

Author(s): Coetzee Bester and Rachel Fischer

The Essential Relationship between Information Ethics and Artificial Intelligence

Abstract:

This article rethinks the position of Information Ethics (IE) vis-à-vis the growing discipline of the ethics of AI. While IE has a long and respected academic history, the discipline of the ethics of AI is much younger. The scope of the latter discipline has exploded in the last decade in sync with the explosion of data driven AI. Currently, the ethics of AI as a discipline can be said to have sub-divided at least into machine ethics, robot ethics, data ethics, and neuro ethics. The argument presented here is that ethics of AI can from one perspective be viewed as a sub-discipline of IE. IE is at the heart of ethical concerns about the potential de-humanising impact of AI technologies, as it addresses issues relating to communication, the status of knowledge claims, and the quality of media-generated information, among many others. Perhaps the single most concerning ethical concern in the context of data-driven AI technology is the rise of new social narratives that threaten humans' special sense of agency and, and this is firstly an IE concern. The article thus argues for the independent position of IE as well as for its position as the core, over-arching discipline, of the ethics of AI.

Keywords: Artificial Intelligence and Autonomous Systems, Cultural Diversity, Ethics of AI, Information Ethics, Intercultural Information Ethics

Outline:

1. Introduction	2
2. Sketching the context of IE and AI	2
3. A Brief history of AI, Ethics and Society	3
3.1. Similarity in the definitions for ethics in AI and IE.....	4
3.2. The relation between AI and IE.....	5
3.3. AI technology and intercultural IE	6
4. The implications of AI on our decision-making and human engagement	7
4.1. The possible danger of manipulated data and algorithms in AI.....	8
4.2. Ownership and responsibility for data-based decisions.....	9
5. Conclusion.....	10
6. References	11

Author(s):

Dr. Coetzee Bester

- coetzee@pamodzicc.co.za

Rachel Fischer

- rachel@3consulting.org

1. Introduction

The debate on whether or not machines can be ethical was highlighted in simplistic social questions presented by the 2004 film called *I, Robot*. Many of the social/ethical questions were asked and eventually not answered. During the past 16 years the *I, Robot* questions and discussion became part of modern reality. During this time philosophically based ethics and Information Ethics (IE) developed to address many ethical questions related to machines and Artificial Intelligence and include amongst others machine ethics, computer ethics, robotic ethics and data ethics. The latest questions in the debate reflect on ethics in AI. Is this a new debate or just another iteration of the field?

Considering the statements of the UNESCO COMEST 2019 report, this article looks at various aspects that address the similarities and differences between IE and Ethics in AI. The understanding of similarities and/or differences seems essential for both information practitioners and end users. To reduce the debate to equating concepts of IE and ethics of AI as being one and the same thing is inaccurate while attempting to force a difference between them just to highlight technicalities is also not conducive to clarity and transparency.

The first part of the following article reflects on the 2019 UNESCO document to understand the platform from which the UNESCO document and views depart. At a very basic and non-technical level the article also looks at comparisons between IE and ethics of AI, particularly from activities related to coding, writing of algorithms, AI design and decision-making. The article concludes with some recommendations to encourage a public debate that serves not only as a philosophical discussion but also as a method to involve AI users and non-specialised scholars.

2. Sketching the context of IE and AI

The question of ethics and machines is not only a technical matter. The scope of the impact of AI technologies covers the whole of the human condition, including, but not limited to, economic, social, political, educational, scientific, legal and healthcare concerns. Over 100 documents on AI ethics have been generated over the past decade or so (AI Ethics Global Inventory). These documents are formulated by intergovernmental entities such as the EU Commission and UNESCO; by the private sector such as Microsoft, and the Future of Life (e.g., Microsoft's 'Responsible AI Principles'), by professional bodies such as the IEEE (e.g., their 'Ethically Aligned Design' document), and by academia and research institutions such as the Berkman Klein Centre at Harvard, the Nuffield Centre for Ethical and Technologies, or the Rathenau Institute in the Netherlands (e.g. the Leverhulme Centre for the Future of Intelligence's report on the 'Ethical and Societal Implications of Algorithms, Data and AI'). Recently even the Vatican called for stricter ethical standards on the development of artificial intelligence, with tech giants IBM and Microsoft being the first companies to sign its new initiative.

In 1452, with Gutenberg's invention of the printing press, the management of information changed. Not only did this invention enable the mass production of written texts, but it also created new opportunities for the translation and distribution of texts, and establishing a 'book culture' which, until the introduction of digital technology gave large numbers of people not only access to information but also the opportunity to share their own ideas. The accessibility of written information became even stronger when libraries, formerly established by the Church and/or the nobility, were opened to the public after the French Revolution. However, with all these wonderful contributions, the printing machine could not interpret information, but could only multiply what was given to it by humans (Britz, 2013).

When referring to a broader, or more popular, application of the term 'ethics', it is typically used either in relation to a code of conduct, or code of ethics, indicating the kind of behaviour regarded as appropriate in specific professions, institutions or organisations, or as the terminology of a specific academic field or sub-field of study. In the latter case, according to Roget's Thesaurus, it refers to the "science" or "philosophy" of "morals" (Roget, 1987), often in disciplines like Philosophy and Theology. Regardless of the field of study, the focus of ethics is the critical analysis of and debate on what it is that should be considered the right thing to do in terms of the way in which human beings live their lives and relate to one another in the face of moral dilemmas (Britz, 2013; Rossouw & Van Vuuren, 2004).

In information societal contexts, Ocholla (2013) teaches that ethics serves three purposes: (i) to "promote what is good in people", (ii) to "avert chaos", and (iii) to "provide norms and standards of behaviour" that are "inclusive" rather than "exclusive". This should be regardless of their cultural, religious, racial, gender,

or other differences. All human beings are therefore morally obliged to treat all other human beings justly and with the same respect with which they would those belonging to their own inner circle, community, society or nation (Mutula, 2013). In this specific context, however, the article is concerned with IE as an example of applied ethics, where IE refers specifically to the academic application of ethics in regard to information and communication technologies. A definition that elucidates the broad application of this field, states that it is:

[...] It is the field that investigates the ethical issues arising from the development and application of information technologies. It provides a critical framework for considering moral issues concerning informational privacy, moral agency (e.g. whether artificial agents may be moral), new environmental issues (especially how agents should one behave in the infosphere), problems arising from the life-cycle (creation, collection, recording, distribution, processing, etc.) of information (especially ownership and copyright, digital divide.

(Le Sueur, Hommes & Bester, 2014)

3. A Brief history of AI, Ethics and Society

In their 2019 Report UNESCO acknowledged that the world is facing a rapid rise of AI technologies and that machine learning algorithms have the capacity to learn and to perform reasoning tasks that used to be limited to human beings. According to the UNESCO (2019) report, we should accept that in the life cycle of information, machines can create, generate, store, interpret, repackage, distribute, protect or destroy information, which implies that data-driven AI is already human-like in their role in information societies. From this view, technological development is likely to have substantial societal and cultural implications.

Based on the perceived social implications, many individuals, institutions and governments and the European Commission are concerned about the ethical implications of AI. As aligned with the World Summit on Information Society (WSIS), UNESCO (2019) took responsibility for the implementation for a number of objectives including Action Lines on Access (C3), E-Learning (C7), Cultural diversity (C8), Media (C9), and Ethical dimension of the information society (C10). UNESCO has two approaches to the matter of AI ethics; (i) a human rights framework and, (ii) a framework of the Internet Universality which works on the UNESCO IFAP platform with IE. It thus seems that ethics in AI and IE is joined in meaning in most critical elements of society (UNESCO, 2019).

The WSIS 2003 message was repeated in October 2004 at a symposium on IE at the Center for Art and Media in Karlsruhe (Germany). This event included 45 delegates from 19 countries across the world were scholars and practitioners respectively schooled in computer science (informatics), computer engineering, library and information science, software engineering, philosophy, law, and management came together. These scientists stated that the focus should be on both the theoretical dimensions and practical applications of computer science and programming language. They concluded that IE should be regarded as something that affected all aspects of human life (Froehlich, 2004).

According to the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems research on Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing, technologists are encouraged to prioritise ethical considerations in the creation of autonomous and intelligent technologies. The following statement in the executive summary indicates the crux of the project:

To fully benefit from the potential of Artificial Intelligence and Autonomous Systems (AI/AS), we need to go beyond perception and beyond the search for more computational power or solving capabilities. We need to make sure that these technologies are aligned to humans in terms of our moral values and ethical principles. AI/AS have to behave in a way that is beneficial to people beyond reaching functional goals and addressing technical problems. This will allow for an elevated level of trust between humans and our technology that is needed for a fruitful pervasive use of AI/AS in our daily lives. Eudaimonia, as elucidated by Aristotle, is a practice that defines human wellbeing as the highest virtue for a society. Translated roughly as "flourishing," the benefits of Eudaimonia begin by conscious contemplation, where ethical considerations help us define how we wish to live. By aligning the creation of AIS with the values of its users and society we can prioritize the increase of human wellbeing as our metric for progress in the algorithmic age.

(IEEE, 2019)

The IEEE's *Ethically Aligned Design* project established various committees to research and contribute to the project, one of which is of central importance in this instance. The committee on "Classical Ethics in Information and Communication Technologies" (Chatila & Havens, 2019) has taken to task the ethical considerations as they relate to AI/AS.

3.1. Similarity in the definitions for ethics in AI and IE

According to Metcalf (2014), ethical codes are mostly written in response to conditions that present themselves in practice from time to time. The most influential ethics codes are hard-won responses to major disruptions, especially medical and behavioural research scandals. Such disruptions re-open questions of responsibility, trust and institutional legitimacy, and thus call for the codification of new social and political arrangements (Metcalf, 2014). In terms of IE and the digital environment, an exponential increase in available digital technologies has led to disruptive changes in human behaviour, knowledge practices and communication.

In 2019 UNESCO noticed the increase in the recent number of declarations of ethical principles on AI and identified the need for a global standard-setting instrument on the ethics of AI. These declarations should accommodate global standard-setting towards pluralistic, multidisciplinary, multicultural and multistakeholder platforms and should be compatible with internationally agreed human rights and standards and be aligned to a human-centred vision (UNESCO, 2019).

Concepts of AI refer to "artificially created" and "intelligent" beings, machines, or tools. Such definitions, according to UNESCO (2019), contribute to a unique fascination with the anthropomorphic possibilities of AI – including its ethical dimensions as amplified by its development and real-world applications. Per definition AI is one of the central issues of the era of converging technologies with profound implications for human beings, cultures, societies and the environment. AI is likely to transform the future of education, science, culture and communication (UNESCO, 2019).

The UNESCO executive at its 206th meeting detailed their view relating to the desirability of a standard-setting instrument on ethics of artificial intelligence. They identified the ethical-global dimensions of peace, cultural diversity, gender equality, and sustainability for AI to contribute to peace, highlighting that AI should never operate outside of human control. This view is supported by the statement that AI is engineer-oriented, consists of the capability to reason, that it includes machine learning, deep learning, computer vision and robotics (Russell & Norvig, 2016) and that it is capable of independently performing tasks that would otherwise require human intelligence and agency (UNESCO, 2019). AI has already left its impact on human life in areas such as transport, medicine, communication, education, science, finance, law, military, marketing, customer services and entertainment, each of which have posed unique ethical concerns. It is clear that at the moment, no AI system can be considered as a general-purpose intelligent agent that can perform well in a wide variety of environments, which is a proper ability of human intelligence (UNESCO, 2019).

According to Britz (2010), the extent of and the ways in which Information and Communication Technology (ICT) supports the different information life cycle activities in society and the workplace play a pivotal role in the shaping, understanding and defining of IE (Britz, 2010). In this context, IE could be defined as the "broader domain of professional ethics", encompassing the ways in which professionals – that is, "the creators/distributors of information products and services, information mediators (including librarians)" – and the general public "engage with, respond [...] and react to [...] ethical issues arising from the use of digital technologies" (Britz, 2010).

Gadamer (1975) said it is only when a person is able to make sense of information that s/he can be regarded as 'knowledgeable'. One could argue that, in order to 'convey knowledge' the information at one's disposal first has to be processed, understood, internalised, and then disseminated and/or put to other uses. The final non-tangible 'product' could then be called 'knowledge' (Bester, 2018). In this instance IE is an example of applied ethics since it is concerned with the sphere of Information and Communication Technologies (ICTs), the life-cycle of information and social justice concerns brought about by ICTs. However, when confronted with an ethical dilemma, one uses normative ethics to guide one's decision making (Bester, 2018).

Similarly, IE is a descriptive and emancipatory discipline dealing with the study of the changes in the relationship between people and the world due to the impact of information and communication technologies. (Britz, 2007). Further to his 2007 remarks Britz (2010) added that IE define the basis of a discourse that focuses

on moral questions related to the life cycle of information, that is, to its generation, gathering, organization, storage, retrieval, and use of digital information (Britz, 2010). As if describing AI, Britz (2010), describes IE as the extent of and the ways in which Information and Communication Technology (ICT) supports the different information life cycle activities in society. He argues that the workplace plays a pivotal role in the shaping, understanding, and defining of IE.

3.2. The relation between AI and IE

While the concept of IE started with the impact of the Gutenberg press in 1450, it was undeniably the invention and use of the computer that served as the impetus towards the development of what is now referred to as 'IE'. The 2003 WSIS Declaration on Information Societies not only speaks to people's technological competence but also the ways in which people should and should not use information communications technologies (ICTs). Premised on the purpose and principles of the Charter of the United Nations and the upholding of the Universal Declaration of Human Rights" (UDHR), the WSIS Declaration emphasizes that, while free access to the technologies of the digital world (broadband and / or computer networks) is critical to the economic development of governments, businesses and individuals, it should under no circumstances be used for abusive purposes.

When Wiener (1948) coined the term, 'computer ethics', he did not define it. It was, however, only in the early 1970s, with the dawning of the Information Age, that "techno-rebels" (Toffler, 1980) increasingly raised concerns about the potentially harmful impact of computer technology on social values and human behaviour gained momentum. These rebels, most of whom worked in the computer industry, realising the opportunities for the beneficial and harmful use of computer technology, no longer focused only on the benefits and dangers inherent in the use of computers for economic gain and/or military clout, but also on the potential effect that their use could have on the ecology, society and humanity as a whole (Toffler, 1980). Arguing that the "fragile biosphere of Planet Earth" could be destroyed by the irresponsible use of computer technologies, they urged people not to become so dependent on these technologies that it eventually ruled their lives. In addition to this, they advocated the development of technologies that were environmentally and human-friendly, used to advance not only the lives of those with economic or political power but of all people (Toffler, 1980).

According to Capurro (2007), there is no such thing as morally neutral technology. Since all technologies create new ways of doing as well as new ways of being, they tend to either positively or negatively affect existing value systems and beliefs, often replacing it with their own (technological) culture and values (Mutula, 2013). It follows that any discussion of IE issues must at least take cognizance of the ways in which modern information communications technology has changed not only the information and knowledge landscape but also the attitudes, values and behaviour of the people whose lives are affected by it (Britz, 2013). More specifically, discussions should focus on contentious issues in terms of different views based on culture, geographical location, literacy levels, and the development status of the country or region concerned. Capurro (2007) also listed key issues of IE and the sustainability of information and knowledge societies include issues arising from access to and accessibility of information, plagiarism, copyright, privacy, safety and security, information poverty and overload, E-waste, and tensions arising from the perceived imposition of global/Western philosophies and values on inhabitants of other parts of the world.

Like IE, the focus of ethics in AI should include the way in which privacy is conceptualized differently by different culture and throughout different time periods. While the concept of privacy is closely related to the self in Western cultures, in others, like the Buddhist cultures it relates to the "non-self", where the idea of privacy, as an adjunct of compassion, becomes quite plausible (Hongladarom, 2007). In other cultures, African cultures, for example, privacy relates to the 'collective self'. Consequently, understandings of, social perceptions of and/or interpretations of privacy will be different (Nakada & Tamura, 2005; Capurro, 2005; Capurro et al, 2013), as will the laws have aimed at its protection. In the future, all information practitioners might even agree that, except for the terminology used, ethics of AI and IE are the same.

The shift from information and knowledge poverty to sufficient information and knowledge equity requires the bridging of what is commonly referred to as the digital divide, a "popular concept or phrase used to explain the inequality of information access and use, largely with respect to ICTs within or between individuals, families, communities, nations and regions" (Ocholla, 2013). AI can help to solve the problems of information poverty and information overload but not without human oversight of the process, while access

and accessibility are norms that are foundational to IE. AI can inform efforts to address information overload and can assist in the complexities of best to overcome the digital / information divide. Factors causing the digital divide include, but are not limited to, education and income levels, unemployment, infrastructure, values assigned to information and the cost of access to information (Lievrouw, 2000; Ocholla, 2013). Also, according to Habermas (1989) other factors include "lack of access to and/or use of modern digital technologies, users' inability to access the kind of information for human development and prosperity, and/or deficiencies in the quality of available information".

Information overload, representing the other side of the information access spectrum, can be just as detrimental to individual and societal development as information poverty is. Too much information can increase users' stress levels and can negatively affect one's physical and psychological health (Britz, 2013). This is not only true for individuals but also for institutions, communities and nations. The volumes of data and information as well as the collection, selection, categorizing, packaging, distribution and storage all forms part of the possible dangers of information overload. Further to this, is what can be called an overload in the secondary phase in the use of data and information which includes activities like analysis, meaning, interpretation, relevance and causality of different sets of data and information. Related to these activities are the expressed ethical concerns in the management of information by selection of specific AI and its programmed algorithms.

3.3. AI technology and intercultural IE

IE guides our concerns related to information poverty, information overload, digital divides, gender discrimination and censorship as well as tensions arising from calls for universal IE on the one hand and intercultural IE on the other. In the field of IE these issues are regarded as objects requiring ethical scrutiny, not only with regard to universal rights and principles but also with regards to the acknowledgement of and respect due to cultures other than those of the Western World (Capurro, 2008). In consideration then of how different cultural traditions have their own moral questions as they relate to ICTs, Intercultural IE (IIE) endeavours to investigate different approaches to IE (Capurro, 2006).

The WSIS Action Line C10 on IE drives the values which information societies should adopt. These values should be 'universal' but there are those who perceive them as an attempt to 'globalise' the world, assimilating and, by implication, destroying not only the values of non-Western nations but also their languages and cultures. In this regard, Mutula (2013) argues that the focus of universal values should be on the promotion of the common good and the prevention of ICT abuse. Capurro and Ocholla (2013) warn that historical and geographical singularities give rise to different kinds of theoretical foundations and practical options and that it is inevitable that with the "global penetration of the Internet and the mobile phone", which are fast merging into a single device, many of the world's cultures could clash with the culture imposed by digital technologies. Hence Mutula (2013), while advocating universal values aimed at the promotion and protection of the common good, also advocates for values that support the respect, preservation, promotion and enhancement of cultural heritage and diverse forms of digital and traditions media.

The key issue here is whether it is possible for different human cultures to survive/flourish in a global digital environment without risking isolation (Capurro, 2007 and Capurro, 2013). This pressure, according to Capurro (ibid), should not be considered merely as a problem of technical access to the Internet but also of how people can better manage their lives, using new interactive digital media while avoiding the danger of cultural exploitation, homophobia, colonialism, and discrimination.

Further to this, there are initiatives to integrate AI with the human brain using a "neural interface": a mesh growing with the brain, which would serve as a seamless brain-computer interface, circulating through the host's veins and arteries (Hinchliffe, 2018). From the perspective of both ethics in AI and IE, this technological development has important implications for the question of what it means to be human, and what "normal" human functioning is.

Related to ethics in AI, these concerns are not merely of academic value, as signified in a 2007 Microsoft report (Harper, Rodden, Rogers & Sellen 2008) which warns that by 2020 "new technologies" could promote "new forms of control or decentralisation, encouraging some forms of social interaction at the expense of other, and promoting certain values while dismissing alternatives". According to this report, AI technology could be used as to develop a culture of "urban indifference", "addiction to social contact" and "subvert traditional forms of governmental and media authority".

Related also to the Microsoft Report, Capurro (2013), confirms that the primary task of digital IE is to serve as a critical and ongoing “interdisciplinary and intercultural reflection on the transformation of humanity through computer technology, emphasizes the value of inclusive, inter-cultural information and knowledge societies”.

The Microsoft Report concludes by reiterating a sentiment that lies at the heart of IE, namely; the importance of interpreting the cultural implications in terms of a “wider context” than the “technical capabilities” of “neural networks, recognition algorithms and data-mining”. It is thus critical to understand the importance of remembering that “computer technologies are not neutral. As a matter of fact, they are laden with human, cultural and social values”. This is also acknowledged by the mentioned recent interaction between the Vatican, IBM and Microsoft (Sonnemaker, 2020).

4. The implications of AI on our decision-making and human engagement

AI means a machine-driven ability to imitate the decision-making process of the human mind by collecting and processing data, and then responding to that data according to pre-set principles for decision-making in anticipation of a specific outcome – a sequence of consecutive steps otherwise known as an algorithm. AI provides a standard ICT-based facility that allows for the collection or acquisition of data via a powerful capacity to store, process and communicate the information in various formats. AI technology is based on three critical components; (i) availability of huge and dynamic huge data sets, (ii) extremely fast processing and (iii) clear algorithmic instructions to re-order, reinterpret and represent the available data as information. The algorithms behind the decisions capacitate AI through instructions (algorithms) to obey and try to optimize, given the data it is provided with (UNESCO, 2019).

According to Ashley (2017) it is highly questionable that AI will soon have the capacity to manage and interpret ambiguous and rapidly evolving data and then execute what human intentions would have been if the human could have coped with complex and multifaceted data. Ashley also concluded that cognitive AI does not make decisions in the same way as humans would since it uses a series of algorithms to select conclusions where algorithms learn rules based on statistical regularities. Ashley argues that “although the machine-induced rules may lead to accurate predictions, they may not be as intelligible or as reasonable” (2017).

Based on Ashley’s insight, it seems important to refer to AI outcomes as ‘results’ of a process rather than as ‘decisions’. The concerns with algorithmic-created AI results are increasingly important as the role of AI and algorithmic capabilities have an impact on news and social media and eventually on issues of access to information, disinformation, discrimination, freedom of expression, privacy, and media and information literacy. These concerns are even bigger if the influence of these algorithmic processes cause or exasperate digital divides between communities, social groups and countries (UNESCO, 2019).

But who are setting the AI filters? According to the IEEE (2019), an organisation of more than 400,000 electrical engineers globally, the ethics of technology has developed various methods to bring ethical reflection, responsibility and reasoning to the design process. In the context of AI, the term “ethically aligned design” (EAD) has been developed to indicate design processes that explicitly include human values. EAD compliant AI developers could also consider other ethical issues such as the prevention of algorithmic bias, minimizing the ability to misuse the technology, and clarification of algorithmic decisions.

From an IE and ethics in AI perspective, one could question the ultimate aim of the collection, filtering, organising and analysis of data: is the aim to conclude which data is usable and what is considered useless? If algorithms prefer certain technology, data or information while rendering others useless or not relevant, then again, the initial human writers of algorithms and the designers of AI machinery are manipulating the value of future data and information.

UNESCO (2019) stated that the proliferation of AI has inaugurated substantial societal and cultural changes, raising issues of freedom of expression, privacy and surveillance, ownership of data, bias and discrimination in algorithms, manipulation of information and trust – which aligns with concerns raised by IE. Biased algorithms, when unchecked, can act as a reality filter, and can inadvertently help reinforce and spread disinformation which in turn influences perceived and established standards of “fact” and “truth”, a concern that has implications for all spheres of meaning, society and human life itself. This phenomenon becomes

particularly concerning in the case of bias in AI deep learning where one level of bias informs the next stage of machine 'learning' and the AI's own interpretation defines newly biased perceptions of reality. Furthermore, due to its magnitude machine learning can exponentially increase bias, inequality, exclusion and create a greater threat to cultural diversity and equity than the original human bias did. As such, the scale and the influence caused by AI technology could soon enhance the digital divide and inequality between individuals, groups and nations. These results may not necessarily be based on reality at all but could be the result of concentrated biases in algorithms for machine learning, data classification and computational resources for storage and processing of data and information. As a result, AI requires careful analysis to address its implications for culture and cultural diversity, education, scientific knowledge, and communication and information. The impact of these matters drives the concerns for global orientation, the global-ethical themes of peace, sustainability, indigenous communities and gender equality (UNESCO, 2019).

4.1. The possible danger of manipulated data and algorithms in AI

AI is likely to have an extreme impact on the natural, social, and environmental sciences, impacting the way that we understand the very foundations of life itself. Just as concerning are the implications for how the combination of AI and human interfaces will interpret knowledge in social contexts. Increasingly powerful speed and systems of machine and deep learning have created challenges in existing conceptions of satisfactory scientific explanation, such as the lack of transparency in black-box analyses (Mayer-Schönberger & Cukier, 2013). Based on extremely powerful data-analytical capacities, AI can produce impressively accurate predictions based on data sets without providing the human observer with any causal (cause and causality) or unifying explanation of its predictions (UNESCO, 2019). This could have implications for our trust in science, since the human observer may no longer be able to replicate the processes that determine the outcome. The legal position described by Shefet (2020) above offers further insight here.

Thus, the successful proliferation of machine learning algorithms increases the likelihood of the delivery of comparable results apart from a scientifically justified model and could have implications for the public perception and evaluation of science and scientific research. This is particularly concerning when the quality of machine learning (also deep learning) depends heavily on the quality and integrity of the available data used to train the algorithms (UNESCO, 2019).

Throughout history, humanity has both embraced and simultaneously feared technological development, and while it could be argued that the human species evolved specifically because of technological development, a fear and suspicion of all new technologies in all areas of development has accompanied technological innovation from the beginning. This relationship is especially true in the Information era where the ubiquitous deployment of digital technologies is all encompassing, even while for the first time in history, we, the creators, begin to lose oversight of our creation.

The digital era has seen multinational tech companies begin to invest in the application of data-driven AI in all their products. Through this process, computing power has become large enough to manage complicated algorithms and analyse "big data" that can then be used for machine learning. According to a 2019 UNESCO study on ethics in AI these companies have access to almost unlimited computing power and to data collected from billions of people to 'feed' AI systems as learning input. Due to the influence of AI in people's daily lives and in professional fields like healthcare, education, scientific research, communications, transportation and security, AI raises new concerns that could affect the trust and confidence people have in these technologies. Lack of trust or concerns range from the risk of criminality, fraud, harassment, hate speech and discrimination to the spreading of disinformation. According to UNESCO (2019) the growing lack of trust in AI is based on our inability to understand its processes, a mistrust that is exacerbated by the absence of transparency in algorithms and the various influences that determine the outcomes of AI systems.

The term, 'Information Age' is applied to the current historical era spanning the first quarter of the twentieth century to date (Le Sueur, Hommes & Bester, 2014). Implied in this terminology are the values attached to the availability of information, the emergence of an information economy, the development of information societies, and the inter-connectivity of people and digital technologies like AI. The Information Age could therefore be defined as a time / period during which information not only became easily accessible through publications and the manipulation of information by computers and computer networks (WordNet.princeton.edu), but also by the increasing use of and reliance on technology to generate, access, process, use and disseminate information (Kawooya, 2004). Put differently, the emphasis on IE and Ethics in

AI within the information sphere are based on the importance of information and/or the dominance of information-based goods and services by information workers in the economic sphere.

The key factor in the onset of the "info-sphere" created by the technologies emerging from these fields was the advent of the computer, a "combination of electronic memory, with programs that tell a machine how to process [...] stored data". Subsequent developments in computing technology transformed the original computer into machines with sufficient capacity to be "deployed for a variety of tasks in corporate headquarters" and later into the "small, powerful mini-computers" with which we are now familiar. Currently computers can do just about anything people want them to, making information from all over the world available to whoever has access to them, and opening multiple opportunities for governments and other parties to track people's movements, even over vast distances. Not only are computers able to "synthesize" information, but also to anticipate the consequences of actions taken by human beings at any moment (Toffler, 1980).

4.2. Ownership and responsibility for data-based decisions

As stated earlier, data-driven AI is the extreme ability to accumulate, analyse, interpret, and repackage data. Debates on the meanings respectively attached to the terms, 'data', 'information' and 'knowledge' are not common. Neither is research on the nuanced differences between them nor of the extent of the impact that they could have on a person's understanding and interpretation of events and circumstances, especially in the absence of the average person's understanding of the range of their influence. And even though researchers, statisticians and those working in the information technology field are familiar with the various nuances of the term, 'data' and with its specific meaning in their particular context, the term has to date seldom featured in common parlance around information and computing, being subsumed instead in the terminology of 'information' or 'knowledge'. Britz (2013) agrees with Gadamer (1975) that knowledge cannot be acquired without understanding but adds to the argument that understanding does not develop "by its own volition", but only through education or training. These are significant principles in discussing responsibility and ownership related to ethics in AI and IE in general.

Laws regulating the right to 'copy' another person's work, known as Copyright (laws) were introduced in Britain in the 18th century, primarily as a means of curbing/restricting the monopolies held by printers and/or to provide authors with some form of short-term legal protection (Britz 2013). With the ever-widening scope and application of these laws, the term, 'copyright', was replaced/supplemented by a new legal concept, namely "intellectual property right". Whereas copyright laws regulated the creation of intellectual products only, intellectual property right laws also regulated access to and the distribution of written texts (Britz, 2013). Rossouw and Van Vuuren (2013) indicate that there are certain responsibilities which all humans are bound to adhere to in terms of their "moral obligation" and that the extent to which and/or the manner in which these responsibilities are carried out signifies whether or not the person, society, profession or organisation is acting ethically or not. Britz (2013) applied this to the ethics of information and concluded that this would mean that every person involved in the generation, processing and use of information must do so responsibly. The question is whether this responsibility, in terms of the protection of ownership and dignity of any of the parties involved, is also applicable to AI? Can AI be held responsible for the maintenance, or contravention, of ownership and copyright?

In creating and writing algorithms, the consequentialist position, is somewhat different in the sense that it evaluates the consequences of an action, not the responsibility of an agent. In applying Britz's argument it is not the intention informing an action but its outcome which determines whether it is ethical or not. Informing this theoretical position is the question, 'For whom is the outcome good?' Put differently, 'Who benefits from the action or outcome of the action?' Regarding the ethics of information, the implication is that the final information product – the outcome of the information process, should benefit as many parties as possible. The creation, use and dissemination of information that is false, sub-standard, or promotes dissent or fear would therefore be regarded as immoral / unethical (Britz, 2013).

In this regard, UNESCO (2019) recognised that copyright is becoming a complex issue, since AI-generated content could in the future be less influenced by human input resulting in a copyright that will be attributed to the algorithms themselves. These challenges related to copyright, are still in place regardless of the fact that the United Nations declared freedom of speech, freedom of access to information, and freedom of the press as universal and basic human rights in 1984, and the inclusion of these rights as principles informing the Constitutions of 'democratic' nation states (Britz, 2013:3). Whereas copyright and intellectual property laws

have the protection of writers in mind, other laws have the preservation of traditional/cultural values or morals as their purpose. The censoring of information – deleting pieces deemed 'unsuitable', 'inappropriate', 'immoral', or 'politically incorrect' from texts, banning the release of texts in their entirety, placing embargoes on the release of information, and/or denying the public access to information regarded as 'sensitive' or 'need to know', are examples (Bester, 2018).

Conversely, Shefet (2020) refers to the GDPR in the European Union (EU), and the ability to remove any reference to an individual from a data set, whether governmental, commercial, or in the custody of some other organisations that manage or store data. The General Data Protection Regulation 2016/679 (GDPR) became applicable in all EU countries on 25 May 2018. The main purpose of the GDPR is to strengthen and unify the protection of personal data. Based on programmes, monitoring and automatic management of private information, many already consider Facebook as a particular kind of AI. Various social media providers, being aware of the dangers inherent in the use of social media, therefore expect users of their services to adhere to some basic standards of ethical behaviour. Facebook, for example, has a Statement of Rights and Responsibilities as well as User Privacy Policy in place which governs their relationship with users and others who interact on their social media platform. The privacy policy provides guidelines on proper interaction between users on the one hand and the ways in which Facebook may collect and use clients' information and content. However, according to Mutula (2013), the mechanisms they have in place to enforce compliance to these standards remain weak. The collection, interpretation and reprocessing of clients' data, information and social media postings by way of machine managed processes is already AI in operation.

5. Conclusion

In conclusion, it must be reiterated that the debate on ethics in AI is not a new debate on new concepts but a continuation of the in-depth discussions on the same topics in IE during the past decade.

As indicated, the focus of this article was to look at and explain the relationship between ethics of AI and IE in general. It has become increasingly clear that ethics in AI should not be separated from IE as that will unnecessarily complicate the debate and ethical guidelines for end users in the so-called Information Society.

Simplified and clear definitions should focus on the impact which digital technologies and ICTs have on society and the environment. The focus of these definitions must therefore remain attached to the ethical issues related to the use of the Internet, digital information and communication media as well as on the responsible use of ICTs in Information Societies as described by the WSIS (Mutula, 2013). In a broader sense, IE is concerned with information and communication beyond just digital media (Capurro, 2013).

At a basic level, AI is nothing more than an extremely powerful reworking of big data as part of the information life cycle to address various human-identified objectives (needs) and human-influenced methods (algorithms). Ethics in AI is, for the time being, just another feature of IE and should be directed by the same guidelines. For now, we should not allow that ethics of information become blurred by AI but rather guided by humanity. AI need not to be a new scary technological era nor is it in itself a new phase, it will develop like any technology, however powerful.

Based on the mentioned conclusions some practical recommendations include that one should not overcomplicate the matter of ethics in ICT – during the past years the focus of IE has expanded from time to time, but the basic science and guidelines that were developed during the past decade, will be able to guide ethics in AI. In addition to the existing IE guidelines for schools and training institutions, specific skills for training of learners in coding, should be formulated in order to ensure basic sensitization to ethics considerations of current and emerging ICTs. Training of creators of algorithms should receive intensive training in IE due to the vast technological, social, cultural, economic and political implications on humanity. And finally, although it was not part of this article, research on matters related to cultural diversity in the training of learners interested in coding and creators of algorithms, should be promoted. There is a gap in the field of IIE and how it relates to ICTs, and therefore promoting cultural diversity will essentially promote the dialogue on value pluralism and how it is applied to how ICTs shape and influence our world.

6. References

- Ashley, K.D. *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge, Cambridge University Press, 2017.
- Bester, B.C. *Development of a Framework for Teaching IE to Various Communities in Southern Africa*. Unpublished thesis for PhD Information Science in the Department of Information Science, Faculty of Engineering, Built Environment and Information Technology, University of Pretoria, 2018.
- Britz, J.J. *The joy of sharing knowledge: But what is there is no knowledge to share? A critical reflection on human capacity building in Africa*. Published in the *Africa Reader on IE* (page 15 to 22). ISBN: 978-0-620-45627-2. Department of Information Science. University of *Pretoria. South Africa. First electronic edition:2007, available online at http://www.africaninfoethics/african_reader.html. First printed edition: 2007
- Britz, J.J. *Understanding IE*. In Dennis Ocholla, Johannes Britz, Rafael Capurro and Coetzee Bester (Eds): *IE in Africa: Cross-cutting Themes*. Pretoria. Groep 7 Drukkers, 2013.
- Britz, J. and Buchanan, E.A. *Ethics from the bottom up? Immersive ethics and the LIS Curriculum*. *Journal of IE*, 19(1):12-16, 2010.
- Britz, J.J. *To Understand or not to Understand: A Critical reflection on Information and Knowledge Poverty*. Dennis Ocholla, Johannes Britz, Rafael Capurro and Coetzee Bester (Eds): *IE in Africa: Cross-cutting Themes*. Pretoria. Groep 7 Drukkers, 2013.
- Sonnemaker, T. *The pope has joined with Microsoft and IBM to create a doctrine for ethical AI, 2020*. Available at: <https://www.businessinsider.com/microsoft-ibm-pope-francis-push-for-ai-principles-2020-2> and <https://www.reuters.com/article/us-vatican-artificial-intelligence/pope-to-endorse-principles-on-ai-ethics-with-microsoft-ibm-idUSKCN20M0Z1>.
- Capurro, R. *Privacy. An intercultural perspective*. In *Ethics and Information Technology*. 7(1):37- 47, 2005.
- Capurro, R. *Towards an ontological foundation of IE*. *Ethics and Information Technology*, 8(4), 175–186, 2006.
- Capurro, R. *IE for and from Africa*. Published in the *Africa Reader on IE*. (page 3 to 14). ISBN: 978-0-620-45627-2. Department of Information Science. University of Pretoria. South Africa, 2007, available online at http://www.africaninfoethics/african_reader.html.
- Capurro, R. *IE in the African Context*. In Dennis Ocholla, Johannes Britz, Rafael Capurro and Coetzee Bester (Eds): *IE in Africa: Cross-cutting Themes*. Pretoria. Groep 7 Drukkers, 7-16, 2013.
- Chatila, R., Havens, J.C., 2019. *The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems*, in: Aldinhas Ferreira, M.I., Silva Sequeira, J., Singh Virk, G., Tokhi, M.O., E. Kadar, E. (Eds.), *Robotics and Well-Being*. Springer International Publishing Boek-titel in italics, Cham, pp. 11–16. https://doi.org/10.1007/978-3-030-12524-0_2
- Dick, A. L. *The Philosophy, Politics and Economics of Information*. Pretoria: Unisa Press.
- Fallis, D. *IE for the twenty-first century library professionals*. *Library Hi Tech*, 25(1): 23-36, 2007.
- Fischer, R. *Normative Ethics Theories towards understanding decision-making in IE and Digital Wellness*. Pretoria: African Centre of Excellence for IE (ACEIE), 2017.
- Froehlich, T. *A brief history of IE*. *Textos universitaris de biblioteconomia i documentació*. Available from <http://www.ub.es/bici/13froel2.htm>
- Gadamer, H.G. *Truth and method*. Translated by W. Glen-Doepi. Stead and Ward. London, 1975.
- Habermas, J. *The structural transformation of the public sphere*. Translated by T. Burger. Cambridge, MA. MIT Press, 1989.
- Harper, R., Rodden, T., Rogers, Y., & Sellen, A. (2008) *Being Human: Human- Computer Interaction in the Year 2020*. Microsoft Research, Cambridge, U.K., 2008. <https://cacm.acm.org/magazines/2009/3/21785-reflecting-human-values-in-the-digital-age/fulltext>, 2008.
- Hinchliffe, T. *Medicine or poison? On the ethics of AI implants in humans*, *The Sociable*. Online. Available at: <https://sociable.co/technology/ethics-ai-implants-humans/House of Lords>, 2018.
- I, Robot. From Wikipedia, the free encyclopedia 20th Century Fox, 2004 (Mann Village Theater) <https://www.google.com/search?q=i+robot+movie&oq=i+&aqs=chrome.0.69i59j69i57j0l6.29587j0j8&sourceid=chrome&ie=UTF-8>

- Le Sueur, C., Hommes, E. and Bester, C. (eds.) *Concepts in IE: an introductory workbook*. Pretoria. Africa Centre of Excellence for IE. Department of Information Science. University of South Africa. Pretoria. Groep 7 Drukkers, 2014.
- Lievrouw, L.A. *The information environment and universal service*. In *The Information Society Vol. 16*:155-160, 2000.
- Kawooya, D. *The Digital Divide: An Ethical Dilemma for Information Practitioners in Uganda?* In T. Mendina and JJ Britz (eds): *IE in the Electronic Age: Current Issues in Africa and the World*. Jefferson, North Carolina. McFarland 28-35, 2004.
- Mayer-Schönberger, V., & Cukier, K. *Big Data: A Revolution that Will Transform how We Live, Work, and Think*. Houghton Mifflin Harcourt, 2013.
- Metcalf, J. *Ethics, Codes: History, Context and Challenges*. Draft version produced for Council for Big Data, Ethics, and Society, 2014.
- Mutula, S.M. *Ethical Dimensions of the Information Society: implications for Africa*. In *IE in Africa: Cross-cutting Themes* by D. Ocholla, J.J. Britz, R. Capurro, and C. Bester (eds): 29-43, 2013.
- Nakada, M. and Tamura, T. *Japanese conceptions of Privacy: An intercultural perspective*. In *Ethics and Information Technology*, 7:27-36, 2005.
- Ocholla, D.N. *What is African IE?* In *IE in Africa: Cross-cutting Themes* by D. Ocholla, J.J. Britz, R. Capurro, and C. Bester (eds): pp. 21-28, 2013.
- Rossouw, D. & Van Vuuren, L. *Business Ethics, Fifth Edition*. Cape Town: Oxford University Press, 2013.
- Russell, S.J. and Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*, 3rd ed. Harlow, Pearson, 2016.
- Shefet, D. *Is the "Right to Be Forgotten" a fundamental right? Is it protected under the Shield Agreement? Why should European data subjects enjoy better protection than Americans?* Published in *TheSciTechLawyer*, Vol 16, No 3, Spring 2020. A Publication of The American Bar Association | Science & Technology Law Section.
- The Institute of Electrical and Electronics Engineers (IEEE). *Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition*, 2019, standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf.
- Toffler, A. *The Third Wave*. Pan Books. In association with Allyn, 1980.
- UNESCO. *Cultural policy: a preliminary study*. Conference: Round-table Meeting on Cultural Policies, Monte Carlo, Monaco, 1967 [2]. Document code:SHC.69/XIX.1a/A <https://unesdoc.unesco.org/ark:/48223/pf0000367821>. Archives reference: CUA/57/1
- UNESCO. *Decisions adopted by the Executive Board at its 206th session*. Conference: UNESCO. Executive Board, 206th, 2019 [586]. Document code:206 EX/DECISIONS. *Preliminary study on the technical and legal aspects relating to the desirability of a standard-setting instrument on the ethics of artificial intelligence*, 2019. Conference: [586] Document code:206 EX/4 <https://unesdoc.unesco.org/ark:/48223/pf0000367821>
- Wiener, H. *Cybernetics: or Control and Communication in the animal and the machine*. New York: John Wiley, 1948.