the "Baby-Bells" is a significant precedent in the USA. In 1982, AT&T[999] announced that it had settled an Antitrust[1000] dispute with the US Department of Justice. AT&T was required to divest its interests in twenty-two Bell Operating Companies (BOCs) that provided local telephone services. The seven new companies, known as the Baby-Bells, each provided telephone services across a number of states. AT&T's firm grip on the US telecommunications services was relinquished and AT&T was allowed to enter new markets to provide other services.

12.8.2.6 Politically Motivated Measures

Political rather than competition motivations can lead government to bans. In 2020 India banned Tiktok,[1001] "The government cited privacy concerns and said that

[999] American Telephone and Telegraph Company

[1000] Antitrust relates competition law on the abuse of dominant position. The laws are regulations that encourage competition by limiting the market power of any particular firm. This often involves ensuring that mergers and acquisitions don't overly concentrate market power or form monopolies, as well as breaking up firms that have become monopolies. Kelly, 2023

[1001] Toktok is a popular social media platform, with worldwide users. It is owned by a Chinese internet company called ByteDance

Chinese apps pose a threat to India's sovereignty and security."[1002] In 2024 the US signed into law legislation that could lead to Tiktok being banned, "US lawmakers and authorities are concerned that data from TikTok's 170 million US users could be accessed by the Chinese state under its national security laws."[1003]

12.8.3 Intellectual Property

AI can produce large volumes of text using platforms such as Google's ChatGPT. The question becomes who owns the AI generated content? What if the text output is a legal opinion, leads to a scientific innovation or is libellous? Who is the owner? ChatGPT's policy is that it does not own the output.[1004] This helped ChatGPT avoid becoming embroiled in plagiarism disputes. Derivative works,[1005] which produce valuable outputs entail several possible owners of inputs[1006] and requesting parties, have the potential for complex legal arguments in courts. In the US there are numerous cases before the courts but as yet undecided.[1007]

[1002] Pathi, 2024
[1003] Milmo, 2024
[1004] Caldwell, 2023
[1005] Derived from one or more already existing works
[1006] Such as data, photographs or song lyrics
[1007] Gil, *et al*, 2023

13 Glossary of Terms

Accountability: Accountability refers to the acceptance of responsibility for honest and ethical conduct toward others. In the corporate world, a company's accountability extends to its shareholders, employees, and the wider community in which it operates. In a wider sense, accountability implies a willingness to be judged on performance.

Actor / Agent: An individual, organisation or state carrying out a particular activity, with either good or malicious intentions. In AI, a software entity that acts autonomously on behalf of its designer.

Adjudication: In warfare (and wargaming), adjudication is the process of determining outcomes, usually by an objective human (although adjudication in some wargames occurs entirely within the simulation, with no human intervention). Adjudication without human intervention occurs in actual war situations.

Algorithm: A set of software instructions for computers to carry out defined tasks. Their uses can range from carrying out a simple calculation, to guiding social media users to items of 'interest' or advertisements. As time goes on algorithms are becoming more sophisticated.

AlgorithmWatch: A human rights organisation based in Berlin and Zurich. It fights to ensure that algorithms and Artificial Intelligence (AI) do not weaken justice, democracy and sustainability, but strengthens them.

Anonymised or Pseudonymised: Provides confidentiality and anonymity of personal data. It is the de-identification of identifiable personal data item values through the use of substitute values. Anonymised or Pseudonymised data can be linked and used for secondary purposes, such as trend analysis, peer comparison or medically related statistical analysis.

Application (App): An application is a type of software that allows specific tasks to be performed. Applications are more commonly referred to as "Apps."

Applications on desktop or laptop computers are called desktop apps, while those on mobile telephones are called mobile apps.

Artificial Capable Intelligence (ACI): Artificial Capable Intelligence is the point between AI and AGI. ACI can achieve a wide range of complex tasks but is still a long way from being fully general.

Artificial General Intelligence (AGI): Artificial General Intelligence is the point at which an AI can perform all human cognitive skills better than the smartest humans.

Artificial Intelligence (AI): Artificial Intelligence is the science of teaching machines to learn humanlike capabilities or the science and engineering of making intelligent machines. (Coined in 1955 by John McCarthy, Stanford University).

Artificial Intelligence Act: The EU's Artificial Intelligence Act became effective 13 March 2024. It sets out definitions and rules relating to AI.

Artificial Intelligence (AI) Impact Assessment: An expert method to identify the positive and negative impacts early in development, to safeguard AI's benefits while at the same time avoiding possible negative attributes. A related benefit is that it can be used for transparency purposes to inform stakeholders and regulators.

Artificial Neural Networks: A computer structure inspired by the biological brain, consisting of a large set of interconnected computational units ('neurons') that are connected in layers. Data passes between these units, as between neurons in a brain. Outputs of a previous layer are used as inputs for the next and there can be hundreds of units and layers. An artificial neural network with more than three layers is considered a deep learning algorithm. Examples of artificial neural networks include transformers (deep learning architectures) or generative adversarial networks.

Automated Influence: The use of AI to collect, integrate and analyse people's data to deliver targeted interventions that shape their behaviour. In this thesis automated Influence is not abbreviated to avoid confusion with AI.

Autonomy: The ability of a person, group, robot or AI to act independently and make their own decisions. AI does not as yet possess autonomy.

Autonomous Robot: Like a biological creature, a robot whose behaviour is self-generated, making use of sensory information.

Bias: A human trait that unfairly favours or discriminates certain people or ethnic origins. Programmers may intentionally or inadvertently impart their biases when designing and building AI systems.

Big Data: Data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse.

Bionics: Generally, the use of technological based body implants to replace or assist bodily functions. This is a combination of an artificial body part (prothesis) or a device, coupled with 'intelligence' to assist the functioning of the body part or device.

Bit: In classical computing bits are used to store and transmit information. A bit can have a value of 0 or a value of 1 to represent two states (on and off), it cannot be both simultaneously. Bits are arranged in sequence to make sense of and make decisions about the inputted data following a prearranged set of instructions.

Bug: See Software Bug.

Byte: In classical computing a byte is a unit of data that is eight binary digits (0 or 1; eight bits) long. A byte is the unit most computers use to represent a character such as a letter, number or typographic symbol.

ChatGPT: ChatGPT is an AI chatbot (simulates and processes written or spoken human conversation) with natural language processing (NLP) that allows a person to have human-like conversations to complete various tasks.

Chinese Room Argument (Turing experiment): A situation whereby a computer can behave as if it understands Chinese. Following the instructions within its programme, it can accept Chinese characters and produce an output with other Chinese characters. In doing so it passes the Turing Test. Creating a similar scenario with a human who has no knowledge of Chinese, a copy of the computer programme, pencils, papers and so on, the human could also achieve the same feat thereby passing the Turing Test.

Classical Computing: Classical computing is the traditional method of computer operations. It is simply a calculator using a sequence of bits, values of 0 and 1 to represent two states (on and off), to makes sense of and make decisions about the inputted data following a prearranged set of instructions. A bit can be 0 or 1, but not both simultaneously.

Conspiracy theory: An attempt to explain harmful or tragic events as the result of the actions of a small powerful group. Such explanations reject the accepted narrative and truth surrounding those events; the official version may be seen as further proof of the conspiracy.

Containment: The ability to monitor, curtail, control and potentially even close down technologies.

Cyber: An adjective or prefix relating to or characteristic of the culture of computers, information technology, and virtual reality.

Cyberattack: A cyberattack occurs when there is an unauthorized action against a computer infrastructure that compromises the confidentiality, integrity, or availability of its content. See also Hacking.

(The) Containment Problem: Technology's predisposition to diffuse widely in waves and to have emergent impacts that are impossible to predict or control, including negative and unforeseen consequences.

Data Protection Impact Assessment: a process used to identify and mitigate against any data protection related risks arising from the collection, storage, sharing and use of individuals' personal data or the data relating to an organisation.

Data Protection Regulation: A set of legal and other rules governing the collection, storage, use, disclosure, manipulation, analysis and other aspects of handling personal data. The permission of the data subject (the person) must be obtained for the collection and any use of their data. In the EU the General Data Protection Regulations (GDPR) govern personal data in all EU member states.

Deep Learning: A type of machine learning that uses artificial neural networks to recognise patterns in data and provide a suitable output, for example, a prediction. Deep learning is suitable for complex learning tasks and has improved AI capabilities in tasks such as voice and image recognition, object detection and autonomous driving.

Digital Divide: The gap between demographics and regions that have access to modern Information and Communications Technology (ICT) and those that do not or have restricted access. This technology can include internet connectivity and access to AI services. It typically exists between urban areas and rural areas; between the educated and the uneducated; between socioeconomic groups; globally, between the more and less industrially developing countries; and the Global North and Global South.

Digital Ethics: Digital ethics is the branch of ethics concerning the set of rules and moral guidelines that govern interpersonal behaviour between individuals and/or companies that is mediated by computer technology, either inside a company or, more broadly, in markets and society. Digital ethics includes transparency, being open to external audit, accountability, being subject to and complying with regulation. See also Ethics and Ethics Washing.

Dignity: The quality of being worthy of esteem, honour or respect. An unearned worth or status that all humans share equally.

Disinformation: False or misleading information which is deliberately intended to mislead. It is the intentional misstating of facts or the circulation of inaccurate information, for malicious, malign or propaganda purposes. See also “misinformation” which is inadvertent and snot malicious.

Disruptive Technology: An innovation that significantly alters the way societies, consumers, industries and businesses operate. Its attributes are recognisably superior, sweeping away the old systems or habits. It brings greater advantages that are immediately obvious, creating a smoother pathway for the new vision.

Drone Swarms / Formation Flying: Formation flying was originally observed in migratory birds. It is two or more aircraft traveling and manoeuvring together in a disciplined, synchronized, predetermined manner. In a tight formation, such as is typically seen at an air show. Drone swarms come as global warfare is entering a third age of drone warfare, defined by autonomy, saturation attacks, increased precision and range, and full-spectrum drone warfare across land, sea, and air.

Email: Electronic Mail, a similar concept to the traditional mail or postal service. Emails are sent and received via the internet.

Ethics: In AI, the issues that AI stakeholders must consider ensuring that the technology is developed and used responsibly. This means adopting and implementing systems that support a safe, secure, unbiased, protective of human dignity and environmentally friendly approach to artificial intelligence. See also Digital Ethics and Ethics Washing.

Ethics Washing: Exaggerating the extent to which an individual, organisation or government practice truly ethical behaviour. See also Digital Ethics and Ethics.

EU: European Union. There are 27 Member States in the EU: Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden.

Existential Risk: The risk of extinction or permanent adverse changes posed to humanity and its environment, affecting the future.

Expert Systems: AI Programmes (usually GOFAI) that can aid humans in making decisions in a particular area of expertise; some expert systems outperform even the best human experts but are not more intelligent.

Foundation Models: A machine learning model trained on a vast amount of data so that it can easily be adapted for a wide range of general tasks, including being able to generate outputs (generative AIs).

GDPR: General Data Protection Regulations, which are in force in all EU member states.

General Purpose AI Systems: AI systems that have a wide range of possible uses, both intended and unintended. They can be applied to many different tasks in various fields, often without substantial modification and fine-tuning. They are increasingly useful commercially due to growing amounts of computational resources available and innovative methods to use them.

Generative: Capable of production or reproduction.

Generative AI: An AI model that generates text, images, audio, video or other media in response to user prompts. It uses machine learning techniques to create new data that has similar characteristics to the data it was trained on. Generative AI applications include chatbots, photo and video filters, and virtual assistants.

Genetic Algorithm: A machine learning method to finding solutions to certain types of problems, akin to the biological process of artificial selection. Candidate solutions are encoded in a 'genome', an initial population of solutions is created and evaluated for its 'fitness' to solve the problem to hand. The fitter candidates are selected and the process repeated until after several generations the population contains mainly highly fit solutions.

Global North / Global South: Global South refers broadly to the regions of Latin America, Asia, Africa, and Oceania (excluding Israel, Japan, South Korea, Australia and New Zealand). It is one of a family of terms, including "Third World" and "Periphery," that denote regions outside Europe and North

America, mostly low-income and often politically or culturally marginalised. The Global North is all countries not in the Global South.

Global South: See Global North.

GOFAI – Good Old-Fashioned AI: AI based on symbolic rules. It is also known as classical, traditional or symbolic AI.

GPT: Generative Pre-training Transformer (deep learning architecture).

Groups through Imitation and Conformity: The process whereby people change their beliefs, attitudes, actions, or perceptions to more closely match those held by groups to which they belong or want to belong or by groups whose approval they desire.

Hacking: The use of unconventional or illicit means to gain unauthorised access to a digital device, computer system or computer network. See also Cyberattack. There are numerous forms of hacking attacks which include:

DOS or DDOS, Denial of Service and Distributed Denial-of-Service with the aim of disabling a website or network, which create vulnerability for another type of attack;

Phishing Attack, when a malicious actor sends emails that seem to be coming from trusted, legitimate sources in an attempt to grab sensitive information from the target;

Ransomware Attack, the victim's system and data is held hostage until they agree to pay a ransom to the attacker;

***Domain Name System* (DNS) Spoofing**, a hacker alters DNS records to send traffic to a fake or "spoofed" website. Once on the fraudulent site, the victim may enter sensitive information that can be used or sold by the hacker.

Hard Law: Hard Law is legislation or individual laws enacted by parliaments. Hard laws take much time to formulate and pass through the various

consideration stages before final enactment in the parliament. Soft Laws are put in place much speedier.

Human Rights: The set of rights that each human individually possess. The United Nations' 1948 "Universal Declaration of Human Rights", states "All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood."

ICT – Information and Communications Technology: collective term for all technical means used to handle information and aid communications.

Intelligence: A process, or an innate capacity to use information to respond to ever-changing requirements. It is a capacity to acquire, adapt, modify, extend and use information to solve problems.

Intelligent Prosthesis: See Bionics.

Internet: An electronic communications network that connects mobile telephones' applications, devices, computer networks and organisations' computer facilities around the world. (See also Internet of Things).

Internet of Things: Items (things) connected to the internet using network addresses and fitted with sensors to collect data. In this context, a 'thing' is any electronic device (domestic or otherwise) connected to the internet. Examples include a domestic fitness tracker (e.g. Fitbit), door-bell camera, a self-service checkout, a sensor in an industrial plant or an internet telephone in a government building.

IT – Information technology: collective term for all computer, internet, data storage and related equipment and technologies.

LaDMA: Language Model for Dialogue Application.

Large Language Model (LLM): A type of Foundation Model that is trained on vast amounts of text to carry out natural language processing tasks. During training phases, Large Language Models learn parameters from factors such

as the model size and training data. Large Language Models often have at least one billion or more parameters.

Law: Laws are enacted and approved by national or local parliaments. A Bill proposing a law may be brought to the parliament for ultimate approval by a government member (a government bill) or another parliamentarian (private member's bill). Each parliament has a set process, (a number of stages for consideration and amendment) before being voted on by the parliament and approved. The final approved legislation is signed into law by the head of state (for example a president) or receives other formal approval (a royal assent in some jurisdictions).

Legislation: Legislation is a set of laws that are directly related to each other, for a particular area.

Logic: A precisely defined formal system or set of principles used for defining the inferential (deductive) relations of reasoning. If coupled with semantic (understanding language) interpretation, logical systems can be used to represent knowledge.

Machine Consciousness: A subfield of AI that attempts to define the architectural requirements and conditions for machines to be conscious.

Machine Learning: A type of AI that allows a system to learn and improve from examples, without all its instructions being explicitly programmed.

Manipulation: To manipulate is to handle or control in a skilful manner. For the purposes of this thesis, manipulation is subterfuge designed to influence or control a person or group of persons, in an underhanded manner. Manipulation facilitates the aims of AI companies or organisations. It also facilitates those wishing to set a narrative for their purposes such as conspiracy groups or the spread of propaganda. Methods used to distort individuals' perception of reality may include seduction, suggestion, persuasion and blackmail to induce submission.

Mega-Analysis: An analytical process by which raw data is pooled across studies.

Meta-Analysis: The statistical combination of results from two or more separate research studies that address a similar research question to generate an average result.

Misinformation: False or misleading information which is circulated in error. It is the unintentional quoting of inaccurate information. There is no malicious or malign intent. See also "Disinformation", which is intentional and malicious.

Moderation: The assessment of online content to ensure it is acceptable, safe and honest, while protecting free speech. Content not meeting these standards is removed. The removal process can be conducted by humans, which is slow and tedious work. Alternatively, the process can be automated, controlled by algorithms.

Moore's Law: Expresses the rate of growth in computing technology. The ability to squeeze twice as many transistors on an integrated circuit board every 24 months. Given that the electrons have less distance to travel, the circuits run twice as fast, providing an overall quadrupling of computational power. (Gordon Moore was the co-founder of the Intel Corporation, the microchip manufacturer.)

Neural Network: In the human brain, a population of biological neurons chemically connected to each other by synapses, through which the brain functions. See also Artificial Neural Networks.

Neurodevelopmental Disorders: The World Health Organisation criteria define neurodevelopmental disorders as "behavioural and cognitive disorders that arise during the developmental period that involve significant difficulties in the acquisition and execution of specific intellectual, motor, or social functions."

Neurodiverse: "Neurodiversity" is a popular term that is used to describe differences in the way individuals' brains work. It is a combination of traits that are seen as both strengths and challenges.

OECD: Organisation for Economic Co-operation and Development.

Online: Online, in telecommunications and technology terms, indicates a state of connectivity. Being online refers to being connected to, served by, or available through a system such as a computer or telecommunications system like the internet. When online individuals can send and receive e:mails, social media messages and view websites.

Opacity / Opaqueness: The direct opposite of transparency. For AI it would be the withholding of crucial information by companies and governments on how AI operates, its risks and infringements on human rights.

Ontogeny: Principle of Ontogeny Non-Discrimination: If two beings have the same functionality and the same consciousness experience, and differ only in how they came into existence, then they have the same moral status. [Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322-23].

Qualia: Subjective conscious experiences (*singular quale*). Sensations felt by living beings; for example, touch, taste, pain, vision and so on. Raw sensations or felt experiences, such as the sensation of the colour red or of pain.

Quantum Computing: This form of computing, still in its relative infancy, originates out of quantum mechanics, the behaviour of nature at the scale of atoms and subatomic particles. It employs quantum physics to solve problems that standard computers are unable to answer. They rely of 'Qubits' (quantum bits) that have multiple possible states; (i) 0 or (ii) 1 or (iii) 0 and 1 simultaneously or (iv) all intermediate states. This is known as Superposition, whereby particles can exist in multiple states at the same time and which allows quantum computers to look at many different variables at once.

Pacing Problem: Laws and regulations enacted by governments become inadequate (or lag behind) in the face of advances, unable to reflect the new reality. This is a general statement that also applies to AI advances.

Personal Data: Any datum (piece of information) relating to a living person. For data protection purposes personal data is two or more pieces of information that

can identify an individual. Data includes, name, address, sex, date of birth, social security number, marital status, health information, sexual orientation and financial information.

Poverty of Stimulus: Relates to human learning. It is the assertion that natural language grammar is unlearnable given the relatively limited data available to children learning a language and therefore that this knowledge is supplemented with some sort of innate linguistic capacity.

Productivity Riddle: A productivity paradox of the perceived discrepancy between measures of investment in information technology and measures of output at the national levels. The concept is attributed to the economist Robert Solow, referring to his comment that "You can see the computer age everywhere but not in the productivity statistics."

Pseudonymised or Anonymised: See anonymised.

Regulations: Regulation refers to the promulgation of targeted rules, typically accompanied by some authoritative mechanism for monitoring and enforcing compliance. Regulations are generally issued by independent regulatory authorities and are enacted much more quickly than laws. In general regulators consult on proposed regulations, with industry players and stakeholders, before implementation. The regulators carry statutory authority, are independent of governments and are established to focus on particular industries or industry segments. Regulators may impose fines or other sanctions in their own right or request a court to do so.

Responsible AI: Often refers to the practice of designing, developing, and deploying AI with certain values, such as being trustworthy, ethical, transparent, explainable, fair, robust and upholding privacy rights and human dignity.

Robot: An artificial physical device carrying out programmed instructions and interacting with its environment.

Sapience: a set of capacities associated with higher intelligence, such as self-awareness and being a reason-responsive agent.

SDGs: See Sustainable Development Goals.

Seasonal Climate Forecasting (SCF): SCF is used to predict severe weather, such as droughts and floods, in order to provide policymakers and farmers with the means to address problems, in an anticipatory rather than a reactive manner.

Semantic Interpretation: The process of mapping a syntactically analysed text of natural language to a representation of its meaning.

Sentience: the capacity for phenomenal experience or qualia, such as the capacity to feel pain and suffer.

Social Contagion: The spread of generally malign ideas, attitudes, beliefs or behaviours. Such contagion can add to support for conspiracy theories.

Social Credit System (China): With China's social credit system, the government collects large volumes of data, to honour good deeds and punish bad deeds, thereby rewarding or disciplining citizens. The data is collected from various sources, including its own registers and private service providers. An extensive network of surveillance cameras also collects data and automatic facial recognition which can be used to identify traffic offenders and publicly pillory them. Citizens obtain scores to either reward or penalise them.

Social Disruption / Socially Disruptive Technology: Social disruption refers to events or circumstances that cause significant changes in societal norms, behaviours, and structures. These disruptions can be caused by various factors such as wars, conflicts, political upheaval, economic crises, technological advancements or natural disasters.

Social Media: Social media is a collective term for websites and applications that focus on communication, community-based input, interaction, content-sharing and collaboration. People use social media to stay in touch and interact with friends, family, business colleagues and various communities.

Soft Law: Soft law refers to measures that are not directly legally enforceable but that can, nonetheless, sometimes create substantive obligations. Examples

include guidelines, sets of principles, codes of conduct, private standards, and partnership programs. In many industries, hard law and soft law exist alongside each other. Soft Law is much quicker to implement and can be easily updated. Guidelines and Code of Practice introduced by regulators carry strong regulatory authority.

Software: The written (or coded) instructions, written by specialists, for computers to carry out the desired purpose or task. (S*ee also algorithms*).

Software Bug: a defect in software that causes the computer to cease functioning. A bug may cause the computer to carry out an unexpected or undesired task.

Stakeholder: A stakeholder is any party that has an interest in or needs knowledge of an organisation, who can either affect or be affected by its operations. The primary stakeholders typically are investors, employees, customers, suppliers, governments, media, interest groups and the wider public.

Strong AI: An aspect of AI devoted to building "persons", *machines with minds*, in the full and literal sense. See also Weak AI.

Structural Unemployment: Structural unemployment is a longer-lasting form of unemployment, caused by fundamental shifts in an economy and exacerbated by extraneous factors such as technology, competition, and government policy. Structural unemployment occurs when workers lack the requisite job skills, language or live too far from regions where jobs are available and cannot move closer. Jobs are available, but there is a serious mismatch between what companies need and what workers can offer. The closing of coal mines and steel works are examples of those skilled in the industries becoming structurally unemployed.

Substrate: Principle of Substrate Non-Discrimination: If two beings have the same functionality and the same conscious experience, and differ only in the substrate of their implementation, then they have moral status. *Substrate refers to the bases for human (biological) and artificial (technological)*

intelligence. [Bostrom and Yudkowsky, in Frankish and Ramsey, 2014, 322-23].

Superintelligence: An as yet theoretical form of AI that has intelligence greater than humans and exceeds their cognitive performance in most domains.

Sustainable Development Goals (SDG): A set of 17 goals developed by the United Nations. They are an urgent call for action by all countries - developed and developing - in a global partnership. They recognise that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests. AI can play a central role to assist implementation of the goals.

Technology: The application of scientific knowledge (in the broadest sense) to produce tools or practical outcomes for humans.

Too Big to Fail: Too big to fail refers to an entity so important to a financial system that a government would not allow it to go bankrupt due to the seriousness of the economic repercussions. For example, the US 2008 Emergency Economic Stabilization Act provided bailout funds for Wall Street banks and US automakers, the financial health of which were considered essential to the United States economy.

Too Big to Jail: Following the financial collapse of 2008, the phrase "too big to jail" was used to describe the theory that certain financial institutions and individuals, even if they engage in criminal conduct, are immune from criminal prosecution. The Attorney General in the US Justice Department asserted that no individual or company, no matter how large or how profitable, is above the law.

Transparency: In the AI context, transparency refers to companies and governments being open and honest about the technology and the implications of its use. In the case of algorithms demonstrating or revealing the inner workings so that fairness, bias, safety, reliability, liability and potential existential risks

can be independently examined. This is particularly important in high-risk scenarios such as healthcare, the handling of large databases or how AI makes critical decisions.

Turing: Alan Turing was an English mathematician, computer scientist, logician, cryptanalyst, philosopher and theoretical biologist. He was influential in the development of theoretical computer science, formalising concepts of algorithm and computation with the Turing machine (considered a model of a general-purpose computer). Viewed as the father of theoretical computer science, he died in 1954.

Turing Test: An experiment in which two participants, a human and a computer, are interrogated by a second human from whom they are hidden. The interrogator's challenge is to say which is the human and the computer. After a series of questions and answers if the interrogator judges no better than 50-50 which is the computer, then the computer is said to have passed the test. Such a computer, according to Turing, should be declared a 'thinking' machine.

Wargaming: A simulation of a military operation, by whatever means, using specific rules, data, methods and procedures. (Also; A scenario-based warfare model in which the outcome and sequence of events affect, and are affected by, the decisions made by the players).

Weak AI: A bodiless computer or 'robots able to operate in the physical environment in a way that is indistinguishable from the behaviours, manifested by embodied human persons navigating the physical world'. It acts intelligently without knowing whether it actually is intelligent, (see Strong AI).

Website: A set of electronic pages providing information, services and facilities. Websites are 'hosted' on the internet by an individual, company, organisation or government body.

UK: The United Kingdom of Great Britain and Northern Ireland. England, Scotland, Wales (collectively Great Britain) and Northern Ireland

US / USA: The United States of America.

14 Bibliography

The bibliography lists all items cited and referenced directly in the body of the thesis, as well as those consulted during research.

Abraham, Yuval, *'Lavender': The AI machine directing Israel's bombing spree in Gaza*, (+972 Magazine, 03 April 2024), <https://www.972mag.com/lavender-ai-israeli-army-gaza/> [accessed 15 April 2024]

Adams, Tim, *The Handover by David Runciman review – is the future out of our control?,* (London: The Guardian, 10 September 2023) <https://www.theguardian.com/books/2023/sep/10/the-handover-how-we-gave-control-of-our-lives-to-corporations-states-ais-by-david-runciman-review-is-the-future-out-of-our-control> [accessed 22 January 2023]

AI Safety Summit, *AI Safety Summit Hosted by The UK 1st and 2nd November 2023 at Bletchley Park*, (AI Safety Summit website, 2023) <https://www.aisafetysummit.gov.uk/> [accessed 27 March 2024]

Alan Turing Institute, *Data science and AI glossary*, (Alan Turing Institute website, 2024) <https://www.turing.ac.uk/news/data-science-and-ai-glossary> [accessed 22 February 2024]

Alibabagroup.com, *About Us,* (Alibabagroup.com website, 2024), <https://www.alibabagroup.com/en-US/about-alibaba> [accessed 08 February 2024]

American Cancer Society, *Mastectomy*, (American Cancer Society website, 20 December 2023) <https://www.cancer.org/cancer/types/breast-cancer/treatment/surgery-for-breast-cancer/mastectomy.html#:~:text=Mastectomy%20is%20breast%20cancer%20surgery,conserving%20surgery%20for%20personal%20reasons> [accessed 11 April 2024]

American Psychiatric Association, *What is Gender Dysphoria?*, (American Psychiatric Association website, August 2022) <https://www.psychiatry.org/patients-families/gender-dysphoria/what-is-gender-dysphoria> [accessed 12 April 2024]

American Psychological Association, *Science of Psychology*, (American Psychological Association, 2013), <https://www.apa.org/education-career/guide/science> [accessed 13 March 2024]

American Psychological Association, *Intelligence (Definition)*, (American Psychological Association, 2018), <https://dictionary.apa.org/intelligence> [accessed 31 January 2024]

Amnesty International, *EU: Artificial Intelligence rulebook fails to stop proliferation of abusive technologies*, (Amnesty International, 13 March 2024) <https://www.amnesty.org/en/latest/news/2024/03/eu-artificial-intelligence-

rulebook-fails-to-stop-proliferation-of-abusive-technologies/> [accessed 02 May 2024]

Amnesty International, *Amnesty International Report 2022/23, The State of the World's Human Rights*, (Amnesty International, 28 March 2023) <https://www.amnesty.org/en/annual-report-archive/#h-2022-23-amnesty-international-annual-report> [accessed 22 February 2024]

Andrews, Edmund L., *The Fast and the Incurious: How High-Speed Stock Trading Pushes Out Smart Money*, (Insights by Stanford Business, Stanford Business School, August 2021) <https://www.gsb.stanford.edu/insights/fast-incurious-how-high-speed-stock-trading-pushes-out-smart-money> [accessed 02 February 2024]

AnyControl, *5 most dangerous social media apps and their possible risks*, (AnyControl website, No Date), <https://anycontrol.app/blog/post/most-dangerous-social-media-apps> [accessed 07 February 2024]

Appel, Gil; Neelbauer, Juliana and Schweidel, David A., *Generative AI Has an Intellectual Property Problem,* (Harvard Business Review, 07 April 2023) <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem> [accessed 18 March 2024]

AppliedBehaviorAnalysisEDU.org, No Date, *5 Mind-Blowing Things Kim Peek Could Do That You Can't*, (AppliedBehaviorAnalysisEDU.org , Wiley University Services USA, No Date), <https://www.appliedbehavioranalysisedu.org/5-mind-blowing-things-kim-peek-could-do-that-you-cant/> [accessed 07 February 2024]

AppliedBehaviorAnalysisEDU.org, No Date, *Autistic Genius: Is Autism Associated with Higher Intelligence?*, (AppliedBehaviorAnalysisEDU.org , Wiley University Services USA, No Date), <https://www.appliedbehavioranalysisedu.org/is-autism-associated-with-higher-intelligence/> [accessed 07 February 2024]

Ashgar, Hadi and Züger, Theresa, *AI for the public. How public interest theory shifts the discourse on AI*, (AI & society, vol. 38, no. 2, pp. 815-828, (USA: Springer Nature, 03 June 2023), <https://link-springer-com.may.idm.oclc.org/article/10.1007/s00146-022-01480-5#Sec13> [accessed 22 January 2024]

Associated Press, *The Latest | Four more Israeli hostages declared dead, and Palestinians seek to join genocide case*, (Associated Press website, 04m June 2024) <https://apnews.com/article/israel-hamas-rafah-latest-06-03-2024-4531e5bc3af4b808352a48f5cbe68f60> [accessed 04 June 2024]

Augustyn, Adam, *Prosthesis*, (Encyclopaedia Britannica, 11 February 2024) <https://www.britannica.com/science/prosthesis> [accessed 27 March 2024]

Augustyn, Adam, *Technology*, (Encyclopaedia Britannica, 20 February 2024) <https://www.britannica.com/technology/technology> [accessed 08 March 2024]

Austin, Robert D. and Pisano, Gary P., *Neurodiversity as a Competitive Advantage*, (Marginalised Groups, Harvard Business Review Magazine website, May-June 2017)

<https://hbr.org/2017/05/neurodiversity-as-a-competitive-advantage> [accessed 29 February 2024]

Australian eSafety Commissioner, *Online Hate*, (eSafety Commissioner, Australian Communications and Media Authority (ACMA), 26 June 2023), <https://www.esafety.gov.au/young-people/online-hate#:~:text=Online%20hate%20can%20be%20defined,characteristics%2C%20at%20the%20same%20time> [accessed 07 February 2024]

Awan, Abid Ali, *What is GPT-4 and Why Does it Matter?*, (datacamp website, nd) <https://www.datacamp.com/blog/what-we-know-gpt4?dc_referrer=https%3A%2F%2Fwww.google.ie%2F> [accessed 14 June 2024]

Bauer, Pat, *Rain Man film by Levinson (1988)*, (Encyclopaedia Britannica, 04 January 2024), <https://www.britannica.com/topic/Rain-Man> [accessed 07 February 2024]

Back, Hilary, *8 Common Threats to Kids on Social Media*, (Experian's Diversity, Equity and Inclusion, Experian, 07 June 2021), <https://www.experian.com/blogs/ask-experian/common-threats-to-kids-on-social-media/> [accessed 07 February 2024]

Backlinko, *The 11 Best ChatGPT Alternatives*, (Backlinko website, 26 March 2024) <https://backlinko.com/chatgpt-alternatives> [accessed 27 March 2024]

BBC East Midlands, *Care home residents react to robot pets*, (BBC East Midlands News, BBC website, 21 April 2024) < https://www.bbc.com/news/uk-england-derbyshire-68827911 > [accessed 24 April 2024]

BBC News, *What is Israel's Iron Dome missile system and how does it work?*, (BBC News Website, 06 November 2023) <https://www.bbc.com/news/world-middle-east-20385306> [accessed 24 April 2024]

Beaney, Michael, *Imagination: the missing mystery of philosophy*, (The Open University website, The Open University 2016) <https://www.open.edu/openlearn/history-the-arts/philosophy/imagination-the-missing-mystery-philosophy/content-section-0/?printable=1#:~:text=Reveal%20discussion-,Discussion,%2C%20existence%20or%20non%2Dexistence> [accessed 13 February 2024]

Bengio, Yoshua and Musk, Elon, *et al*; *Pause Giant AI Experiments: An Open Letter*, (*An Open Letter signed by numerous Academics and Technology Industry CEOs*, Future of Life Institute, 22 March 2023) <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> [accessed 08 March 2024]

Benn, C. and Lazar, S, *What's Wrong with Automated Influence*, (Cambridge: Canadian Journal of Philosophy. Cambridge University Press, 52(1), pp. 125–148. doi: 10.1017/can.2021.23, 2022) <https://www-cambridge-org.may.idm.oclc.org/core/journals/canadian-journal-of-philosophy/article/whats-wrong-with-automated-influence/98F5E24BEADE585050B773D2CBEB1F39> [accessed 22 January 2024]

Bertelsmann Stiftung, *China's Social Credit System: Fact Check*, (The Magazine of the Bertelsmann Stiftung, Bertelsmann Stiftung website, 15 February 2022), <https://www.bertelsmann-stiftung.de/fileadmin/files/aam/Asia-Book_A_03_China_Social_Credit_System.pdf> [accessed 14 March 2024]

Bhagwat, Ash, *The Impact of 9/11 on Freedom of Expression in the United States*, (Verfassungsblog website – a journalistic and academic forum, 31 January 2022) <https://verfassungsblog.de/os4-us-foe/> [accessed 21 February 2024]

Birbeck University London, *65% of Neurodivergent Employees Fear Discrimination Despite UK Businesses Promoting Neurodiverse Workplaces*, (Birbeck University London website, 16 March 2023) <https://www.bbk.ac.uk/news/neurodivergent-employees-fear-discrimination-despite-uk-businesses-promoting-neurodiverse-workplaces> [accessed 29 February 2024]

Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., Newman, P., Parry, V., Pegman, G., Rodden, T., Sorrell, T., Wallis, M., Whitby, B. & Winfield, A. 2017, *Principles of robotics: regulating robots in the real world*, (Connection science, vol. 29, no. 2, pp. 124-129. 2017) <https://www-tandfonline-com.may.idm.oclc.org/doi/full/10.1080/09540091.2016.1271400> [accessed 29 February 2024]

Bostrom, Nick, *Superintelligence: paths, dangers, strategies*, First edn., (Oxford: Oxford University Press, Oxford, 2014)

Botelho, Bridget and Bigelow, J., *Big Data – Definition*, (TechTarget website, January 2022) <https://www.techtarget.com/searchdatamanagement/definition/big-data> [accessed 19 February 2024]

Barzashka, Ivanka, *Wargames and AI: A dangerous mix that needs ethical oversight*, (The Bulletin of Atomic Sciences, website, 04 December 2023) <https://thebulletin.org/2023/12/wargames-and-ai-a-dangerous-mix-that-needs-ethical-oversight/> [accessed 03 July 2024]

Bringsjord, Selmer and Naveen Sundar Govindarajulu, Artificial Intelligence, California: The Stanford Encyclopaedia of Philosophy (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/entries/artificial-intelligence/> [accessed 17 January 2024]

The Britannica Dictionary (a), *technology*, (The Britannica Dictionary, Encyclopaedia Britannica, 25 June 2024) <https://www.britannica.com/dictionary/technology> [accessed 28 June 2024]

The Britannica Dictionary (b), *self*, (The Britannica Dictionary, Encyclopaedia Britannica, nd) <https://www.britannica.com/dictionary/technology> [accessed 28 June 2024]

Britannica, The Editors of Encyclopaedia, *Boolean Algebra,* (Encyclopaedia Britannica, 23 February 2024) <https://www.britannica.com/topic/Boolean-algebra> [accessed 03 April 2024]

Britannica, The Editors of Encyclopaedia, *Delusion,* (Encyclopaedia Britannica, 11 January 2024) <https://www.britannica.com/science/delusion> [accessed 11 March 2024]

Britannica, The Editors of Encyclopaedia, *Deontological ethics,* (Encyclopaedia Britannica, 16 February 2024) <https://www.britannica.com/topic/deontological-ethics> [accessed 14 March 2024]

Britannica, The Editors of Encyclopaedia, *Euthanasia,* (Encyclopaedia Britannica, 11 March 2024) <https://www.britannica.com/topic/euthanasia> [accessed 05 April 2024]

Britannica, The Editors of Encyclopaedia, *Strategic Defence Initiative*, (Encyclopaedia Britannica, 04 Jun. 2024) <https://www.britannica.com/topic/Strategic-Defense-Initiative> [accessed 21 June 2024]

Bunz, Mercedes and Janciute, Laima, *Artificial Intelligence and The Internet of Things – UK Policy Opportunities and Challenges,* (Westminster: University of Westminster Press, 2018) <https://library.oapen.org/bitstream/handle/20.500.12657/29709/UWP-025-bunz.pdf?sequence=1> [accessed 28 September 2023]

Burrows, Emma and Litvinova, Dasha, *Protests, poisoning and prison: The life and death of Russian opposition leader Alexei Navalny*, (The Associated Press website, 16 February 2024) <https://apnews.com/article/russia-navalny-life-timeline-0722708e19e51b10699b2cc73ece0bae> [accessed 21 February 2024]

Caldwell, Keegan, *AI And Intellectual Property: Who Owns It, And What Does This Mean For The Future?*, (Forbes Media LLC., 31 October 2023) <https://www.forbes.com/sites/forbesbusinesscouncil/2023/10/31/ai-and-intellectual-property-who-owns-it-and-what-does-this-mean-for-the-future/> [accessed 18 March 2024]

Cambridge Online Dictionary, *robot*, (Cambridge Online Dictionary, nd) <https://dictionary.cambridge.org/dictionary/english/robot> [accessed 06 June 2024]

Cambridge Online Dictionary, *Sapient*, (Cambridge Online Dictionary, nd) <https://dictionary.cambridge.org/dictionary/english/sapient> [accessed 17 June 2024]

Campos, Andre Santos, (2013), *Responsibility and Justice in Aristotle's Non-Voluntary and Mixed Actions*, (Journal of Ancient Philosophy, 7(2), pp. 100-121. doi: 10.11606/issn.1981-9471.v7i2p100-121.

Canales, Katie and Mok, Aaron, *China's 'social credit' system ranks citizens and punishes them with throttled internet speeds and flight bans if the Communist Party deems them untrustworthy*, (Business Insider website, 28 November 2022) <https://www.businessinsider.com/china-social-credit-system-punishments-and-rewards-explained-2018-4?r=US&IR=T#:~:text=China%27s%20%27social%20credit%27%20system%20ranks,

Communist%20Party%20deems%20them%20untrustworthy> [accessed 14 March 2024]

Capitol Technology University, *AI at War: The Next Revolution for Military and Defence*, (Capitol Technology University, website, 04v August 2023), <https://www.captechu.edu/blog/ai-war-next-revolution-military-and-defense> [accessed 08 July 2024]

Dr. Cass, Hilary, *Independent review of gender identity services for children and young people*, (The Cass Review website, April 2024) <https://cass.independent-review.uk/home/publications/final-report/> [accessed 12 April 2024]

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E. and Tavares, M. M., *Gen-AI: Artificial Intelligence and the Future of Work,* (IMF Staff Discussion Notes, IMF, January 2024) <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379?cid=bl-com-SDNEA2024001> [accessed 28 March 2024]

Centre for Disease Control and Prevention, *The BRCA1 and BRCA2 Genes*, (Centre for Disease Control and Prevention website, 20 June 2023) <https://www.cdc.gov/genomics/disease/breast_ovarian_cancer/genes_hboc.htm#:~:text=The%20genes%20most%20commonly%20affected,the%20BRCA1%20and%20BRCA2%20genes> [accessed 11 April 2024]

The Centre for the Study of Existential Risk, (The Centre for the Study of Existential Risk, website, 2018-2021) <https://www.cser.ac.uk/research/paradigms-AGI/> [accessed 02 April 2024]

Centric Medical Health (Ireland), *Neurodiversity: What It Means and Why It Matters,* (Centric Medical Health (Ireland) website, 10 October 2023) <https://mentalhealth.ie/blog/neurodiversity-what-it-means-and-why-it-matters> [accessed 29 February 2024]

CFI Education Inc., *Flash Crashes; Sudden plunge in the price of bonds, stocks, or commodities followed by a recovery*, (Corporate Finance Institute, CFI, website 2024) <https://corporatefinanceinstitute.com/about-cfi/> [accessed 10 June 2024]

Chang, Kenneth, *Quantum Computing Advance Begins New Era, IBM Says*, (The New York Times, 19 June 2023) <https://www.nytimes.com/2023/06/14/science/ibm-quantum-computing.html> [accessed 13 February 2024]

Chasse, J. Dennis, *Coercion, Freedom, and Democracy in Hayek, Dewey, and Commons,* (Journal of Economic Issues, 56:1, 18-39, DOI, 25 March 2022), <https://www-tandfonline-com.may.idm.oclc.org/doi/full/10.1080/00213624.2022.2008751> [accessed 26 January 2024]

Chenu, Marie-Dominique, *St. Thomas Aquinas*, (Encyclopaedia Britannica, 29 Apr. 2024) <https://www.britannica.com/biography/Saint-Thomas-Aquinas> [accessed 24 June 2024]

Cherry, Kendra, *What Is a Genius IQ Score?,* (Verywell Mind, Dotdash Media Inc., 11March 2023), <https://www.verywellmind.com/what-is-a-genius-iq-score-2795585> [accessed 08 February 2024]

Cherry, Kendra, *What Is an IQ Test?,* (Verywell Mind, Dotdash Media Inc., November 2022), <https://www.verywellmind.com/how-are-scores-on-iq-tests-calculated-2795584> [accessed 31 January 2024]

CIA, *About CIA*, (CIA Website, nd) <https://www.cia.gov/about/> [accessed 24 April 2024]

Cleveland Clinic, *Ebola Virus Disease*, (Cleveland Clinic, website, 22 November 2023) <https://my.clevelandclinic.org/health/diseases/15606-ebola-virus-disease> [accessed 19 April 2024]

Cleveland Clinic, *Neurodivergent*, (Cleveland Clinic website, 02 June 2022), <https://my.clevelandclinic.org/health/symptoms/23154-neurodivergent> [accessed 29 February 2024]

The Club of Rome, *History*, (The Club of Rome, website, nd), <https://www.clubofrome.org/history/> [accessed 08 July 2024]

The College of Physicians of Philadelphia, *Misconceptions About Vaccines, History of Anti-Vaccination Movements*, (History of Vaccines, The College of Physicians of Philadelphia website, nd) <https://historyofvaccines.org/vaccines-101/misconceptions-about-vaccines/history-anti-vaccination-movements> [accessed 03 July 2024]

CompaniesMarketcap.com, *Top publicly traded tech companies by revenue*, (CompaniesMarketcap.com website, Q3 2023), <https://companiesmarketcap.com/tech/largest-tech-companies-by-revenue/> [accessed 08 February 2024] *Note: data used covers period Q4 2022 to Q3 2023*.

Congress Research Service, *Constitution of the United States: Analysis and Interpretation*, (The Congressional Research Service, constitution.congress.gov website, No Date) <https://constitution.congress.gov/about/constitution-annotated/> [accessed 21 February 2024]

Conover, Emily, *A new supercomputer just set a world record for speed*, (Science News Explores website, 11 July 2022) <https://www.snexplores.org/article/supercomputer-world-record-speed-computing-calculations-exascale-frontier> [accessed 13 February 2024]

Copeland, Jack, *Artificial intelligence: a philosophical introduction*, (Blackwell, Oxford [England], 1993)

Crawford, K, *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*, (New Haven, USA: Yale University Press, New Haven, 2021)

Cressman, Colleen and Brown, Allison, Introduction to Philosophy: Ethics, (The Open Education Network, University of Minnesota, 2019),

<https://open.umn.edu/opentextbooks/textbooks/797> [accessed 09 February 2024]

Cruz-Benito, Juan, *AI in Education*, (Switzerland: MDPI Switzerland, 2022) <https://mdpi-res.com/bookfiles/book/5611/AI_in_Education.pdf?v=1700465327> [accessed 28 September 2023]

Curado, M, & Gouveia, SS (eds), *Automata's Inner Movie: Science and Philosophy of Mind*, (Vernon Art and Science Inc., Wilmington, DE., 2019) Available from: ProQuest Ebook Central. [16 February 2024].

Dados, Nour and Connell, Raewyn, *The global south*, (Jargon Key Concepts in Social Research, American Sociological Association, 2012) <https://journals.sagepub.com/doi/pdf/10.1177/1536504212436479> [accessed 12 April 2024]

Debes, Remy, *Dignity*, (The Stanford Encyclopaedia of Philosophy (Spring 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.)) <https://plato.stanford.edu/entries/dignity/> [accessed 22 March 2024]

De Cosmo, Leonardo, *Google Engineer Claims AI Chatbot Is Sentient: Why That Matters*, (Scientific American, 12 July 2022) <https://www.scientificamerican.com/article/google-engineer-claims-ai-chatbot-is-sentient-why-that-matters/> [accessed 02 July 2024]

DeLapp, Kevin M., *Metaethics*, (Internet Encyclopaedia of Philosophy, nd) <https://iep.utm.edu/metaethi/#:~:text=Metaethics%20is%20a%20branch%20of,on%20what%20morality%20itself%20is> [accessed 02 July 2024]

DePauw University, *What is Philosophy?*, (DePauw University, 2024), <https://www.depauw.edu/academics/departments-programs/philosophy/what-is-philosophy/> [accessed 31 January 2024]

Dictionart.com, *technology*, (Dictionary.com website, nd) <https://www.dictionary.com/browse/technology> [accessed 28 June 2024]

Diederich, Joachim, *Philosophical Aspects of a Resistance to Artificial Intelligence*, (Australia: School of Information Technology and Electrical Engineering, The University of Queensland, 2021) < file:///C:/Users/35387/Downloads/Philosophical%20Aspects%20of%20a%20Resistance%20to%20AI.pdf> [accessed 18 January 2024]

DigitalStakeout, *What is a Social Media Threat?*, (DigitalStakeout Inc. website, No Date), <https://www.digitalstakeout.com/post/what-is-a-social-media-threat> [accessed 07 February 2024]

Dimock, Michael, *Defining generations: Where Millennials end and Generation Z begins*, (Pew Research Centre, website, 17 January 2019) <https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/> [accessed 03 July 2024]

Dong, Yanyan; Hou, Jie; Zhang, Ning and Zhang, Maocong, *Research on How Human Intelligence, Consciousness, and Cognitive Computing Affect the Development of Artificial Intelligence* (Complexity, vol. 2020, Article ID 1680845, 10 pages, Hindawi, London, 2020) <https://www.hindawi.com/journals/complexity/2020/1680845/> [accessed 26 April 2024]

Dovier, Agostino, with Montanari, Angelo and Orlandini, Andrea, *AIxIA 2022 – Advances in Artificial Intelligence,* (Switzerland: Springer Nature Switzerland, 2022) <https://link.springer.com/book/10.1007/978-3-031-27181-6 > [accessed 03 November 2023]

Downs, Jane and Lindsay, Martin , *Nine-year-olds added to malicious WhatsApp groups*, (BBC News website, 12 April 2024) <https://www.bbc.com/news/articles/cy0l4z8n1p9o> [accessed 12 April 2024]

Driver, Julia, *Moral Theory*, (The Stanford Encyclopaedia of Philosophy (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.)) <https://plato.stanford.edu/archives/fall2022/entries/moral-theory/> [accessed 15 February 2024]

Duggal, Nikita, *Explaining the Concepts of Quantum Computing*, (Simplilearn Solutions website, 29 January 2023) <https://www.simplilearn.com/tutorials/machine-learning-tutorial/what-is-quantum-computing> [accessed 13 February 2024]

Duignan, Brian, *January 6 U.S. Capitol attack*, (Encyclopaedia Britannica website, 07 February 2024) <https://www.britannica.com/event/January-6-U-S-Capitol-attack> [accessed 15 February 2024]

Dutch News, *Electricity grid capacity shortage now threatens households*, (Dutch News website, 23 January 2024] <https://www.dutchnews.nl/2024/01/electricity-grid-capacity-shortage-now-threatens-households/> [accessed 05 April 2024]

Eastgate Software, *10 Potential Negative Effects of AI in Education*, (Eastgate Software website, 26 January 2024) <https://eastgate-software.com/10-potential-negative-effects-of-ai-in-education/> [accessed 27 March 2024]

Ebert, Roger, *The China Syndrome* (film review), (RogerEbert.com website, 01 January 1979) <https://www.rogerebert.com/reviews/the-china-syndrome-1979> [accessed 08 February 2024]

Edwards, M, *Time and the Science of the Soul in Early Modern Philosophy*, (BRILL, Boston, 2013). Available from: ProQuest Ebook Central. [16 February 2024].

Eijk, Ph. J. van der, *Medicine and philosophy in classical antiquity: doctors and philosophers on nature, soul, health and disease*, (Cambridge University Press, Cambridge, UK; New York, 2005).

Eirgrid (Ireland), *Ireland Capacity Outlook 2022-2031* (Eirgrid, SONI, 2022) <https://www.eirgrid.ie/site-

files/library/EirGrid/EirGrid_SONI_Ireland_Capacity_Outlook_2022-2031.pdf> [accessed 05 April 2024]

European Commission (a), *The Digital Services Act, Ensuring a safe and accountable online environment*, (European Commission website, no date) <https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en> [accessed 22 February 2024]

European Commission (b), *The enforcement framework under the Digital Services Act*, (European Commission website, no date) <https://digital-strategy.ec.europa.eu/en/policies/dsa-enforcement> [accessed 22 February 2024]

European Commission (c), *Regulation (EU) 2016/679 of The European Parliament and of The Council*, (Official Journal of the European Union, The European Commission, 27 April 2016)

European Commission (d), *Antitrust: Commission accepts Microsoft commitments to give users browser choice*, (Press Release, European Commission, 16 December 2009) <https://ec.europa.eu/commission/presscorner/detail/en/IP_09_1941> [accessed 25 April 2024]

European Data Protection Board, *1.2 billion euro fine for Facebook as a result of EDPB binding decision*, (Press Release, European Data Protection Board, 22 May 2023) <https://www.edpb.europa.eu/news/news/2023/12-billion-euro-fine-facebook-result-edpb-binding-decision_en> (accessed 15 March 2024]

European Group on Ethics in Science and New Technologies, *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*, (European Group on Ethics in Science and New Technologies, Directorate-General for Research and Innovation, The European Commission, March 2018), <https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1> [accessed 26 January 2024]

European Group on Ethics in Science and New Technologies, 2018, *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*, 9 March 2018, European Commission, Directorate-General for Research and Innovation, Unit RTD.01. <https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1> [accessed 26 January 2024]

European Parliament (a), *Proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act)*, (European Parliament, 13 March 2024) <https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.html> [accessed 15 March 2024]

European Parliament (b), *Artificial Intelligence Act: MEPs adopt landmark law, Press Release*, (Press release, European Parliament, 14 March 2024) <https://www.europarl.europa.eu/news/en/press-

room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law> [accessed 15 March 2024]

Evangelisti, Dave, *IQ Score Ranges | Are You Average, High, or Genius IQ?*, (Test-Guide.com website, 16 November 2023), <https://www.test-guide.com/iq-score-range.html> [accessed 08 February 2024]

Evers, Jeannie, *El Niño*, (National Geographic Society, 19 October 2023) <https://education.nationalgeographic.org/resource/el-nino/> [accessed 25 March 2024]

Faiola, Anthony and Zakrzewski, Cat, *Governments used to lead innovation. On AI, they're falling behind*, (The Washington Post, 06 November 2023) <https://www.washingtonpost.com/technology/2023/11/02/ai-regulation-bletchley-park/> [accessed 04 July 2024]

The Federal Bureau of Investigation (USA), *Frequently Asked Questions*, (The Federal Bureau of Investigation, website, nd) <https://www.fbi.gov/about/faqs> [accessed 04 June 2024]

Figà Talamanca, G., and Arfini, S., *Through the Newsfeed Glass: Rethinking Filter Bubbles and Echo Chambers*, (Philosophy & technology, 35(1), 20, 2022) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8923337/#:~:text=Firstly%2C%20filter%20bubbles%20describe%20forms,are%20enacted%20by%20users%20themselves> [accessed 11 April 2024]

Fiveable, *Social Disruption*, (US College Education Company, Fiveable Inc website, nd) <https://library.fiveable.me/key-terms/ap-hug/social-disruption> [accessed 25 April 2024]

Fjelland, Ragnar, *Why general artificial intelligence will not be realised*, (Humanities and Social Sciences Communications, Nature website, 17 June 2020) <https://www.nature.com/articles/s41599-020-0494-4> [accessed 25 April 2024]

Fleming, Sean , *A short history of jobs and automation*, (Artificial Intelligence, World Economic Forum website, September 2020) <https://www.weforum.org/agenda/2020/09/short-history-jobs-automation/> [accessed 28 March 2024]

Fletcher, Richard and Khan, Mehreen, *AI is the talk of the town, but what happens next?*, (London: The (London) Times, 20 January 2024).

Flynn, Kerry, *How Gen Z gets its news*, (Axios News website, 16 February 2024) <https://www.axios.com/2024/02/16/tiktok-news-gen-z-social-media#> [accessed 03 July 2024]

Forston, Danny; Rise of the Ballot Bots, Why 2024 Will be the Year of the Deepfake Election; The Sunday Times, 31 December 2023.

Frankel, Julia, *Half a year into the war in Gaza, here's a look at the conflict by the numbers,* (World News, The Associated Press, 06 April 2024),

<https://apnews.com/article/israel-hamas-gaza-war-statistics-95a6407fac94e9d589be234708cd5005> [accessed 15 April 2024]

Frankena, W.K., Stevenson, C.L., Brandt, R. B., Goldman, A.I. & Kim, J., *Values and morals: essays in honor of William Frankena, Charles Stevenson, and Richard Brandt*, (D. Reidel, Boston, 1978)

Frankish, Keith, and Ramsey, William, M., *The Cambridge Handbook of Artificial Intelligence*, (Cambridge University Press, Cambridge, UK. 2014)

Franks, B., *Taming the big data tidal wave: finding opportunities in huge data streams with advanced analytics, 1st edn*, (John Wiley & Sons, Inc, Hoboken, New Jersey, 2012)

Freedman, Lawerence, *The Future of War*, (The Sunday Times Magazine, The Sunday Times, 04 February 2024)

Frontline, *The Power of Big Oil*, (Frontline, Public Broadcasting Service (PBS), 2022) <https://www.pbs.org/wgbh/frontline/documentary/the-power-of-big-oil/> [accessed 12 April 2024]

Gajjar, Devyani, *Artificial intelligence: An explainer*, (London: Parliamentary Office of Science and Technology (POST) UK Parliament, 2023) <https://researchbriefings.files.parliament.uk/documents/POST-PB-0057/POST-PB-0057.pdf> [accessed 15 January 2024]

Gary, Debbie, *Formation flying*, (Encyclopaedia Britannica, nd) <https://www.britannica.com/technology/formation-flying> [accessed 10 July 2024]

GBIF, *Data hosting*, (Global Biodiversity Information Facility (GBIF) website, nd) <https://www.gbif.org/data-hosting> [accessed 03 July 2024]

GCHQ (Government Communications Headquarters), *Overview*, (GCHQ Website, nd) <https://www.gchq.gov.uk/section/mission/overview> [accessed 24 April 2024]

Georgieva, Kristalina, *AI Will Transform the Global Economy. Let's Make Sure It Benefits Humanity*, (Artificial intelligence, IMF (International Monetary Fund) Blog, IMF 14 January 2024) <https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity> [accessed 28 March 2024]

Gert, Bernard and Joshua Gert, *The Definition of Morality*, (The Stanford Encyclopaedia of Philosophy (Fall 2020 Edition), Edward N. Zalta (ed.)) <https://plato.stanford.edu/archives/fall2020/entries/morality-definition/> [accessed 15 February 2024]

Gillis, Alexander S., *Algorithm*, (TechTarget website, 2024) <https://www.techtarget.com/whatis/definition/algorithm#:~:text=An%20algorithm%20is%20a%20procedure,throughout%20all%20areas%20of%20IT> [accessed 14 March 2024]

Giovanola, Benedetta, *Justice, emotions, socially disruptive technologies*, (Critical Review of International Social and Political Philosophy, 26:1, 104-119, 2023) <https://www-tandfonline-com.may.idm.oclc.org/doi/full/10.1080/13698230.2021.1893255> [accessed 01 February 2024]

Glickman, J., *Moral philosophy: an introduction*, (St. Martin's Press, New York, 1976)

Glover, Ellen, *Strong AI vs. Weak AI: What's the Difference?*, (Built In website, 20 April 2024) <https://builtin.com/artificial-intelligence/strong-ai-weak-ai> [accessed 25 April 2024]

Graham, Christine, *What Companies Need to Know About Abuse of Dominance*, (Cooley Legal Firm, Cooley LLP website, 31 January 2023) <https://www.cooley.com/news/insight/2022/2022-12-31-what-companies-need-to-know-about-abuse-of-dominance> [accessed 25 April 2024]

Gray, John, *The New Leviathans: Thoughts After Liberalism*, (Great Britain: Penguin Random House, 2023)

Gregerson, Erik, *Isaac Asimov*, (Encyclopaedia Britannica, 2024) <https://www.britannica.com/biography/Isaac-Asimov> [accessed 26 April 2024]

Grierson, Jamie, *Anger from campaigners as WhatsApp lowers age limit to 13 in UK and EU*, (The Guardian, 12 April 2024) <https://www.theguardian.com/technology/2024/apr/12/anger-from-campaigners-as-whatsapp-lowers-age-limit-to-13-in-uk-and-eu> [accessed 12 April 2024]

Guest, Peter, The UK's Controversial Online Safety Act Is Now Law, (Wired website, 26 October 2023) <https://www.wired.com/story/the-uks-controversial-online-safety-act-is-now-law/> [accessed 22 February 2024]

Gunkel, D.J., *The machine question: critical perspectives on AI, robots, and ethics*, (MIT Press, Cambridge, Mass., 2012)

Hanna, Katie Terrell, *Digital Divide*, (TechTarget website, nd) <https://www.techtarget.com/whatis/definition/digital-divide#:~:text=The%20digital%20divide%20is%20a,t%20or%20have%20restricted%20access> [accessed 22 March 2024]

Hacker, P. M. S., *Is There Anything It Is like to Be a Bat?* (Philosophy, vol. 77, no. 300, 2002, pp. 157–74. JSTOR) <https://www.jstor.org/stable/3752108?seq=7> [accessed 23 February 2024]

Harbath, Katie and Fernekes, Collier, *History of the Cambridge Analytica Controversy*, (Bipartisan Policy Centre, March 2023) <https://bipartisanpolicy.org/blog/cambridge-analytica-controversy/> [accessed 02 February 2024]

Harris, Elizabeth, *9 Amazing Examples of Disruptive Technology (Inspired by the WTIA)*, (Resultist Consulting, No Date), <https://www.resultist.com/blog/9-amazing-

examples-of-disruptive-technology-inspired-by-the-wtia> [accessed 07 February 2024]

Harvard University Press, *Too Big to Jail?*, (Harvard University Press, 03 February 2014) <https://harvardpress.typepad.com/hup_publicity/2014/02/too-big-to-jail-brandon-garrett.html> [accessed 09 July 2024]

Hayek, F A, *The Constitution of Liberty, The Collected Works of F. A. Hayek Volume XVII*, (Chicago: The University Chicago Press, 2011), <https://ia800702.us.archive.org/26/items/TheConstitutionOfLiberty/The%20Constitution%20of%20Liberty.pdf> [accessed 30 November 2023]

Hayek, F A, *The Road to Serfdom*, (London: Routledge, London and New York, 1944, repr. 2006) <https://ctheory.sitehost.iu.edu/img/Hayek_The_Road_to_Serfdom.pdf> [accessed 30 November 2023]

Healthline Media, *12 Signs of Low Emotional Intelligence — Plus Tips for Building It*, (Healthline Media LLC, 2024) <https://www.healthline.com/health/mental-health/low-emotional-intelligence> [accessed 13 March 2024]

Hearne, Dr. Rory, *The Irish water war, austerity and the 'Risen people'*, (Department of Geography, Maynooth University, April 2015) <https://www.maynoothuniversity.ie/sites/default/files/assets/document/TheIrishWaterwar_0.pdf> (accessed 15 March 2024]

Hern, Alex (a), *OpenAI deems its voice cloning tool too risky for general release*, (The Guardian Newspaper, website, 31 March 2024) <https://www.theguardian.com/technology/2024/mar/31/openai-deems-its-voice-cloning-tool-too-risky-for-general-release> [accessed 02 April 2024]

Hern, Alex (b), *Big tech has distracted world from existential risk of AI, says top scientist*, (The Guardian Newspaper, website, 25 May 2024) <https://www.theguardian.com/technology/article/2024/may/25/big-tech-existential-risk-ai-scientist-max-tegmark-regulations> [accessed 27 May 2024]

Hernandez-Orallo, Jose and O'hEigeartiagh, Sean, *Paradigms of Artificial General Intelligence and Their Associated Risks*, (Centre for the Study of Existential Risk, University of Cambridge, UK, July 2019) <https://www.cser.ac.uk/media/uploads/files/Paradigms_of_AGI_abridged.pdf> [accessed 02 April 2024]

Hirst, Graeme, Semantic Interpretation and Ambiguity, (Department of Computer Science, University of Toronto, Toronto, Canada M5S 1A4,1988) <https://www.cs.toronto.edu/pub/gh/Hirst-semInt-88.pdf> [accessed 17 April 2024]

History.com, *Watergate Scandal*, (History.com website, 18 October 2023) <https://www.history.com/topics/1970s/watergate> [accessed 14 February 2024]

Holoyda, Brian, *QAnon, conspiracy theory*, (Britannica, 25 May 2024) <https://www.britannica.com/topic/QAnon> [accessed 20 June 2024]

Hopster, Jeroen, *What are socially disruptive technologies?*, (Technology in Society, Volume 67, 2021) <https://www.sciencedirect.com/science/article/pii/S0160791X21002256> [accessed 25 April 2025]

Hosking, Patrick, *The 'grand old man of economics' is still walking the road to freedom*, (The Times, 04 May 2024)

Horgan-Jones, Jack, *People with just six months to live should be allowed to avail of assisted dying, committee to recommend,* (The Irish Times, 07 March 2024) <https://www.irishtimes.com/politics/oireachtas/2024/03/07/committee-on-assisted-dying-people-with-just-six-months-to-live-should-be-allowed-to-avail-of-assisted-dying/> [accessed 05 April 2024]

Howard, Jeffrey W., *Freedom of Speech*, (The Stanford Encyclopaedia of Philosophy (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.)) <https://plato.stanford.edu/entries/freedom-speech/>, [accessed 28 January 2024]

Howarth, Josh, *Top 32 Social Media Platforms (January 2024)*, (Exploding Topics website, January 2024) <https://explodingtopics.com/blog/top-social-media-platforms> [accessed 02 February 2024]

Huckins, Grace, *Minds of machines: The great AI consciousness conundrum*, (Massachusetts Institute of Technology (MIT) Technology Review website, 16 October 2023) <https://www.technologyreview.com/2023/10/16/1081149/ai-consciousness-conundrum/> [accessed 02 July 2024]

Hunt, Tamlyn, *AI Safety Research Only Enables the Dangers of Runaway Superintelligence*, (Scientific American, website, 09 January 2024) <https://www.scientificamerican.com/article/ai-safety-research-only-enables-the-dangers-of-runaway-superintelligence/> [accessed 02 July 2024]

IBM (x), *What is Strong AI?,* (Topics, IBM website, nd) <https://www.ibm.com/topics/strong-ai> [accessed 05 July 2024]

IBM (a), *What is hacking?,* (Security Topics, IBM website, nd) <https://www.ibm.com/topics/cyber-hacking> [accessed 04 April 2024]

IBM (b), *What is Strong AI?,* (IBM website, nd) <https://www.ibm.com/topics/strong-ai> [accessed 25 April 2024]

IMDb editors, *2001: A Space Odyssey Quotes*, (IMDb editors, IMDb website, 2024) <https://www.imdb.com/title/tt0062622/quotes/> [accessed 14 March 2024]

International Banker, *Should We Fear Artificial Superintelligence?*, (International Banker, website, 18 January 2021) <https://internationalbanker.com/technology/should-we-fear-artificial-superintelligence/> [accessed 02 July 2024]

International Committee of The Red Cross (IRCR), *Fundamentals of International Humanitarian Law*, (How Does Law Protect in War? International Committee of The Red Cross website, nd) <https://casebook.icrc.org/law/fundamentals-ihl> [accessed 24 April 2024]

JD.com, *About Us*, (JD.com website, 2017) <https://corporate.jd.com/ourBusiness> [accessed 08 February 2024]

Jiang, Xinyan, *What Kind Of Knowledge Does a Weak-Willed Person Have? – A Comparative Study of Aristotle and The Ch'eng-Chu School*, (Grand Valley State University, Philosophy East and West, vol. 50, no. 2, 2000, pp. 242–53), <https://www.jstor.org/stable/1400144> {accessed 13 February 2024]

Johnson, Bobbie, *Privacy no longer a social norm, says Facebook founder*, (London: The Guardian 11 January 2010), <https://www.theguardian.com/technology/2010/jan/11/facebook-privacy> [accessed 18 January 2024]

Johnson, Holly, *More than 15,000 children killed in Gaza since October, says Education Ministry*, (The National website, 04 June 2024) <https://www.thenationalnews.com/news/mena/2024/06/04/west-bank-gaza-israel/> [accessed 04 June 2024]

Johnson, Robert and Cureton, Adam, *Kant's Moral Philosophy*, (The Stanford Encyclopaedia of Philosophy (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.)) <https://plato.stanford.edu/entries/kant-moral/> [accessed 14 March 2024]

Jolly, Jasper, *Neurodivergent women sought for jobs at GCHQ and BAE Systems*, (The Guardian 16 November 2022), <https://www.theguardian.com/society/2022/nov/16/neurodiverse-women-sought-for-jobs-at-gchq-and-bae-systems> [accessed 07 February 2024]

Kallenborn, Zachary, *Swarm Clouds on the Horizon? Exploring the Future of Drone Swarm Proliferation*, (Modern War Institute, 20 March 2024) <https://mwi.westpoint.edu/swarm-clouds-on-the-horizon-exploring-the-future-of-drone-swarm-proliferation/> [accessed 10 July 2024]

Kanesan, Pavitra and Fauzan, Norsiah, *Models of Emotional Intelligence: A Review*, (ProQuest, Journal of Social Sciences and Humanities, Vol. 16. No.7 (1-9), ISSN: 1823-884x, 2019) <https://www.proquest.com/docview/2396318735?fromopenview=true&pq-origsite=gscholar&parentSessionId=Ov1T60L9QShIhxYomwvlqFnR1uenuVktdSH%2BFAcJQZY%3D&sourcetype=Scholarly%20Journals> [accessed 31 January 2024]

Kang, Cecilia and Satariano, Adam, *As A.I. Booms, Lawmakers Struggle to Understand the Technology*, (The New York Times website, 03 March 2023) <https://www.nytimes.com/2023/03/03/technology/artificial-intelligence-regulation-congress.html> [accessed 11 April 2024]

Kasperbauer, T.J., *Conflicting roles for humans in learning health systems and AI-enabled healthcare*, (Journal of evaluation in clinical practice, vol. 27, no. 3, pp. 537-542, 2021) <https://maynoothuniversity.userservices.exlibrisgroup.com/discovery/openurl?institution=353MAY_INST&vid=353MAY_INST:Services&%3Fctx_ver=Z39.88-2004&ctx_enc=info:ofi%2Fenc:UTF-8&rfr_id=info:sid%2Fsummon.serialssolutions.com&rft_val_fmt=info:ofi%2Ffmt:kev:mtx:journal&rft.genre=article&rft.atitle=Conflicting%20roles%20for%20humans%20in%20learning%20health%20systems%20and%20AI%E2%80%90enabled%20healthcare&rft.jtitle=Journal%20of%20evaluation%20in%20clinical%20practice&rft.au=Kasperbauer,%20T.%20J.&rft.date=2021-06-01&rft.pub=John%20Wiley%20%26%20Sons,%20Inc&rft.issn=1356-1294&rft.eissn=1365-2753&rft.volume=27&rft.issue=3&rft.spage=537&rft.epage=542&rft_id=info:doi%2F10.1111%2Fjep.13510&rft.externalDocID=JEP13510&paramdict=en-us> [accessed 19 March 2024]

Kates, Graham and MacFarlane, *Scott, Partner of Brian Sicknick, officer who died after Jan. 6 attack, sues Trump, 2 others*, (CBS News website, 06 January 2023) <https://www.cbsnews.com/news/brian-sicknick-jan-6-attack-lawsuit-trump-george-tanios-julian-khater/> [accessed 15 February 2024]

Kelly, Robert, *Antitrust Laws: What They Are, How They Work, Major Examples,* (Investopedia website, 31 January 2023) <https://www.investopedia.com/terms/a/antitrust.asp> [accessed 05 April 2024]

Kenny, Anthony,. *A New History of Western Philosophy: in Four Parts,* (Clarendon Press, Oxford ([England]; New York; 2010)

Kenton, Will, *Externality: What It Means in Economics, With Positive and Negative Examples,* (Economics, Investopedia, February 2024) <https://www.investopedia.com/terms/e/externality.asp> [accessed 07 May 2024]

Kenton, Will, *Structural Unemployment: Definition, Causes, and Examples*, (Macroeconomics, Investopedia website, 06 January 2023) <https://www.investopedia.com/terms/s/structuralunemployment.asp> [accessed 05 April 2024]

Kington, Tom, *Don't play God, Pope's AI guru tell tech titans*, (London: The (London) Times, 20 January 2024).

Kleinman, Zoe, *NHS AI test spots tiny cancers missed by doctors*, (BBC News website, 21 March 2024) <https://www.bbc.com/news/technology-68607059> [accessed 21 March 2024]

Knight, Kevin, *Question 40. War – Summa Theologiae*, (Second Part of the Second Part, The Summa Theologiæ of St. Thomas Aquinas, Literally translated by Fathers of the English Dominican Province, 2017) <https://www.newadvent.org/summa/3040.htm> [accessed 24 April 2024]

Korstanje, Maximiliano & Skoll, Geoffrey, *An Essay on the Social Costs and Benefits of Technology Evolution*, (Human & Social Studies. Research and Practice. 2. 10.2478/hssr-2013-0002, 2013) <https://www.researchgate.net/profile/Geoffrey-Skoll/publication/275995652_An_Essay_on_the_Social_Costs_and_Benefits_of_Technology_Evolution/links/56efe24f08ae52f8ad7f7f43/An-Essay-on-the-Social-Costs-and-Benefits-of-Technology-Evolution.pdf> [accessed 08 March 2024]

Krar, Steve, *How Computers Work*, (Manufacturing Automation website, nd) <https://www.automationmag.com/images/stories/LWTech-files/25%20How%20Computers%20Work.pdf> [accessed 03 April 2024]

Kremer, William, *More than a million people die on roads every year. Meet the man determined to prevent them,* (Future, BBC News website, 19 May 2024) <https://www.bbc.com/future/article/20240517-vision-zero-how-europe-cut-the-number-of-people-dying-on-its-roads> [accessed 25 May 2024]

Kuhn, George, *How Target Used Data Analytics to Predict Pregnancies*, (Drive Research website, 01 August 2023), <https://www.driveresearch.com/market-research-company-blog/how-target-used-data-analytics-to-predict-pregnancies/> {accessed 09 February 2024]

Lazar, Seth, *War*, (The Stanford Encyclopaedia of Philosophy (Spring 2020 Edition), Edward N. Zalta (ed.)) <https://plato.stanford.edu/archives/spr2020/entries/war/> [accessed 23 April 2024]

Legal Information Institute, *Too big to fail,* (Legal Information Institute, Cornell Law School, NY, August 2021) <https://www.law.cornell.edu/wex/too_big_to_fail#:~:text=%E2%80%9CToo%20big%20to%20fail%E2%80%9D%20refers,seriousness%20of%20the%20economic%20repercussions> [accessed 09 July 2024]

Liao, Shen-yi and Tamar, Gendler, *Imagination*, (The Stanford Encyclopaedia of Philosophy (Summer 2020 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/entries/imagination/> [accessed 13 February 2024]

(The) Library of Congress, *Today in History - December 15 - The Bill of Rights*, (The Library of Congress website, nd) <https://www.loc.gov/item/today-in-history/december-15/#:~:text=On%20December%2015%2C%201791%2C%20the,of%20peaceful%20assembly%20and%20petition> [accessed 16 March 2024]

Lorenz, Hendrik, *Ancient Theories of Soul*, (The Stanford Encyclopaedia of Philosophy (Summer 2009 Edition), Edward N. Zalta (ed.)) <https://plato.stanford.edu/entries/ancient-soul/> [accessed 15 February 2024]

Lu, Marcus, *The 50 Most Valuable Companies in the World in 2023*, (Visual Capitalist website, 20 August 2023), <https://www.visualcapitalist.com/the-50-most-valuable-companies-in-the-world-in-2023/> {accessed 08 February 2024]

Lum, Kristian and Chowdhury, Rumman, *What is an "algorithm"? It depends whom you ask*, (MIT Technology Review, 26 February 2021)

<https://www.technologyreview.com/2021/02/26/1020007/what-is-an-algorithm/> [accessed 14 March 2024]

MacFarlane, Sir Alistair, *Information, Knowledge & Intelligence*, (Issue 98, Philosophy Now, 2013) <https://philosophynow.org/issues/98/Information_Knowledge_and_Intelligence> [accessed 31 January 2024]

Manning, Christopher, Professor, *Artificial Intelligence Definitions*, (Artificial Intelligence Definitions, Stanford University, Human-Centred Artificial Intelligence, April 2022) <https://hai.stanford.edu/sites/default/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf> [accessed 22 February 2024]

Marr, Bernard, (a), *15 Things Everyone Should Know About Quantum Computing*, (Bernard Marr website, Bernard Marr & Co., 2021) <https://bernardmarr.com/15-things-everyone-should-know-about-quantum-computing/> [accessed 13 February 2024]

Marr, Bernard, (b) *What Is Quantum Computing? A Super-Easy Explanation For Anyone*, (Bernard Marr website, Bernard Marr & Co., 2021) <https://bernardmarr.com/what-is-quantum-computing-a-super-easy-explanation-for-anyone/> [accessed 13 February 2024]

Masters, James, *The Rules of Go or Wei Chi*, (Masters Traditional Games, Masters of Games website, 2024) <https://www.mastersofgames.com/rules/go-rules.htm> [accessed 02 February 2024]

Marchant, Gary, *Why Soft Law is the Best Way to Approach the Pacing Problem in AI,* (Carnegie Council for Ethics in International Affairs website, 29 September 2021) <https://www.carnegiecouncil.org/media/article/why-soft-law-is-the-best-way-to-approach-the-pacing-problem-in-ai> [accessed 16 March 2024]

McClain, Colleen, *More so than adults, U.S. teens value people feeling safe online over being able to speak freely*, (Pew Research Centre, 30 August 2022), <https://www.pewresearch.org/short-reads/2022/08/30/more-so-than-adults-u-s-teens-value-people-feeling-safe-online-over-being-able-to-speak-freely/> [accessed 07 February]

McCombs School of Business, *Deontology*, (Ethics Unwrapped, McCombs School of Business, nd) <https://ethicsunwrapped.utexas.edu/glossary/deontology#:~:text=Deontology%20is%20an%20ethical%20theory%20that%20uses%20rules%20to%20distinguish,Don't%20cheat.%E2%80%9D> [accessed 13 June 2024]

McGreal, Chris, *'What we now know … they lied': how big oil companies betrayed us all*, (The Guardian, 21 April 2022) <https://www.theguardian.com/tv-and-radio/2022/apr/20/what-we-now-know-they-lied-how-big-oil-companies-betrayed-us-all> [accessed 12 April 2024]

McKinsey & Company, *What is Davos?*, (McKinsey & Company website, 11 December 2023) <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-davos> [accessed 08 February 2024]

McMahon, Liv and Kleinman, Zoe, *AI Safety: UK and US sign landmark agreement*, (BBC News website, 02 April 2024) <https://www.bbc.com/news/technology-68675654> [accessed 05 April 2024]

Mediratta, Anuj, *Big Data: Terms, Definitions, and Applications*, (Ace Data Services Pvt. Ltd., EMC Corporation, 2015), Published on Dell's Education website <https://education.dell.com/content/dam/dell-emc/documents/en-us/2015KS_Mediratta-Big_Data_Terms,_Definitions_and_Applications.pdf> [accessed 19 February 2024]

Mental Health Foundation, *The truth about self-harm,* (Mental Health Foundation website, UK, nd) <https://www.mentalhealth.org.uk/explore-mental-health/publications/truth-about-self-harm> [accessed 12 April 2024]

Mercer County Community College, *Intelligence (Types)*, (Mercer County Community College, 2024), <https://www.mccc.edu/~jenningh/Courses/documents/Handout-Intelligence_000.pdf> [accessed 31 January 2024]

Merriman Webster Dictionary, *animatronic*, (Merriman Webster Dictionary, nd) <https://www.merriam-webster.com/dictionary/animatronic> [accessed 05 April 2024]

Merriam Webster Dictionary, *geopolitics*, (Merriam Webster Dictionary website, nd) <https://www.merriam-webster.com/dictionary/geopolitics> [accessed 02 July 2024]

Merriman Webster Dictionary, *bionics*, (Merriman Webster Dictionary, website, nd) <https://www.merriam-webster.com/dictionary/bionics#:~:text=Kids%20Definition-,bionics,the%20solution%20of%20engineering%20problems> [accessed 09 May 2024]

Merriman Webster Dictionary, *Presentism,* (Merriman Webster Dictionary, nd) <https://www.merriam-webster.com/dictionary/presentism> [accessed 17 June 2024]

Merriman Webster Dictionary, *Subliminal influence,* (Merriman Webster Dictionary, nd), <https://www.merriam-webster.com/dictionary/subliminal> [accessed 05 April 2024]

Merriman Webster Dictionary, *Machiavellian*, (Merriman Webster Dictionary, website, nd) <https://www.merriam-webster.com/dictionary/Machiavellian> [accessed 05 April 2024]

Meta, *Actions*, (Meta website, 2024), <https://about.meta.com/actions/> [accessed 16 February 2024]

Meta, *2023 Responsible Business Practices Report*, (Meta website 2023 Report) <https://scontent-dub4-1.xx.fbcdn.net/v/t39.8562-6/10000000_1974692352892843_2880908794103790486_n.pdf?_nc_cat=107&ccb=1-7&_nc_sid=e280be&_nc_ohc=JWdh3kcbuw8AX8YOo9c&_nc_ht=scontent-dub4-1.xx&oh=00_AfAf8aAIH9kG7IIKMegqhKEpT0e3PK5_v9QsbGF6qBERRg&oe=65D4A7C5> [accessed 16 February 2024]

Meta, *Facebook Community Standards*, (Meta website, nd) <https://transparency.meta.com/en-gb/policies/community-standards/> [accessed 19 April 2024]

Microsoft, *What is a cyberattack?*, (Microsoft, website, nd) <https://www.microsoft.com/en-ie/security/business/security-101/what-is-a-cyberattack> [accessed 24 April 2024]

Milmo, Dan and Hern, Alex, *What will the EU's proposed act to regulate AI mean for consumers?*, (The Guardian website, 14 March 2024) <https://www.theguardian.com/technology/2024/mar/14/what-will-eu-proposed-regulation-ai-mean-consumers> [accessed 15 March 2024]

Milmo, Dan, *Why is US threatening to ban TikTok and will other countries follow suit?*, (The Guardian website, 24 April 2024) <https://www.theguardian.com/technology/2024/apr/24/why-is-us-threatening-to-ban-tiktok-and-could-other-countries-follow-suit> [accessed 25 April 2024]

Ministry of Defence (UK), *Wargaming Handbook*, (Development, Concepts and Doctrine Centre, Ministry of Defence, UK, 2017) <https://assets.publishing.service.gov.uk/media/5a82e90d40f0b6230269d575/doctrine_uk_wargaming_handbook.pdf> [accessed 04 July 2024]

Minow, M., *Dignity and Human Rights: Aspirations and Challenges in an Age of Political Divisions, Distrust, and AI*, (Telos (New York, N.Y.), , no. 203, pp. 21-39, 2023) <https://maynoothuniversity.userservices.exlibrisgroup.com/discovery/openurl?institution=353MAY_INST&vid=353MAY_INST:Services&%3Fctx_ver=Z39.88-2004&ctx_enc=info:ofi%2Fenc:UTF-8&rfr_id=info:sid%2Fsummon.serialssolutions.com&rft_val_fmt=info:ofi%2Ffmt:kev:mtx:journal&rft.genre=article&rft.atitle=Dignity%20and%20Human%20Rights:%20Aspirations%20and%20Challenges%20in%20an%20Age%20of%20Political%20Divisions,%20Distrust,%20and%20AI&rft.jtitle=Telos%20(New%20York,%20N.Y.)&rft.au=Minow,%20Martha&rft.date=2023-07-01&rft.pub=Telos%20Press%20Ltd&rft.issn=0090-6514&rft.eissn=1940-459X&rft.issue=203&rft.spage=21&rft.epage=39&rft_id=info:doi%2F10.3817%2F0623203021&rft.externalDBID=n%2Fa&rft.externalDocID=001052960500003&paramdict=en-us> [accessed 05 April 2024]

Mokyr, Joel, *Long-Term Economic Growth and the History of Technology*, (Handbook of Economic Growth, Science Direct website, 2005)

<https://www.sciencedirect.com/science/article/abs/pii/S1574068405010178> [accessed 08 March 2024]

Morreall, John, *Philosophy of Humour*, The Stanford Encyclopaedia of Philosophy, Summer 2023, Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/entries/humor/> [accessed 25 March 2024]

Morrison, Nick, *School Leaders Warn AI Is A 'Real And Present' Danger To Education*, (Forbes Media LLC, website, May 2023) <https://www.forbes.com/sites/nickmorrison/2023/05/20/school-leaders-warn-ai-is-a-real-and-present-danger-to-education/> [accessed 28 March 2024]

Müller, Vincent C., Ethics of Artificial Intelligence and Robotics, (California: The Stanford Encyclopaedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/fall2023/entries/ethics-ai/> [accessed 17 January 2024]

Müller, Vincent C., *The history of digital ethics*, (in Carissa Ve liz (ed.), Oxford Handbook of Digital Ethics, Oxford University Press, 2022)

Müller, Vincent C, *Philosophy and Theory of Artificial Intelligence 2017,* (Leeds: University of Leeds, Springer Nature Switzerland, 2018) <https://www.researchgate.net/publication/327228207_Philosophy_and_theory_of_artificial_intelligence_2017> [accessed 28 September 2023]

Nagel, Thomas, *What Is It Like to Be a Bat?* (The Philosophical Review, vol. 83, no. 4, 1974, pp. 435–50. JSTOR) <https://www.jstor.org/stable/2183914> [accessed 23 February 2024]

National Library of Medicine, *Prosthetics through the ages*, (National Library of Medicine website, 11 May 2023) < https://magazine.medlineplus.gov/article/prosthetics-through-the-ages> [accessed 27 March 2024]

Nederman, Cary, *Niccolò Machiavelli*, (The Stanford Encyclopaedia of Philosophy, Winter 2023 Edition, Edward N. Zalta & Uri Nodelman (eds.)) <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=machiavelli> [accessed 13 March 2024]

Nordquist, Richard, *The Theory of Poverty of the Stimulus in Language Development*, (ThoughtCo, Dotdash Meredith Publishing, 03 July 2019) <https://www.thoughtco.com/poverty-of-the-stimulus-pos-1691521> [accessed 08 April 2024]

O'Donovan, Brian, *Trial of new connected vehicle technology to be launched*, (RTE, 01 February 2024) <https://www.rte.ie/news/ireland/2024/0201/1429859-car-tech/> [accessed 01 February 2024]

Ocklenburg, Sebastian and Güntürkün, Onur, *Recognising Yourself and Others*, (The Lateralized Brain, 2018, from ScienceDirect website 2024)

<https://www.sciencedirect.com/topics/neuroscience/mirror-test> [accessed 16 March 2024]

Ofcom, *Light-hearted news on social media drawing Gen Z away from traditional sources*, (Ofcom website, 20 July 2023) <https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/light-hearted-news-social-media-drawing-gen-z/> [accessed 03 July 2024]

Office of Public Affairs, No Company or Individual is "Too Big to Jail", (Office of Public Affairs, U.S. Department of Justice, 18 January 2023) <https://www.justice.gov/opa/video/no-company-or-individual-too-big-jail> [accessed 08 July 2024]

Ohio State University, *What is Psychology?*, (Ohio State University, 2014) <https://psychology.osu.edu/about/what-psychology#:~:text=Psychology%20is%20the%20scientific%20study,%2C%20brain%20functions%2C%20and%20behavior> [accessed 13 March 2024]

Oi, Mariko, *Neuralink: Musk's firm says first brain-chip patient plays online chess,* (BBC News website, 21 March 2024), <https://www.bbc.com/news/business-68622781> [accessed 21 March 2024]

Oladele, Ayo, *AI Opportunities: Top 9 Expectations*, (Velvetech website, 2024) <https://www.velvetech.com/blog/ai-opportunities-top-6-expectations/> [accessed 14 March 2024]

OpenAI, Introducing ChatGPT, (OpenAI website, 2024), <https://openai.com/blog/chatgpt> [accessed 27 March 2024]

Ortiz, Sabrina, *What is ChatGPT and why does it matter? Here's what you need to know,* (ZDNET, website, 17 June 2024) <https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-everything-you-need-to-know/> [accessed 24 June 2024]

Palanca-Castan N, Sánchez Tajadura B, Cofré R., *Towards an interdisciplinary framework about intelligence*, (Heliyon, v.7(2), National Library of Medicine, February 2021) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7902546/> [accessed 01 February 2024]

Palm, Jenny, *Exploring Limited Capacity in the Grid: Actors, Problems, and Solutions*, (Front. Energy Res. 9:663769. doi: 10.3389/fenrg.2021.663769, 2021) <https://www.frontiersin.org/articles/10.3389/fenrg.2021.663769/full#:~:text=Lack%20of%20capacity%20in%20the%20electricity%20networks%20has%20led%20to,e.g.%2C%20T%C3%B6rnwall%2C%202019> [accessed 05 April 2024}

Panovski, Antonio, *How Did Philosophy Help Develop Artificial Intelligence?*, (TheCollector.com, December 2023),<https://www.thecollector.com/philosophy-artificial-intelligence-development/> [accessed 01 February 2024]

Pathi, Krutika, *Here's what happened when India banned TikTok*, (Associated Press, for Public Broadcasting Service (PBS), 24 April 2024) <https://www.pbs.org/newshour/world/heres-what-happened-when-india-banned-tiktok> [accessed 25 April 2024]

Pauly, Marc , (University of Groningen The Netherlands, The Internet Encyclopaedia of Philosophy website, No Date) <https://iep.utm.edu/conspiracy-theories/> [accessed 14 February 2024]

PCMag, *Transistor Density*, (PCMag website, nd) <https://www.pcmag.com/encyclopedia/term/transistor-density#:~:text=The%20number%20of%20transistors%20per,%2Dof%2Dthe%2Dart> [accessed 03 April 2024]

Pearl, Lisa, *Poverty of the Stimulus Without Tears*, (University of California, nd) <https://sites.socsci.uci.edu/~lpearl/courses/readings/Pearl2019Ms_PovStimWithoutTears.pdf> [accessed 08 April 2024]

Pepper, Diarmuid, *WhatsApp to lower minimum age requirement across Europe from 16 to 13*, (The Journal website, 15 February 2024) <https://www.thejournal.ie/whatsapp-lowers-minimum-age-requirement-across-europe-6300396-Feb2024/> [accessed 12 April 2024]

Pettinger, Tejvan, *Supernormal Profits*, (Economicshelp.org website, 10 September 2019) <https://www.economicshelp.org/blog/3181/economics/supernormal-profits/> [accessed 02 July 2024]

Peyre, Henri M., *Émile Durkheim*, (Encyclopaedia Britannica, 22 May. 2024) <https://www.britannica.com/biography/Emile-Durkheim> [accessed 24 June 2024]

Pfeiffer, Lee, *2001: A Space Odyssey*, (Encyclopaedia Britannica, 13 February 2024) <https://www.britannica.com/topic/2001-A-Space-Odyssey-film-1968> [accessed 14 March 2024]

Psynso, *Poverty of the Stimulus*, (Psynso website, nd) <https://psynso.com/poverty-of-the-stimulus/> [accessed 08 March 2024]

Poel, I.v.d., Frank, L., Hermann, J., Hopster, J., Lenzi, D., Nyholm, S., Taebi, B. & Ziliotti, E., *Ethics of Socially Disruptive Technologies: An Introduction*, (First edn, Open Book Publishers, Cambridge, UK., 2023)

Poel, I.v.d., *Socially Disruptive Technologies, Contextual Integrity, and Conservatism About Moral Change*, (Philosophy & technology, vol. 35, no. 3, pp. 82-82. 2022)

Popkin, Helen A.S., *Privacy is dead on Facebook. Get over it.*, (USA: NBC News, 13 January 2010), <https://www.nbcnews.com/id/wbna34825225> [accessed 18 January 2024]

Press Trust of India, *Foxconn invites bids for construction of chip assembly plant in India*, (Press Trust of India, Business Standard website, 06 February 2024)

<https://www.business-standard.com/india-news/foxconn-invites-bids-for-construction-of-chip-assembly-plant-in-india-124020601795_1.html> [accessed 08 February 2024]

Premier Energy, *The Electricity Capacity Crisis – SSEN Ramped Connection*, (Premier Energy Consultants (UK), 12 April 2023) <https://www.premierenergy.co.uk/blog/the-electricity-capacity-crisis-ssen-ramped-connection/> [accessed 05 April 2024]

PWC, 2021, *Conti cyber-attack on the HSE – Independent Post Incident Review,* (Price Waterhouse Coopers (PWC), Commissioned by the HSE Board in conjunction with the CEO and Executive Management Team, 03 December 2021) < https://www.hse.ie/eng/services/publications/conti-cyber-attack-on-the-hse-full-report.pdf> [accessed 20 December 2023]

Qiu, Linda, *Fact-Checking the Breadth of Trump's Election Lies,* (The New York Times website, The New York Times Company, 17 August 2023) <https://www.nytimes.com/2023/08/17/us/politics/trump-election-lies-fact-check.html> [accessed 15 February]

Quist, Terry C., (2022) *What philosophy can do for intelligence*, (Intelligence and National Security, Tayllor & Francis Group, Informa UK Limited, May 2022), <https://www.tandfonline.com/doi/full/10.1080/02684527.2022.2076328> {accessed 31 January 2024]

Rajtmajer, Sarah and Susser, Daniel, 2020. *Automated Influence and the Challenge of Cognitive Security*, In Hot Topics in the Science of Security Symposium (HotSoS '20), April 7–8, 2020, Lawrence, KS, USA. ACM, New York, NY, USA, 9 pages < https://philpapers.org/rec/RAJAIA-2 > [accessed 29 January 2024]

Rao, Pallavi, *Visualizing the $105 Trillion World Economy in One Chart*, (Visualcapitalist.com website, 06 August 2023) <https://www.visualcapitalist.com/visualizing-the-105-trillion-world-economy-in-one-chart/#:~:text=The%20U.S.%20is%20expected%20to,%2419.4%20trillion%20GDP%20in%202023> [accessed 08 February 2024]

Reichental, Jonathan, PhD, *Quantum Artificial Intelligence Is Closer Than You Think*, (Forbes Magazine Website, Forbes Media LLC, 20 November 2023) <https://www.forbes.com/sites/jonathanreichental/2023/11/20/quantum-artificial-intelligence-is-closer-than-you-think/> [accessed 13 February 2024]

Reid, Scott A; *Conspiracy theory*, (Encyclopaedia Britannica, 6 March 2024) <https://www.britannica.com/topic/conspiracy-theory> [accessed 11 March 2024]

Rexel, *Competition law prohibits abusive behaviour by companies in a dominant market position*, (Rexel website, 2020) <https://ethique.rexel.com/en/competition-law/abuse-of-a-dominant-position/#:~:text=A%20dominant%20position%20is%20not,player%20in%20a%20fragmented%20market> [accessed 25 April 2024]

Riedel, Stefan, *Biological warfare and bioterrorism: a historical review*, (National Centre for Biotechnology Information, National Library of Medicine, 2004) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1200679/> [accessed 22 April 2024]

Riggs, Allynn, *Quotes, Quotable Quote from The Blood*, (Goodreads Website, August 2014) <https://www.goodreads.com/quotes/1328710-you-didn-t-have-to-look-your-enemy-in-the-eye> [accessed 16 July 2024]

Riordan, Michael, T*ransistor Electronics,* (Encyclopaedia Britannica, 29 March 2024) <https://www.britannica.com/technology/transistor/Innovation-at-Bell-Labs> [accessed 03 April 2024]

Risse, Mathias, *Brave New World: Political Philosophy and AI*, (Cambridge: Cambridge University Press, 18 January 2023), <https://www.cambridgeblog.org/2023/01/brave-new-world-political-philosophy-and-ai/> [accessed 22 January 2024]

Rohlf, Michael, *Immanuel Kant*, California: The Stanford Encyclopedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/entries/kant/> [accessed 15 January 2024]

Rosati, Connie S., *Moral Motivation*, (The Stanford Encyclopaedia of Philosophy (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.)) <https://plato.stanford.edu/Archives/fall2022/entries/moral-motivation/> [accessed 15 February 2024]

Rothschild, Mike, *The Storm is Upon Us, How QANON became a Movement, Cult, and Conspiracy Theory of Everything*, (Monoray, Octopus Publishing Group, 2021)

Runciman, David, *Rescuing Democracy in the Age of the Internet*, (Cambridge: Ethics & International Affairs, 29, no. 3 (2015), pp. 331-344, Cambridge University Press), <https://www-cambridge-org.may.idm.oclc.org/core/journals/ethics-and-international-affairs/article/rescuing-democracy-in-the-age-of-the-internet/736CFD1A158ADEF668BAF844AC75AD91> [accessed 15 January 2024]

Sachs, Joe, *Aristotle: Ethics*, (Internet Encyclopaedia of Philosophy, nd) <https://iep.utm.edu/aristotle-ethics/> [accessed 10 July 2024]

Samsung, *About Us,* (Samsung website, 2024), < https://www.samsung.com/ie/about-us/company-info/> [accessed 08 February 2024]

Schmidtz, David and Boettke, Peter, *Friedrich Hayek*, California: The Stanford Encyclopaedia of Philosophy (Summer 2021 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/entries/friedrich-hayek/>. [accessed 15 January 2024]

Schneider, Barry. R, *Chemical Weapon*, (Encyclopedia Britannica, 03 Februaty 2024) <https://www.britannica.com/technology/chemical-weapon> [Accessed 22 April 2024]

Seal, Thomas, *5G-coronavirus conspiracy theory spurs rash of telecom tower arson fires*, (Fortune Magazine, website, Fortune Media IP Limited, 06 April 2020) <https://fortune.com/2020/04/06/5g-coronavirus-conspiracy-theory-telecom-tower-fires/> [accessed 03 July 2024]

Seddon, Peter B; Constantinidis, Dora; Tamm, Toomas and Dod, Harjot, *How does business analytics contribute to business value?*, (University of New South Wales, Australia, John Wiley & Sons, 2016), <https://onlinelibrary.wiley.com/doi/full/10.1111/isj.12101> [accessed 30 January 2024]

Sentient Digital, Inc., *The Most Useful Military Applications of Ai in 2024 and Beyond*, (Sentient Digital, Inc., website, March 2024) <https://sdi.ai/blog/the-most-useful-military-applications-of-ai/> [accessed 05 July 2024]

Sheldon, R., Hanna, K. T., *Cyberterrorism*, (TechTarget website, January 2022) <https://www.techtarget.com/searchsecurity/definition/cyberterrorism> [accessed 20 December 2023]

Sheldon, Robert, *App,* (TechTarget, website, nd) <https://www.techtarget.com/searchmobilecomputing/definition/app#:~:text=App%20is%20an%20abbreviated%20form,cases%2C%20for%20another%20software%20program> [accessed 01 May 2024]

Shulman, Carl and Bostrom, *Sharing the World with Digital Minds*, (in Clarke, S., Zohny, H. & Savulescu, J. (eds.): Rethinking Moral Status (Oxford University Press, 2021) <https://nickbostrom.com/papers/digital-minds.pdf> [accessed 06 March 2024]

Singh, Spandana, *Everything in Moderation, An Analysis of How Internet Platforms Are Using Artificial Intelligence to Moderate UserGenerated Content*, (New America, Research and Policy Recommendations, website, 17 July 2019) <https://d1y8sb8igg2f8e.cloudfront.net/documents/Everything_in_Moderation_2019-07-15_142127_tq36vr4.pdf> [accessed 03 July 2024]

Smallman, M., *Multi Scale Ethics—Why We Need to Consider the Ethics of AI in Healthcare at Different Scales*, (Science and engineering ethics, vol. 28, no. 6, pp. 63-63, 2022) <https://www.proquest.com/docview/2740754838?https://search-proquest-com.may.idm.oclc.org/central/&parentSessionId=reK6G0zbWlYZHcNheksx6e1YlxDL8uaAIabclunnZnA%3D&pq-origsite=summon&accountid=12309&sourcetype=Scholarly%20Journals> [accessed 19 March 2024]

Smith, Tim, *Disruptive Technology: Definition, Example, and How to Invest*, (Investopedia, Dotdash Meredith Publishing, April 2022) <https://www.investopedia.com/terms/d/disruptive-technology.asp> [accessed 01 January 2024]

Smowltech, *Digital ethics: definition and examples*, (Smowltech website, 15 May 2024) <https://smowl.net/en/blog/digital-

ethics/#:~:text=Digital%20ethics%20encompasses%20principles%2C%20guidelines,trust%2C%20responsibility%2C%20and%20accountability> [accessed 02 July 2024]

Snow, Robert, *Combine, Cluster, Complement: 3 Fundamentals of Combinatorial Digital Innovation*, (Gartner Inc. website, 02 February 2021) <https://www.gartner.com/smarterwithgartner/combine-cluster-complement-3-fundamentals-of-combinatorial-digital-innovation> [accessed 18 March 2024]

Sobolev, Michael, *Digital Nudging: Using Technology to Nudge for Good*, (in SSRN Electronic Journal, ResearchGate, March 2021) <https://www.researchgate.net/publication/353411014_Digital_Nudging_Using_Technology_to_Nudge_for_Good?enrichId=rgreq-63f95a2252d2d7d0d649e5b27ceec42c-XXX&enrichSource=Y292ZXJQYWdlOzM1MzQxMTAxNDtBUzoxMDQ4ODQ1NTg4OTA1OTg0QDE2MjcwNzU2ODM2MzU%3D&el=1_x_2> {accessed 08 July 2024]

Somers, James, *How the Artificial-Intelligence Program AlphaZero Mastered Its Games*, (The New Yorker, 28 December 2018), <https://www.newyorker.com/science/elements/how-the-artificial-intelligence-program-alphazero-mastered-its-games> [accessed 01 February 2024]

Specialisterne Foundation, *The Autism Advantage*, (Specialisterne Foundation (Denmark), 16 May 2022), <https://dk.specialisterne.com/the-autism-advantage/> [accessed 07 February 2024]

Stahl, Bernd Carsten, with Schroeder, Doris and Rodrigues, Rowena, *Ethics of Artificial Intelligence Case Studies and Options for Addressing Ethical Challenges,* (Switzerland: Springer Briefs in Research and Innovation Governance, Springer Nature Switzerland, 2023) <https://link.springer.com/book/10.1007/978-3-031-17040-9> [accessed 28 September 2023]

Stash Banking Services, *15 Largest AI Companies in 2024*, (Investing, Stash Banking Services, 7 February 2024) <https://www.stash.com/learn/top-ai-companies/> [accessed 12 April 2024]

Sternberg, Dr. Robert J., *Intelligence*, (Dialogues Clin Neuroscience, National Centre for Biotechnology Information, National Library of Medicine (USA), 2012), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3341646/> [accessed 07 February 2024]

Stevenson, L., *Twelve Conceptions of Imagination*, (The British journal of aesthetics, vol. 43, no. 3, pp. 238-259, 2003) <https://academic.oup.com/bjaesthetics/article/43/3/238/90315?login=true> [accessed 14 February 2024]

Stevenson, Leslie, *The Study of Human Nature: readings*, (Oxford University Press, New York, 1981)

Sulyeman, Mustafa, *The Coming Wave,* (London: Penguin Random House, 2023)

Susser, Daniel and Grimaldi, Vincent; *Measuring Automated Influence: Between Empirical Evidence and Ethical Values* (2021); Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3848919> [accessed 29 January 2024]

Szczepański, Marcin, *Economic impacts of artificial intelligence (AI)*, (Briefing, EPRS | European Parliamentary Research Service, European Union, July 2019) <https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf> [accessed 28 March 2024]

Szczepański, Marcin, *The productivity riddle*, (Briefing, EPRS | European Parliamentary Research Service, European Union, December 2018) <https://www.europarl.europa.eu/RegData/etudes/BRIE/2018/630319/EPRS_BRI(2018)630319_EN.pdf> [accessed 28 March 2024]

Tableau.com, *What is the history of artificial intelligence (AI)?*, (Tableau.com website, nd) <https://www.tableau.com/data-insights/ai/history#:~:text=Birth%20of%20AI%3A%201950%2D1956&text=The%20term%20%E2%80%9Cartificial%20intelligence%E2%80%9D%20was,intelligence%20called%20The%20Imitation%20Game> [accessed 27 June 2024]

Target.com, *About*, (Target.com website, 2024), <https://corporate.target.com/about> [accessed 09 February]

Tenenbaum, Sergio, *The Judgment of a Weak Will*, (University of New Mexico, Philosophy and Phenomenological Research Vol. LIX, No. 4, December 1999) <https://fitelson.org/seminar/tennenbaum.pdf> [accessed 13 February 2024]

Tescoplc.com, *Tesco Financial Results 2023 – Five-year record*, (Tescoplc.com website, 2023), <https://www.tescoplc.com/investors/reports-results-and-presentations/financial-performance/five-year-record/> [accessed 09 February 2024]

Teuscher, Christof, *Alan Turing: Life and Legacy of a Great Thinker*, (Germany: Springer-Verlag, 2004)

Thierer, Adam, *The Pacing Problem and the Future of Technology Regulation*, (The Mercatus Centre at George Mason University website, 08 August 2018) <https://www.mercatus.org/economic-insights/expert-commentary/pacing-problem-and-future-technology-regulation> [accessed 16 March 2024]

The Times of India, *'Creep' AI robot touches woman inappropriately on TV in Saudi Arabia*, (The Times of India, 09 March 2024) <https://timesofindia.indiatimes.com/world/middle-east/creep-ai-robot-touches-woman-inappropriately-on-tv-in-saudi-arabia/articleshow/108354164.cms> [accessed 11 March 2024]

Trend Micro, *Cybercriminals*, (Trend Micro website, nd) <https://www.trendmicro.com/vinfo/us/security/definition/cybercriminals#:~:text=

Cybercriminals%20are%20individuals%20or%20teams,personal%20data%2C%20and%20generating%20profit> [accessed 21 June 2024]

Tucker, Patrick, *Refugee or Terrorist? IBM Thinks Its Software Has the Answer*, (USA: Defence One, 27 January 2016), <https://www.defenseone.com/technology/2016/01/refugee-or-terrorist-ibm-thinks-its-software-has-answer/125484/> [accessed 18 January 2024]

Turing, A.M. and Copeland, B.J., *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life, plus the secrets of Enigma*, (Oxford: Clarendon Press, New York; Oxford; 2004, repr. 2005)

Turnitin, *We are Turnitin*, (Turnitin website, 2024), <https://www.turnitin.com/about/> [accessed 27 March 2024]

UK Parliament, *Quantum Technologies*, (Research Briefing, UK Parliament, 24 April, 2017), <https://post.parliament.uk/research-briefings/post-pn-0552/> [accessed 03 April 2024]

UK National Archives, *Online Safety Act 2023*, (UK National Archives website, 23 October 2023) <https://www.legislation.gov.uk/ukpga/2023/50/enacted> [accessed 22 February 2024]

United Nations (a), *Universal Declaration of Human Rights*, (United Nations, 1948), <https://www.un.org/en/about-us/universal-declaration-of-human-rights> [accessed 22 March 2024]

United Nations (b), *AI for Good*, (AI for Good website, nd), <https://aiforgood.itu.int/about-ai-for-good/#:~:text=The%20goal%20of%20AI%20for,United%20Nations%20platform%20on%20AI> [accessed 22 March 2024]

United Nations (c), *What are the Sustainable Development Goals?*, (United Nations Development Programme, United Nations, nd) <https://www.undp.org/sustainable-development-goals> [accessed 12 April 2024]

United Nations (d), *What are Biological Weapons?*, (Office for Disarmament Affairs, United Nations, nd) <https://disarmament.unoda.org/biological-weapons/about/what-are-biological-weapons/#:~:text=Biological%20weapons%20disseminate%20disease%2Dcausing,agent%20and%20a%20delivery%20mechanism> [accessed 22 April 2024]

United Nations (e), *Biological Weapons Convention*, (Office for Disarmament Affairs, United Nations, nd) <https://disarmament.unoda.org/biological-weapons> [accessed 22 April 2024]

United Nations (f), *Chemical Weapons Convention*, (Office for Disarmament Affairs, United Nations, nd) <https://www.un.org/en/genocideprevention/documents/atrocity-crimes/Doc.42_Conv%20Chemical%20weapons.pdf> [accessed 22 April 2024]

University of California San Francisco, *Tobacco CEO's Statement to Congress 1994 News Clip*

"Nicotine is not addictive.", (University of California San Francisco, website, nd), <https://senate.ucsf.edu/tobacco-ceo-statement-to-congress> [accessed 15 April 2024]

Vallance, Chris, *Should we fear an attack of the voice clones?*, (BBC Technology News, 03 February 2024) <https://www.bbc.com/news/technology-68074257> [accessed 05 April 2024]

Valta, Maximilian; Menzel, Johannes; Maier, Christian; Pflügner, Katharina; Meier, Marco; and Weitzel, Tim; *Digital Nudging: A Systematic Literature Review and Future Research Directions*, (In Proceedings of the 2022 Computers and People Research Conference (SIGMIS-CPR '22), Association for Computing Machinery, New York, NY, USA, Article 8, 1–10, 2023) <https://doi.org/10.1145/3510606.3550211> [accessed 08 July 2024]

Visual Journalism Team, *Ukraine in maps: Tracking the war with Russia*, (BBC News website, 17 May 2024) <https://www.bbc.com/news/world-europe-60506682> [accessed 04 June 2024]

von Braun, Joachim, with Archer, Margaret S and others, 2021; *Robotics, AI, and Humanity – Science, Ethics, and Policy,* (Switzerland: Springer Switzerland, 2021, <https://link.springer.com/book/10.1007/978-3-030-54173-6> [accessed 28 September 2023]

Wald, Chelsea, *Precision formation flight astounds scientists*, (Nature Magazine, Springer Nature Limited, 15 January 2014) <https://www.nature.com/articles/nature.2014.14537> [accessed 10 July 2024]

Wallis, Emma, *Fire in former hotel intended to house asylum seekers,* (InfoMigrants, 18 December 2023), <https://www.infomigrants.net/en/post/53968/ireland-fire-in-former-hotel-intended-to-house-asylum-seekers> [accessed 04 January 2024]

Wang, Maya, *China's Chilling 'Social Credit' Blacklist*, (Human Rights Watch website, 12 December 2017) <https://www.hrw.org/news/2017/12/12/chinas-chilling-social-credit-blacklist> [accessed 14 March 2024]

The Week, *Why the undersea cables connecting the world are under threat*, (The Explainer, The Week, website, 22 March 2024) <https://theweek.com/news/technology/955812/undersea-cables-connect-world-subject-concern> [accessed 05 July 2024]

Weinrib, Jacob, *Dignity and Autonomy,* (Max Planck Encyclopaedia of Comparative Constitutional Law, February 18, 2019) <http://dx.doi.org/10.2139/ssrn.3336984> [accessed 28 June 2024]

Wendling, Mike, *QAnon: What is it and where did it come from?*, (BBC News website, 6 January 2021) <https://www.bbc.com/news/53498434> [accessed 11 March 2024]

Whalley, Jason and Curwen, Peter, *Whatever Happened to the Baby Bells? Internationalization and De-internationalization in the Telecommunications*

Industry, (8 MINN. J.L. SCI. & TECH. 149, 2007) <https://scholarship.law.umn.edu/mjlst/vol8/iss1/6> [accessed 05 April 2024]

Whipple, Tom and Blakely, *How AI may want to wipe us out and what we can do to stop it*, (The (London) Times, 03 June 2023)

Williams, Zach, *Conspiracy Theories Give a Boost to State Efforts to Tackle AI*, (Bloomberg Law website, Bloomberg Industry Group,, 11 September, 2023) <https://news.bloomberglaw.com/artificial-intelligence/fears-of-ai-toxins-government-takeover-surface-in-statehouses> [accessed 14 February 2024]

Wilson, Philip K., *Eugenics*, (Encyclopaedia Britannica, 19 June 2024) <https://www.britannica.com/science/eugenics-genetics> [accessed 09 July 2024]

Wilson, Steven Lloyd and Wiysonge, Charles, *Social media and vaccine hesitancy*, (British Medical Journal website, BMJ Global Health 2020;5:e004206, 17 October 2020) <https://gh.bmj.com/content/5/10/e004206> [accessed 03 July 2024]

Winder, Davey, *Tesla Hacked As Electric Cars Targeted In $1 Million Hacking Spree*, (Cybertechnology, Forbes Magazine, 27 January 2024) <https://www.forbes.com/sites/daveywinder/2024/01/27/tesla-hacked-as-electric-cars-targeted-in-1-million-hacking-spree/> [accessed 05 April 2024]

Wolf, Martin, *The Crisis of Democratic Capitalism,* (Allen Lane, Penguin Books, Penguin Random House UK, 2023)

Workplace Relations Commission (Ireland), *Protection of Whistleblowers*, (Workplace Relations Commission - Ireland website, no date) <https://www.workplacerelations.ie/en/what_you_should_know/employer-obligations/protection-of-whistleblowers/> [accessed 21 February 2024]

The World Bank, *GDP (current US$) - European Union (2022),* (The World Bank website, nd) <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=EU> [accessed 02 May 2024]

World Economic Forum, *Explainer: What is neurodivergence? Here's what you need to know*, (Mental Health, World Economic Forum website, 10 October 2022) <https://www.weforum.org/agenda/2022/10/explainer-neurodivergence-mental-health/> [accessed 29 February 2024]

The World Intellectual Property Organisation (WIPO), *What is Intellectual Property?* (The World Intellectual Property Organisation website, 2024) <https://www.wipo.int/about-ip/en/> [accessed 18 March 2024]

Wright, James, *Inside Japan's long experiment in automating elder care*, (MIT Technology Review, 09 January 2023), <https://www.technologyreview.com/2023/01/09/1065135/japan-automating-eldercare-robots/> [accessed 22 March 2024]

Yadav, A., and Jha, A., *Intelligence and Consciousness*, (World Futures, 80(2), pp. 153–173, 2024) <https://www.tandfonline.com/doi/full/10.1080/02604027.2024.2330256> [accessed 25 April 2024]

Yang, Zeyi, *China just announced a new social credit law. Here's what it means,* (USA: MIT Technology Review, 22 November 2022) <https://www.technologyreview.com/2022/11/22/1063605/china-announced-a-new-social-credit-law-what-does-it-mean/> [accessed 18 January 2024]

Zafra, Mariano; Hunder, Max; Rao, Anurag and Kiyada, Sudev, *How drone combat in Ukraine is changing warfare,* (Reuters News Website, 26 March 2024) <https://www.reuters.com/graphics/UKRAINE-CRISIS/DRONES/dwpkeyjwkpm/> [accessed 24 April 2024]

Zauderer, Steven, *Are Autistic People Smart? Autism Geniuses,* (Cross River Therapy, USA), 20 September 2023), <https://www.crossrivertherapy.com/autism/geniuses#are-people-with-autism-smart> [accessed 07 February 2024]

Zimmermann, A., Vredenburgh, K. and Lazar, S., *The Political Philosophy of Data and AI*, (Cambridge: Canadian journal of philosophy, 2022, vol. 52, no. 1, pp. 1-5) <https://www-cambridge-org.may.idm.oclc.org/core/journals/canadian-journal-of-philosophy/article/political-philosophy-of-data-and-ai/B7C38B3D585739080CC94EC48A6E6D7C> [accessed 22 January 2024]

Zvornicanin, Enes, *Why Is ChatGPT Bad at Math?*, (Baeldung website, 11 June 2024) <https://www.baeldung.com/cs/chatgpt-math-problems> [accessed 28 June 2024]

-- END --