Ethics 4EU

#Ethics4EU

European Values for Ethics in Digital Technology

Co-funded by the Erasmus+ Programme of the European Union

Executive Summary

Digital Ethics deals with the impact of digital Information and Communication Technologies (ICT) on our societies and the environment at large. It covers a wide spectrum of societal and ethical impacts including issues such as data governance, privacy and personal data, Artificial Intelligence (AI), algorithmic decision-making and pervasive technologies. Importantly, it is not only about hardware and software, but it also concerns systems, how people and organizations and society and technology interact. In addition, with Digital Ethics comes the added variable of assessing the ethical implications of artefacts which may not yet exist, or artefacts which may have impacts we cannot predict.

The Ethics4EU Project is an Erasmus+ transnational project that will explore issues around teaching Digital Ethics in Computer Science. Ethics4EU will develop new curricula, best practices and learning resources for Digital Ethics for Computer Science students. It follows a 'train the trainer' model for up-skilling Computer Science lecturers across Europe.

This research report on European Values for Ethics in Technology is the first Intellectual Output of the Ethics4EU project and it is presented in two parts:

- Part 1 used a semi-systematic literature review methodology to discuss and present the origins of Digital Ethics, recent views from EU working groups on Digital Ethics, geographical perceptions of Digital Ethics and a summary overview of pertinent Digital Ethics topics and challenges for an increasingly interconnected ICT world. These topics include data ethics, including data management and practices, AI Ethics including ethical concerns when building AI systems, automated decision making and AI policy, ethics for pervasive computing including topics such as surveillance, privacy and smart technologies, social media ethics including topics such as balancing free speech and access to accurate information and the relationship between Digital Ethics, digital regulations and digital governance with a specific focus on the GDPR legislation.

- Part 2 presents the results of focus groups conducted with three key groups of stakeholders – academics, industry specialists and citizens. The analysis captures their insights with regard to ethical concerns they have about new technologies, the skills or training future computer professionals should have to protect themselves in the online world and who should be responsible for teaching Digital Ethics. We analyse the similarities between the topics uncovered in the literature review and those highlighted by the focus group participants.

Key outputs

Based on the literature reviewed in part 1 and the analysis of sessions with stakeholders in part 2, we have developed a set of 50 values for teaching Digital Ethics. These values are dispersed throughout the report at relevant parts as well as in Appendix A at the end of the report. These values provide direction to educators for the development and delivery of teaching content for Digital Ethics by elucidating important aspects highlighted in the literature and underscored by members of our focus groups. The 50 values were classified using a thematic analysis to produce 8 categories of Digital Ethics topics. In the Ethics4EU project we will use these categories and values to guide the creation of curricula and learning materials for teaching Digital Ethics to Computer Science students. It should be noted that these 8 categories are not crisp and there is often overlap between topics from different categories. The 8 categories of Digital Ethics topics are:

- Foundations of Digital Ethics
- Digital Ethics Values
- Data Ethics
- AI Ethics
- Ethics for Pervasive Computing, Privacy and Surveillance
- Social Media Ethics
- The Relationship between Digital Ethics, Digital Regulations and Digital Governance
- Professional Ethics

Contents

Part 1

Review of Digital Ethics

1. Introduction

Ethics is in general the discipline that deals with moral issues, while morality is the whole set of opinions, decisions, and actions with which people, individually or collectively, express what they think is good or right. The systematic reflection on morality provided by ethics increases our ability to cope with moral problems. It is important to stress that ethics is not a set of predefined answers to the moral issues, but rather a process for reflecting on questions and answers concerning the moral choices people can make and for searching the right kind of morality.

Digital Ethics deals with the impact of digital Information and Communication Technologies (ICT) on our societies and the environment at large. It encompasses a range of issues and concerns from privacy and agency around personal information, digital literacy, big data including governance and accountability, the dominance of a small number of large network platforms, pervasive technology, the Internet of Things and surveillance applications, Artificial Intelligence (AI) and algorithmic decision-making including the fairness, accountability, and transparency of those automated decisions, and automating human intelligence for robotics or autonomous vehicles, AI and the future of work and governance, and intelligent technologies in healthcare and medicine. Importantly, it is not only about hardware and software, but it also concerns whole techno-social systems, how people and organisations and society and technology interact. In addition, with Digital Ethics comes the added variable of assessing the ethical implications of artefacts which may not yet exist, or artefacts which may have impacts we cannot predict. Digital Ethics seeks to understand the application of ethics to the development and (mis)use of systems which include digital and electronic components.

Many of the subjects discussed above are currently undergoing scrutiny in the media, for example the use of automated profiling using illegally harvested data in the last U.S. election and the Brexit referendum or automated decision-making software displaying gender and racial biases when shortlisting applicants for jobs, while others such as the impact of surveillance technology have been studied for decades. The last decade has seen the world of technology experience rapid, and often unchecked, growth and innovation, and technology development is happening at a much more rapid pace than the relevant ethical debates. Increasingly there is a sense that we are developing technology faster than we are assessing its moral implications.

The Ethics4EU Project is an Erasmus+ transnational project that will explore issues around teaching Digital Ethics in Computer Science. Ethics4EU will develop new curricula, best practices and learning resources for Digital Ethics for Computer Science students. It follows a 'train the trainer' model for up-skilling Computer Science lecturers across Europe.

This research report on European Values for Ethics in Technology is the first Intellectual Output of the Ethics4EU project and it is presented in two parts:

- Part 1 used a semi-systematic literature review methodology to discuss and present the origins of Digital Ethics, recent views from EU working groups on Digital Ethics, geographical perceptions of Digital Ethics and a summary overview of pertinent Digital Ethics topics and challenges for an increasingly interconnected ICT world. These topics include data ethics, including data management and practices, AI Ethics including ethical concerns when building AI systems, automated decision making and AI policy, ethics for pervasive computing including topics such as surveillance, privacy and smart technologies, social media ethics including topics such as balancing free speech and access to accurate information and the relationship between Digital Ethics, digital regulations and digital governance with a specific focus on the GDPR legislation.

- Part 2 presents the results of focus groups conducted with three key groups of stakeholders – academics, industry specialists and citizens. The analysis captures their insights with regard to ethical concerns they have about new technologies, the skills or training future computer professionals should have to protect themselves in the online world and who should be responsible for teaching Digital Ethics. We analyse the similarities between the topics uncovered in the literature review and those highlighted by the focus group participants.

Based on the literature reviewed in part 1 and the analysis of sessions with stakeholders in part 2, we have developed a set of 50 values for teaching Digital Ethics. These values are dispersed throughout the report at relevant parts as well as in Appendix A at the end of the report. These values provide direction to educators for the development and delivery of teaching content for Digital Ethics by elucidating important aspects highlighted in the literature and underscored by members of our focus groups. The 50 values were classified using a thematic analysis to produce 8 categories of Digital Ethics topics. In the Ethics4EU project we will use these categories and values to guide the creation of curricula and learning materials for teaching Digital Ethics to Computer Science students. It should be

noted that these 8 categories are not crisp and there is often overlap between topics from different categories. The 8 categories of Digital Ethics topics are:

- Foundations of Digital Ethics
- Digital Ethics Values
- Data Ethics
- AI Ethics
- Ethics for Pervasive Computing, Privacy and Surveillance
- Social Media Ethics
- The Relationship between Digital Ethics, Digital Regulations and Digital Governance
- Professional Ethics

2. Methodology

Part 1 of this report was created using a semi-systematic literature review methodology. A semi-systematic review looks at how research within a selected field has progressed over time or how a topic has developed across research traditions. The semi-systematic review approach is designed for topics that have been conceptualized differently and studied by various groups of researchers within diverse disciplines and that hinder a full systematic review process. The semi-systematic literature review is a survey of scholarly and other sources on Digital Ethics topics. The topics were chosen by consultation with other members of the research team and include the origins of Digital Ethics, recent views from EU working groups on Digital Ethics, geographical perceptions of Digital Ethics and a summary overview of pertinent Digital Ethics topics and challenges for an increasingly interconnected ICT world. Google Scholar and Scopus were the main databases accessed to find papers for the literature review. The grey literature was also consulted in order to expand the literature beyond academic insights. The review presents and provides an overview of current knowledge and existing research in Digital Ethics.

In the ethical analysis it is central to involve all the relevant stakeholders, and get their understanding and domain knowledge in the ethical analysis/ethical screening. Stakeholders are those people who affect the solutions of the problem, as well as those who are affected by the decisions following from the ethical analysis of a problem. It is very important to clarify and define different roles that various stakeholders have. For example, general public, experts, financing bodies, have different expertise and may have conflicting interests and understanding of the problem to be solved. This is typical for transdisciplinary and interdisciplinary approaches where involved parties including researchers have different perspectives. The differences in understanding are managed through discussions and mutual exchange of information and knowledge, as established in transdisciplinary problem solving methodology (Hirsh Hadorn, et al., 2008), (Swiss Academies of Arts and Sciences, 2020), (Frodeman, 2017).

3. Background – A Brief History of Digital Ethics

"Digital Ethics" can be considered a subset of "technoethics" - a term coined in 1975 by the philosopher Mario Bunge to denote the need for technologists and engineers to develop ethics as a branch of technology (Bunge, 1975) - with its focus on information and communication technologies, including computing.

Although Bunge was one of the first to use the term, a similar concept - "technethic" - was discussed a little earlier by Norman Faramelli in 1971 (Faramelli, 1971). Both these authors recognise that the origins of the term go back to early civilisations, where philosophers like Plato and Aristotle acknowledged that new technologies (in the arts and engineering) would give rise to new ethical questions. The transition from the Rennaisance to our early modern era in the 17th century gave rise to new thinking about technology; and philosophers like Francis Bacon saw technology as a fundamental part of natural and moral philosophy. It is also clear that technoethics has been influenced by the rise of pragmatism, led by William James and John Dewey in the 1870s (Fesmire, 2003), during the industrial revolution.

The fundamental question that is posed by technoethics throughout the whole of its history is how to guide new technological advances for the benefit of society in a variety of social and ethical environments. It is important to highlight that, as the philosophical and pragmatic premises of ethics vary between different frameworks (e.g. utilitarianism, deontological ethics, virtue ethics, etc.), the systematic reflection provided by ethics varies consequently, impacting technoethics and Digital Ethics. The study, and especially the teaching of Digital Ethics, cannot prescind from gaining insights on the main different frameworks in which it developed.

Rafael Capurro defined Digital Ethics in his 2009 paper as Digital Ethics deals with the impact of digital Information and Communication Technologies (ICT) on our societies and the environment at large (Capurro, 2009). Digital Ethics is closely related to other ethical topics such as computer ethics - the study of the ethical questions that arise as a consequence of the development and deployment of computers and computing technologies (Johnson, 1985) - and information ethics - the uses and abuses of information, information technology, and information systems for personal, professional, and public decision making (Hauptman, 1988). The terms Digital Ethics, computer ethics and information ethics are often used interchangeably. In this report, we focus on the term Digital Ethics as we seek to understand the impacts of technologies on society.

The writings of Bynum provide a comprehensive overview of Digital Ethics (Bynum, 1999), (Bynum, 2000), (Bynum, 2006), (Bynum, 2018) which are summarized in the following

paragraphs. Since the second half of the last century computer scientists, such as Norbert Wiener and Joseph Weizenbaum, spoke of the ethical challenges imminent in computer technology. In the beginning, the discussion was focused on the moral responsibility of computer professionals. But for scientists like Wiener and Weizenbaum the impact of computer technology was understood to be something that concerned society as a whole. For example, while Wiener was helping to develop the science of cybernetics during the Second World War, he foresaw enormous social and ethical implications of cybernetics combined with electronic computers. He predicted that, after the War, the world would undergo “a second industrial revolution” – an “automatic age” with “enormous potential for good and for evil” that would generate a staggering number of new ethical challenges and opportunities (Wiener, 1948), (Wiener, 1954). In his book “*Computer Power and Human Reason: From Judgment to Calculation*” (Weizenbaum, 1976), Weizenbaum lays out the case that while AI may be possible, we should never allow computers to make important decisions because computers will always lack human qualities such as compassion and wisdom.

As early as 1973, the U.S. Department of Health, Education & Welfare published a report "about changes in American society which may result from using computers to keep records about people" (Ware, 1973). The report - "*Records, computers and the rights of citizens*" focused on the impact of computer technology on citizen rights. It rightly identifies that technology brings dangers, as well as social advantages - "the danger that some record keeping applications of computers will appear in retrospect to have been oversimplified solutions to complex problems, and that their victims will be some of our most disadvantaged citizens." The report also discusses the risks of automation without transparency, data security, privacy and trust. It proposes a code of fair information practice for government departments to follow when implementing and deploying information systems.

The use of computers to replace (or assist) humans in problem solving activities is discussed in a paper - "*Ethics in computer-aided design: a polemic*" (Gero, 1975) - published by Gero in 1975. Although the paper focus on the problem of computer-aided design, the lessons that it provides are widely applicable to all domains where automated reasoning is applied. Gero considers whether such systems can incorporate "values" in any meaningful way. In particular, he considers who/what is to blame when decisions made by such systems lead to unwanted consequences. The paper concludes by recommending that all such systems should be developed within a well-defined ethical context (boundary).

Also in 1976, Walter Maner noticed that medical ethical questions and problems often became more complicated or significantly altered when computers got involved. Sometimes

the addition of computers, it seemed to Maner, actually generated wholly new ethics problems that would not have existed if computers had not been invented. He concluded that there should be a new branch of applied ethics similar to already existing fields like medical ethics and business ethics. After considering the name "information ethics", he decided instead to call the proposed new field "computer ethics". In 1980 Maner developed a 'starter kit' for teaching computer ethics which contained curriculum materials and pedagogical advice for university teachers (Maner, Starter Kit in Computer Ethics, 1980). It also included a rationale for offering such a course in a university, suggested course descriptions for university catalogues, a list of course objectives, teaching tips, and discussions of topics like privacy and confidentiality, computer crime, computer decisions, technological dependence and professional codes of ethics. Maner himself stated that it was difficult to interest the academic establishment – either philosophers or computer scientists in the subject. Maner was a key scholar in the field for many years and his publications and presentations expanded the field to a wider audience. For example, the following passage, from Maner's ETHICOMP95 conference keynote address, drew a number of other people into the discussion:

"I have tried to show that there are issues and problems that are unique to computer ethics. For all of these issues, there was an essential involvement of computing technology. Except for this technology, these issues would not have arisen, or would not have arisen in their highly altered form. The failure to find satisfactory non-computer analogies testifies to the uniqueness of these issues. The lack of an adequate analogy, in turn, has interesting moral consequences. Normally, when we confront unfamiliar ethical problems, we use analogies to build conceptual bridges to similar situations we have encountered in the past. Then we try to transfer moral intuitions across the bridge, from the analog case to our current situation. Lack of an effective analogy forces us to discover new moral values, formulate new moral principles, develop new policies, and find new ways to think about the issues presented to us" (Maner, 1996).

Deborah Johnson's "Computer Ethics" published in 1985 was the first major textbook in the field (Johnson, 1985). On page one, she noted that computers "pose new versions of standard moral problems and moral dilemmas, exacerbating the old problems, and forcing us to apply ordinary moral norms in uncharted realms." Contrary to Maner, Johnson believes that computers transform old ethical problems in new ways, but she doesn't agree that computers create whole new problems from an ethical point of view (Bynum, 2018) . "Computer Ethics" quickly became the primary text used in computer ethics courses offered at universities in English-speaking countries. The textbook set the computer ethics research agenda on topics such as ownership of software and intellectual property, computing and privacy, responsibilities of computer professionals, and fair distribution of technology and human power. In later editions in 1994, 2001 and 2009, Johnson added new ethical topics

such as “hacking” into people’s computers without their permission, computer technology for persons with disabilities, and ethics on the Internet.

Also in 1985, James Moor’s classic paper, “What Is Computer Ethics?” was published in a special computer-ethics issue of the journal “Metaphilosophy” (Moor, 1985). Moor provided an account of the nature of computer ethics that was broader and more ambitious than the definitions of Maner or Johnson. He went beyond descriptions and examples of computer ethics problems by offering an explanation of why computing technology raises so many ethical questions compared to other kinds of technology:

“Computers are logically malleable in that they can be shaped and moulded to do any activity that can be characterized in terms of inputs, outputs and connecting logical operations… Because logic applies everywhere, the potential applications of computer technology appear limitless. The computer is the nearest thing we have to a universal tool. Indeed, the limits of computers are largely the limits of our own creativity”.

According to Moor, computer technology makes it possible for people to do a vast number of things that it wasn’t possible to do before. Since no one could do them before, the question may never have arisen as to whether one ought to do them. In addition, because they could not be done before, perhaps no laws or standards of good practice or specific ethical rules had ever been established to govern them. Moor claimed this could lead to policy vacuums and conceptual vacuums where computer ethics was concerned. Furthermore computers provide us with new capabilities and these in turn give us new choices for action. Often, either no policies for conduct in these situations exist or existing policies seem inadequate. One difficulty is that along with a policy vacuum there is often a conceptual vacuum. Although a problem in computer ethics may seem clear initially, a little reflection reveals a conceptual muddle. What is needed in such cases is an analysis that provides a coherent conceptual framework within which to formulate a policy for action (Moor, 1985). In the same paper Moor introduces a very powerful concept, that is the invisibility factor. It refers to the fact that most of the time computer operations are invisible. Moor distinguishes three types of invisibility that have an impact from an ethical point of view. The first one is the invisibility of abuse. Invisible abuse is the intentional use of the invisible operations of a computer to conduct actions that are unethical. The second one is the invisibility of programming values. Invisible programming values are those values that - intentionally or unintentionally - are embedded in a computer program. The third one is the invisibility of calculation. Computers perform today calculations that are too complex for human inspection and understanding.

Beginning with the computer ethics works of Norbert Wiener, a common thread has run through much of the history of computer ethics; namely, concern for protecting and advancing central human values, such a life, health, security, happiness, freedom, knowledge, resources, power and opportunity. This is known as a "human-values approach" to computer ethics. In the late 1990s, a similar approach to computer ethics, called "Value-Sensitive computer design", emerged based upon the insight that potential computer-ethics problems can be avoided, while new technology is under development, by anticipating possible harm to human values and designing new technology from the very beginning in ways that prevent such harm (Flanagan, Howe, & Nissenbaum, 2008), (Brey, 2012), (Van den Hoven, 2007).

The same idea is emphasized by Donald Gotterbarn, who believed that computer ethics should be seen as a professional ethics devoted to the development and advancement of standards of good practice and codes of conduct for computing professionals. Thus, in 1991, in the article "Computer Ethics: Responsibility Regained", Gotterbarn said: "There is little attention paid to the domain of professional ethics – the values that guide the day-to-day activities of computing professionals in their role as professionals. By computing professional I mean anyone involved in the design and development of computer artifacts. … The ethical decisions made during the development of these artifacts have a direct relationship to many of the issues discussed under the broader concept of computer ethics" (Gotterbarn, 1991). Thus in the 1990s the focus changed from computer ethics to professional responsibility and advanced the professionalization and ethical maturation of computing practitioners. This resulted in the development of a number of codes of ethics and codes of conduct for computing professionals, for example, the ACM Code of Ethics (ACM, 2018).

Professional bodies continue to play a very important role in producing and disseminating ethical and standard guidelines for ICT professionals, for example the IEEE Ethically Aligned Design Guidelines provide guidance for ICT professionals involved in the design and development of autonomous and intelligent systems (The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019). It should be noted that in contrast to other professions such as Medicine or Law, which have codes of ethics and possible penalties in place for noncompliance, the ICT profession still lacks a coherent umbrella ethical framework. Within the ICT field there are multiple professional bodies based on particular ICT specialities, some of which have their own codes of ethics, but there is no coordinated approach to ethics taken by the ICT industry as a whole (Thornley, et al., 2018). There are also concerns about the extent to which IT education adequately prepares IT professionals for the ethical dimensions of their profession (Al-Saggaf, Burmeister, & Schwartz, 2017).

In 1995, a rapid growth of information and computer ethics spread to Europe when Terrell Bynum joined with Simon Rogerson of De Montfort University in England to create the Centre for Computing and Social Responsibility and to organize the first computer ethics conference in Europe, ETHICOMP95. In 1996 Krystyna Górniak-Kocikowska (Gorniak-Kocikowska, 1996), argued at ETHICOMP that computer ethics eventually will evolve into a global ethic applicable in every culture on earth. The "Górniak Hypothesis", predicted that such a theory would emerge over time because of the global nature of the Internet and the resulting ethics conversation among all the cultures of the world. Developments since then appear to be confirming Górniak's hypothesis and have resulted in the metaphysical information ethics theory of Luciano Floridi (see, for example, (Floridi, 1999), (Floridi, 2005), (Floridi, 2014)). In 1999 a seminal book was published devoted to the discussion of the structure and nature of regulation of the Internet (Lessing, 1999). The title of the book (*Code, and Other Laws of Cyberspace*) is evocative of the approach, as the author argues that computer code regulates conduct in cyberspace as much as legal code does in our societies. As such for the rest of this report we focus on Digital Ethics issues that have resulted from increasing technological connectivity, for example data management, AI, pervasive computing and social media platforms and applications.

We introduce now the first set of values that the reader will encounter throughout the report. These are intended as directions to educators for the development and delivery of teaching content for Digital Ethics. We believe that students might benefit from an introduction on the evolution of the discipline in order to gain a deeper understanding of what Digital Ethics is and in order to understand the context from which it originated. Developing content on the history of the discipline will also provide students with the relevant terminology. We condensed these directions in the values below, these can be used as guidelines for educators that deliver or intend to deliver content on the foundations of Digital Ethics.

Value 1

Develop teaching content that explains the history and evolution of Digital Ethics, the debate about the different approaches, the difference between ethics and morality and that ethics is not a manual with answers: it reflects on questions and arguments concerning the moral choices people can make. Ethics is a process for searching for the right kind of morality, also in the case of Digital Ethics.

Value 2

Create timelines for the key events in the history of Digital Ethics and introduce and discuss with students some reference to the main ethical frameworks (e.g. utilitarianism, virtue ethics, etc.) developed by ethics in general.

4. Recent Views from the EU

Against the background described above, the EU has been calling for a broad understanding of Digital Ethics as core values central to protecting human dignity, autonomy and the democratic functioning of our societies. This has led to a number of recent reports in the area described in the following paragraphs.

One of the first reports that referenced Digital Ethics was a report by the European Group on Ethics and Technology (EGE, 2012). The report focuses on the ethics of information and communication technology and on the ethical issues arising from the fast expansion of ICT. It was written at the request of the president of the European Commission, Manuel Barroso. The aim was that this piece would be used by the Commission as a reference point for promoting responsible policies across Europe for the EU's Digital Agenda and facilitating their societal acceptance. The report highlights the ethical aspects of various concepts such as digital identity, the right to privacy, and personal data safety.

In the European context (see for example the funding programmes, and H2020 in particular) the so-called Responsible Research and Innovation (RRI) (Von Schomberg, 2013) approach is highly encouraged. It is an approach that anticipates and assesses potential implications and societal expectations with regard to research and innovation, with the aim to foster the design of inclusive and sustainable research and innovation. RRI implies that societal actors (researchers, citizens, policy makers, business, third sector organisations, etc.) work together during the whole research and innovation process in order to better align both the process and its outcomes with the values, needs and expectations of society.

In 2013, the ALLEA Permanent Working Group on Science and Ethics (ALLEA, 2013) published a statement that was addressed to research organizations like universities and academies, and focuses on recommendations and practical examples about the importance of teaching ethics to scientists of all fields. The purpose of which is to help new scientists to make decisions based on a concrete understanding of the scientific, ethical and legal issues of various matters during their practice. It is also recommended that originations should provide resources that ensure that all research is undertaken by professionals who are sufficiently trained in ethics and are aware of the interchange between science and society. This is a welcome early report that notes that ethics impacts across many strands of society.

More recent reports have focused on emerging technologies, for example, the European Group on Ethics in Science and New Technologies published a statement on AI, robotics and 'autonomous systems' in 2018 (EGE, 2018). This statement examines the moral reflections and the ethical, societal and legal challenges posed by advances in the aforementioned

technology fields and the various uncoordinated initiatives, attempting to provide solutions. The authors call for an internationally uniform, ethical and legal framework and principles that deal with issues such as design, production, use and governance of AI, robotics and 'autonomous' systems. The EGE proposes a set of ethical principles built along with democratic criteria based on the EU treaties and the EU Charter of Fundamental Rights.

The High-Level Expert Group on Artificial Intelligence drafted a report in which ethics guidelines for achieving trustworthy AI are examined (EGAI, 2019). Trustworthy AI is defined by three key components, namely lawfulness, ethicality and robustness. The report provides an applied approach on operationalizing the basic principles behind the guidelines in sociotechnical systems in order to reach trustworthiness. This is described in three levels of abstraction from the necessary foundational ethical principles, onto the requirements that should be met by AI systems and finally how to assess these systems. The report concludes that the EU at its core places the most importance on the citizens, first and foremost. This leads the way towards building pioneering AI systems under Europe's fundamental values of fundamental rights, democracy and the rule of law.

A project report on the ethics of computer coding was written in 2018 by Colman et al. (Colman, Bühlmann, O'Donnell, & van der Tuin, 2018). It is argued that the integration of computers in modern societies since the 1970s has reached a level where almost everything works through a networked algorithmic environment. The authors call this the "algorithmic condition" and discuss how every form of ICT operates under this condition. Therefore, the ethical codes and guidelines that lead research in this condition should be examined and stated clearly and explicitly. Drawing from the philosophies of Hannah Arendt and Jean Lyotard they have attempted to define and address the ethical issues of ICT-related research and innovation and by doing that also redefine what the ethics for ICT could be. The report aims to encompass any infrastructure, system, network, organization or individual that is part of or actively engaged in a so-called "algorithmic condition".

The High Performance and Embedded Architecture and Compilation European Network outlined the need for Digital Ethics training for all professionals in computing and IT in 2019 (HiPEAC, 2019). This call for training was deemed necessary due to the widespread growth of ICT systems for which privacy, security and safety are critical factors, such as in buses, automobiles, and airplanes. In their view, computing professionals should consider if all potential developments in this field should be actualized or not and these decisions should be made according to ethical principles. They include examples of organizations' initiatives towards establishing ethical codes and guidelines such as the ACM Code of Ethics and the AI4People European forum among others (ACM, 2018), (AI4People, 2020). The researchers

suggest creating an international, independent scientific organization that would provide the world with a clear scientific view on ICT and AI.

The EU has led many important initiatives in Digital Ethics in the previous decade. In line with other organizations, many recent publications have tended to focus on ethical issues and challenges related to AI, machine learning, data science and algorithmic decision-making. It is fair to say that the recent interest in Digital Ethics is largely driven by these topics and their potential impact on wider society.

In a variety of ethics fields such as Ethics of technology/Ethics in technology/Technoethics, Professional ethics/Technology ethics/Engineering ethics/ Computer ethics and AI Ethics, among others, one of the central questions is about the consequences of technology for society and environment. In the emerging technologies it is often difficult to predict those consequences because of the very nature of novel and radically disruptive technologies. In those situations Precautionary principle requires that proponents of a new technology adopt precautionary measures ”when scientific evidence about an environmental or human health hazard is uncertain and the stakes are high” (European Union, 2016), (World Commission on the Ethics of Scientific Knowledge and Technology, 2006), that is for decision-making under uncertainty. The precautionary principle has three aspects (European Union, 2016): link to innovation, risk governance (risk assessment, management and communication), science-policy interfaces. Decisions applying the precautionary principle must be open, informed, and democratic and must include all important stakeholders. The World Charter for Nature, adopted by the United Nations General Assembly in 1982, was the first endorsement of the Precautionary principle. Its international implementation was through the Montreal Protocol, the Rio Declaration and Kyoto Protocol. In the European Union law, the Precautionary principle is a statutory requirement.

4.1. EU Digital Ethics: A Philosophical Perspective

Some researchers take a more philosophical perspective on these issues, for example, (Jeandesboz, 2011) looked at the ethics of the use of technology at European borders, focusing on EUROSUR - the European Border Surveillance System which represents a coupling of European ‘digital’ bordering practices with surveillance of geographical borders, with technologies such as improved sensors, satellites or unmanned vehicles. This system also allows for the proactive surveillance of populations as they travel towards the EU, and also looks for patterns in the observation of circulatory behaviours, and the profiles of individuals, thereby changing the concept of what a border is, and thereby requiring a concomitant change in our ethical conceptualization of these ideas. Furthering this notion, (Manners, 2008) argues that the EU itself is a game changing concept, and creates a new "normal" in terms of ethics, and world politics. He examines the actions of the EU with respect to their nine core principles, and concludes that they generally act in accordance

with those principles, but they must make their decisions more transparent to member states, national parliaments, the public and the media.

4.2. EU Digital Ethics: A Geographical Perspective

When drawing up the proposal for the Ethics4EU project, the partners involved thought it would have been interesting to explore the geographical perspective of Digital Ethics. As common sense suggests moral values vary somewhat geographically, it was felt teaching Digital Ethics could be impacted by this diversity and should be taken into account when developing teaching material. The researches included in this section describe very specific perspectives that might be debatable, but we felt it was worth it to include them as they point out at the general lack of sources specifically addressing Digital Ethics from a geographical perspective.

The notion of a homogeneous European perspective on ethics is an inherently flawed one, as that perspective varies from country to country, and within those, it varies in each different county, canton, region, and assembly. It can also vary within each household, but such a degree of granularity is difficult to model, therefore for this section of the review, we look at broad ethical distinctions between the north and south of Europe, and the east and west of Europe (Gottardello & Pàmies, 2019).

(Polonsky, Brito, Pinto, & Higgs, 2001) looked at a comparison of northern and southern views on business ethics. Their research looked at 962 university students across four Northern EU countries (Germany, Denmark, Scotland, The Netherlands) and four Southern EU countries (Portugal, Spain, Italy, Greece) using the Consumer Ethics Scale (CES). The paper notes that historically the Northern countries have tended to be more industrialized, and Southern countries have been more agrarian, and as such, ethics (particularly in business practices, in this case) and moral judgements vary in these two regions. They also note that Southern countries generally have slightly lower population growth rates, lower *per capita* income, fewer years of schooling, higher infant mortality rates, and lower proportions of their population living in urban areas. The research showed that there are small but significant differences specifically in views expressed on the perceived appropriateness of *Actively Benefiting from Illegal Activity*, and *Actively Benefiting from Questionable Activity*, but highlight that this might be an issue with the instrument used in this research.

(Batog, et al., 2019) compare demographics of Eastern and Western Europe, highlighting that Eastern Europe has a dwindling working-age population (partially because of emigration), although the average age is younger in the West, but population growth rates

are lower, it also has a lower *per capita* income, and lower participation in schooling and of adults in lifelong learning. (Steurer & Konrad, 2009) looked at a comparison of eastern and western views on business ethics. They looked at two qualitative approaches; an analysis of 19 corporate responsibility reports (12 from Eastern Europe and 7 from Western Europe) complemented by two surveys of 22 companies (11 from Eastern Europe and 11 from Western Europe). In their review of the literature, they note the following characteristics:

- Eastern Europeans view social responsibility as primarily being the government's responsibility, and companies generally see adhering to legislation as their primary responsibility.
- Eastern European companies have higher levels organisational corruption.
- Eastern European countries lack the systematic government incentives and initiatives for social and environmental issues.

Their research shows when comparing East to West, the East has lower corporate responsibility reporting, lower compliance with EU environmental standards, lower transparency about fraud and corruption issues, and lower interaction with civil society organizations.

4.3. EU Digital Ethics: Privacy as a Social Good

Within this discussion it is important to recognize the importance of privacy both as an individual and a social good, by recognizing how privacy is intertwined with other fundamental values, such as autonomy, equality, and democracy. First of all, privacy plays an essential role for human relationships. For instance privacy is necessary to maintain a diversity of relationships. Indeed the kind of relationships we have with others is precisely a function of the information we have about each other; if everyone had the same information about us, we would not have a diversity of relationships (Rachels, 1975). However, when privacy is framed exclusively as an individual good and individual privacy is balanced against social goods, such as security or public health, personal privacy always loses. For this reason, some scholars argue that privacy should be framed as a social good and not only as an individual one (Regan, 2015). The lack of privacy therefore has effects on democracy, and not only on the diversity of relationships. The idea of democracy is that citizens have the freedom to exercise their autonomy. Democracy, hence, requires citizens capable of critical thinking and freedom. The problem here is not just that we are being tracked and monitored. The problem is that the norms by which we are measured, evaluated, and treated are often not subject to public discussion and negotiation. They are invisible to the individuals being watched, evaluated, and treated. All this given, the importance of privacy cannot be underestimated and the public debate around the impact of ICT on this should fully include a discussion about privacy both as an individual and as a social good.

As a European project, Ethics4EU values are aligned with the EU's call for a broader understanding of Digital Ethics and awareness of ethical challenges posed by technology. Computer Science students should be trained to be aware of the interchange between their future profession and the context they live in. Educational content for Digital Ethics should be created taking into account Europe's fundamental values as it is very likely graduates will either work within the European context or market. The concept of geographical relativity should be introduced, though we believe foundational concepts such as digital identity, privacy and data safety should always be taught regardless of the geographical context in which the educational content is delivered.

Value 3

Develop teaching content that provides an overview of the ethical challenges related to machine learning, data science, algorithmic decision-making and other related AI issues.

Value 4

Develop teaching content that explains Europe's fundamental values of fundamental rights, democracy and the rule of law. Also create content that explains ethics are not static concepts, but rather change as other things change, such as technologies, conceptualizations and assemblies; and moral perspectives vary somewhat geographically.

Value 5

Develop teaching content that explains digital identity, the right to privacy, and personal data safety.

5. Survey of Recent Literature

The survey focuses on contemporary aspects of Digital Ethics, namely data and data management, AI and algorithmic decision-making, pervasive technologies, social media and governance and regulation such as GDPR.

5.1. Data and Data Management

Data ethics is a relatively new branch of ethics that studies (and evaluates) moral problems related to data management (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including those using AI and machine learning) and corresponding practices (including responsible innovation, programming, hacking and

professional codes), in order to formulate and support morally good solutions for data (Floridi & Taddeo, 2016).

The growing importance of data has led to a large increase in the value of digital data. Data has become a key input for driving growth, enabling businesses to differentiate themselves, and maintain a competitive edge. In the private sector, some have tried to place a hard monetary value on data, for instance, each Facebook user is worth $34 a year to that organisation (Meeker, 2018). The terms of the relationship between individuals and their public sector data have not yet been fully defined however decisions about healthcare, transport, credit, job opportunities, child safety, access to bail, and other sectors are made using data that people may or may not know that they have created or given away (Coldicutt, 2019). Value is particularly high from the mass collection and aggregation of data, particularly by companies with data-driven business models. A common aggregation purpose is to get more information about particular groups based on specific variables such as age, profession, or income. However, the use of aggregated data underpins the risks to individuals' privacy at a very fundamental level. Therefore, it is vital to highlight data management frameworks that promote the collection, processing and aggregation of data according to values and ethical approaches.

Big Data in particular, together with the promise of revolutionizing the scientific and technological practices, has raised several ethical issues. According to Boyd and Crawford (Boyd & Crawford, 2012) this is a socio-technical phenomenon whose conceptual complexity needs to be clarified. On the one hand, it can be seen as a powerful tool to address various societal problems (from curing cancer to addressing climate change), on the other hand Big Data raises various concerns related to surveillance, privacy, civil freedom, and increasing corporate and governmental control. Because these data are committed to record details about human behavior, they are perceived as a threat to fundamental values, such as autonomy, fairness, justice, due process, property, solidarity, and privacy. The usual tools to deal with these issues - anonymity and informed consent - do not offer final solutions and other creative approaches are under scrutiny (Barocas & Nissenbaum, 2014).

The connection between data and surveillance has been raised by many, and the new term surveillance capitalism (Zuboff, 2019) has been proposed to describe the phenomenon that uses human experience as free raw material for hidden commercial practices of extraction, prediction, and sales. The threats individuated by the process of the dispossession of data from citizens to companies and governments concern not only the individuals and their free will, but also the very conception of our democratic societies.

Data management is an important aspect of data ethics as it concerns the processes involved in the acquiring, validating, storing, protecting and processing of data. The

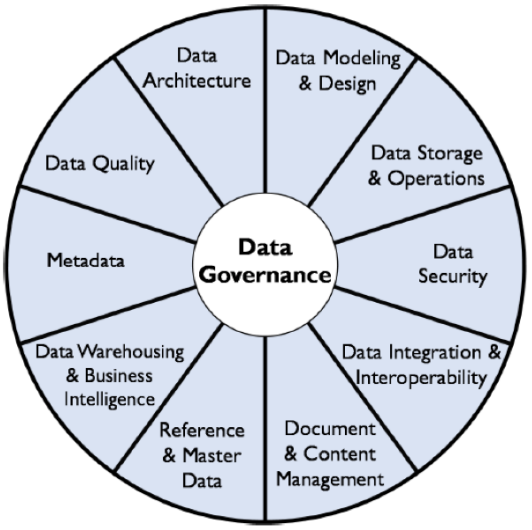
reviewed works that follow are divided into two sections. The first section reviews the various data management frameworks which are commonly used in industrial practice and their focus on data ethics. The second section reviews various research and industry published papers covering topics relating to data management and specific use cases.

There are a large number of data management frameworks currently in use in industry. After discussions with industry partners on commonly used data management tool, we review three of the popular data management frameworks in this report. They are:

- Data Management Association – Data Management Body of Knowledge (DM-BOK) (DAMA International, 2017)
- Zachman Framework (Zachman, 2009)
- Castlebridge – Ethical Enterprise Information Management Framework (E2IM) (OKeefe & O'Brien, 2018)

5.1.1. Data Management Association – Data Management Body of Knowledge (DM-BOK)

The Data Management Association (DAMA) has published their framework as the Data Management Body of Knowledge (DAMA International, 2017). In their latest edition of this framework, they have added a new chapter on data ethics and its importance for the management and governance of data within organisations. With organisations operating a data-centred approach there has been an increasing need to ensure that data is managed and handled ethically regarding the effects on all stakeholders and to ensure a focus on minimizing data related risks, DAMA have developed the DM-BOK Wheel to outline the cycle of data management in organisations.



The DM-BOK Wheel (DAMA International, 2017)

Although ethics does not appear in the DM-BOK Wheel, it is a core element of each component of it. Commencing with data governance, and typically a function operated at the highest levels of an organization, it is important that correct ethical handling of data is passed down to all parts of the organization. The ethics of data handling is complex and has the following concepts:

- *Impact on People*: data represents characteristics of individuals and is used to make decisions that affects people's lives. It is important to manage the quality and reliability of this data.
- *Potential of Misuse*: misuse of data can have a negative impact on the organization and the individuals involved.
- *Economic Value of Data*: Data has value. Ethics of data ownership should determine how that value can be accessed and by whom, to ensure it is not exploited.

The DAMA DM-BOK framework goes into detail on what is required in each of these concepts, giving guidance on what needs to be addressed, and how this related to the legal requirements for the organization. For example the EU GDPR, the United States Privacy Program and Canadian Privacy Statutory Obligations. The DAMA framework provides an outline for creating an Ethical Data Handing Strategy and Roadmap.

5.1.2. Zachman Framework

The Zachman Framework (Sowa & Zachman, 1992), (Zachman, 2008) is a logical model that provides a comprehensive representation of an information technology enterprise. It allows for multiple perspectives and the categorization of business artefacts. It is built on a six-by-six matrix. The six rows are "Scope", "Business Model", "System Model", "Technology Model", "Components", and "Working System" and the six columns are "Who", "What", "When", "Where", "Why", and "How". John Zachman defines 'architecture' as the set of design artefacts or descriptive representations that are relevant for describing an object such that it can be predicted to requirements (quality) and well as maintained over the period of its useful life (change).

	DATA *What*	FUNCTION *How*	NETWORK *Where*	PEOPLE *Who*	TIME *When*	MOTIVATION *Why*
Objective/Scope (contextual) *Role: Planner*	List of things important in the business	List of Business Processes	List of Business Locations	List of important Organizations	List of Events	List of Business Goal & Strategies
Enterprise Model (conceptual) *Role: Owner*	Conceptual Data/ Object Model	Business Process Model	Business Logistics System	Work Flow Model	Master Schedule	Business Plan
System Model (logical) *Role:Designer*	Logical Data Model	System Architecture Model	Distributed Systems Architecture	Human Interface Architecture	Processing Structure	Business Rule Model
Technology Model (physical) *Role:Builder*	Physical Data/Class Model	Technology Design Model	Technology Architecture	Presentation Architecture	Control Structure	Rule Design
Detailed Reprentation (out of context) *Role: Programmer*	Data Definition	Program	Network Architecture	Security Architecture	Timing Definition	Rule Speculation
Functioning Enterprise *Role: User*	Usable Data	Working Function	Usable Network	Functioning Organization	Implemented Schedule	Working Strategy

The Zachman Framework (Sowa & Zachman, 1992)

Although this framework was developed long before the recent emerging importance of ethics, it has ethics built into it using the "Why" column, which recognises the importance of capturing this level of detail. The original purpose of the "Why" column was to the goals, strategies and means of the organisation. Given the context of ethics, the "Why" column expands to include all new constraints as they arise. This enables positive interrogation by asking the question "Why we are doing these things?" and as a negative interrogation by asking the question "Why are we not doing these things?" This is further expanded with the different perspectives of various roles in the framework including "Executive", "Business Manager", "Architect", "Engineer", "Technician" and "User", therefore including ethical considerations at various roles within the organisation and at various levels within the data and information architecture.

5.1.3. Ethical Enterprise Information Management Framework

The Ethical Enterprise Information Management Framework (E2IM) by Castlebridge (O'Keefe & O'Brien, 2018) was developed based on their own experiences of working on data management projects while also having assisting companies with developing legal and ethical frameworks for their clients. Building upon the work in the DM-BOK and Zachman frameworks, they also bring in extended elements of a generic information management model developed by Prof. Rik Maes at the University of Amsterdam in the 1990s. The E2IM looks to address some of the issues with delivering ethical solutions. They propose that ethics needs to be considered at two levels. These are a society level and an organisational level. These then feed into the everyday processes of the organisation and the various frameworks and approaches they follow, linking the quality of the outcome to the ethical approaches followed.

The ethics in society level influences strategic and tactical governance and planning in an organisation through the definition and enforcement of laws and regulations, and the development of standards and codes of practice to support the implementation of other

legislative requirements and wider concepts of good practice. For example, customer feedback and complaints drive changes in organisational business practices, information management and technology components.

The ethics of an organisation level can influence society through lobbying at strategic level, contribution to establishing what good practices are through benchmarking and contribution to standards working groups, and through education of the customer and the wider market as to the benefits of product or the societal value of the proposed information processing.

Castlebridge posit that a framework for ethical information management practices will need to look to the future to ensure processes are designed with regard to respect for human dignity and fundamental rights such as privacy and data protection. Communication of these values and ethics in an organisation must be cross functional and extend across silos and will need to be supported by a data governance framework that ensures accountability. Following good information governance practices and ensuring ethical requirements are considered at the beginning stage of the information life cycles will help to ensure that new development in process and technologies enhances the dignity and empowerment of the person.

Castlebridge state that technology itself is neutral, but our use of technology must be ethical. The fundamental requirement in any design or plan to use technology in a novel way is to ensure that the outcomes of the new use do not result in any violations of human dignity whether by design in which the individual is seen as a means rather than an end, or by unintended consequences of a well-intended process. Rather from initial planning and design, the ethical values of upholding human dignity must be integrated and communicated as a vital consideration in the design and implementation of new technology and processes.

The neutrality of technology has been severely criticized by many. For example, Winner (Winner, 1980), one of the fathers of the current philosophy of technology, discusses the different ways in which artefacts have a political connotation. Also Deborah Johnson in her seminal book Computer Ethics (Johnson, 1985) stresses how information and communication technologies are not neutral and evidences how their design impacts on the ethical issues associated.

5.1.4. Data Management: Recent Literature

There have been a number of recent articles on data management and data governance. A broad overview of the concerns to be addressed by data-based businesses are given in (Loi, Heitz, Ferrario, Schmid, & Christen, 2019) who outline the structure and content of a code of

ethics for companies engaged in data-based business, i.e. companies whose value propositions strongly depends on using data. The code provides an ethical reference for all people in the organization who are responsible for activities around data. Ethical decisions around data usage are coupled to the question of how and to what end data is being used. These questions arise along the four steps of:

- data acquisition,
- data storage and access control,
- data processing and knowledge generation,
- usage of data-created knowledge in a concrete context.

There has also been an emphasis on effective data management design for data governance (Ladley, 2012). With much data management moving to the cloud, the role of accountability and trust for data protection in cloud ecosystem has attracted significant attention (Felici, Koulouris, & Pearson, 2013). Without accountability, cloud consumers will lack confidence to put personal and/or confidential data in the cloud and in the ability of the providers to handle their assets in a responsible way (Tountopoulos, Felici, Pannetrat, Catteddu, & Pearson, 2014). Moving data to the cloud involves a shift in responsibilities across organizational boundaries. This redistribution of responsibilities across the cloud ecosystem changes risk fundamentals (e.g. likelihood of occurrence and severity) as well as risk perceptions of such threats. This paper looks at how the customers' perceptions of the risks can affect data and IT governance.

Another important element of data management is data sharing and this is a pertinent concern for researchers and often written about in the health and biomedical domain. For example in (Rahimzadeh, Dyke, & Knoppers, 2016), the authors highlight the ethical and information governance issues raised in the development of a research project that sought to access and analyse children's social care data. It documents the process involved in identifying, accessing and using data held in Birmingham City Council's social care system for collaborative research with a partner organisation. This includes identifying the data, its structure and format; understanding the Data Protection Act 1998 and 2018 (DPA) exemptions that are relevant to ensure that legal obligations are met; data security and access management; the ethical and governance approval process.

One project that addressed a health participant's attitudes to data governance was the DIRECT project (Shah, et al., 2019). The DIRECT project collected substantial amounts of health and genetic information from patients at risk of, and with, Type II Diabetes. They then conducted a survey to understand participants' future data governance preferences. The top three priorities for data sharing on the part of participants were: highly a secure database,

DIRECT researchers to monitor data used by other researchers, and researchers cannot be allowed to identify participants. Preferences of how data should be governed, and what data could be shared and with whom varied between countries.

Another area where data sharing is of vital importance is security. In his paper on ethical dilemmas in terrorism risk management, (Eijkman, 2013) the author focuses on digital security governance in the context of the air travel and accessing sharing data from key security programmes including the Passenger Name Record (PNR), the Advance Passenger Information (API) and the Terrorist Finance Tracking System (TFTP) programmes. Particularly, it considers the ethical dilemmas of using and sharing digital personal data as well as accountability for this type of risk management. Because there are broader socio-political, legal and technological issues connected to the use of information and communication technology for digital security governance. In addition to the rule of law and good governance, public and private authorities have to be aware and take responsibility for the side effects of digital security governance on the basis of personal data. These side effects may include violating the right to seek redress if the information is incorrect or the right to privacy when religious meal preferences are used as an indicator of a threat analysis for terrorism risk management.

5.1.5. Summary

Data and data management is a fundamental pillar of data ethics. Ethics and ethical questions should be considered at each stage of the data management lifecycle - data collection and acquisition, data storage, data processing and data usage. Key issues to be deliberated include consent, the type of data being gathered (e.g. identifiable, personal or relating to vulnerable groups), the need to store data securely and who has access, the reuse of data, and what data protection rights apply. A comprehensive resources including a questionnaire for organizations to consider data ethics dilemmas has been created by the think tank Dataethics.eu (DataEthics, 2020).

The growing importance of data has led to a large increase in the value of digital data. Data has become a key input for driving growth, enabling businesses to differentiate themselves, and maintain a competitive edge. A Computer Science student entering the professional world is almost certainly likely to work with data at one stage or another of the data management life cycle. For this reason it is essential developing content in Digital Ethics that takes the importance of data into consideration. The values this content should include are synthesized below.

Value 6

Develop teaching content that prompts learners to ask relevant questions at each stage of the data management life cycle, looking at the Dataethics.eu questionnaire as a teaching tool.

Value 7

Develop teaching content that explores the power of, and the ethical challenges associated with, aggregate data.

Value 8

Develop teaching content that explores the additional ethical considerations of using cloud architectures to store data.

Value 9

Develop teaching content that highlights specific examples of sensitive data, such as health data, biomedical data, data about children, and air travel data.

5.2. Artificial intelligence

Artificial intelligence (AI) is the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. Advances in AI have brought Digital Ethics firmly into the public eye and are driving many of the discussions around ethics, data and decision making. In this section of the report, we review some recent work in AI Ethics where pertinent issues include transparency; inclusion; responsibility; impartiality; reliability; security and privacy.

(Müller, 2020) presents a wide overview of the ethical issues arising by the use of AI and robotics technologies. The main themes of this overview are the following. The ethical issues that arise with AI systems as objects, i.e., tools made and used by humans, such as privacy and manipulation, opacity and bias, human-robot interaction, employment, and the effects of autonomy. The ethical issues that arise with AI systems as subjects, i.e., ethics for the AI systems themselves in machine ethics and artificial moral agency. The problem of a possible future AI superintelligence leading to a “singularity”.

(Yampolskiy, 2013) focuses on safety engineering, and being able to prove that systems that are developed are safe, even if they are recursively self-improving. This is a significant challenge, and the author suggests that even if significant safety constraints are added to

the initial version of an intelligent system, it is extremely difficult to ensure that the successive generations of the system will retain the initial safety constraints, as well as ensuring the safety added in every generation. The author also argues that research into the development of strong AI, that is the idea that appropriately programmed computers can literally said to understand and have other cognitive states (Searle, 1980), is inherently unethical as it may result in the creation of human-level intelligence which may result in the AI suffering, as well as humans being replaced.

(Dignum, 2018) examines some of the ethical dilemmas associated with AI, including the topic of AI safety and explores the moral dilemmas that arise from ethical algorithms. The author notes that in 2016 the European Union created a draft report with recommendations to the Commission on Civil Law Rules on Robotics, discussing in particular autonomous and intelligent systems, that outlines a possible legal framework for the implications of AI unleashing the upcoming industrial revolution. Dignum also goes on to point out that a fixed set of ethical rules can lead to contradictions, and predicaments; and using the example of self-driving cars to illuminate some of these challenges, and argue that AI systems should have in-built ethical systems that align with the value of their owners rather than a universal set of ethical values. For example, universal values may unwittingly discriminate against certain groups and not have adequate focus on inclusion. Dignum further discusses the need to have a flexible ethical system in the context of automated military weapons, and highlight the importance of the need for a human who can decide when to (and when not to) deploy a weapon. Dignum concludes by highlighting the importance of AI safety, i.e. creating systems with more emphasis on developing safe “agents” and safe approaches to AI.

5.2.1. The Cases For and Against AI

The case against AI, as argued by (Helbing, et al., 2019), suggests that as a result of AI in combination with Big Data we have reached a dangerous crossroads, where there is a potential that AI might pose a serious threat to humanity, views that are echoed by tech visionaries like Elon Musk, Bill Gates, and Steve Wozniak. We are already seeing how individual privacy has been compromised and as a result controlled individuals are in countries like Singapore, where computer programs are influencing economic and immigration policy, the property market and school curricula, under the aim of protecting citizens from terrorism. Software systems are already using behavioural economics and “persuasive computing” to program people to behave in particular ways and this trend will continue unabated unless legislative action is taken (Zuboff, 2019). In fact when combining this type of “nudging” with Big Data we get a new phenomenon called “Big Nudging” or “Hyper Nudging” where the persuasion is being targeted at each specific individual instead of being aimed at a societal level. The author argues that to combat these trends more empowerment of each citizen is needed (through the development of new tools like digital assistants), more transparency is needed (at a societally level, at a government level, and at

a technology level), and decentralisation is vital – of services, of data, and of computer systems. This is also related to the moral general debate on the moralization of technologies (Verbeek, 2011). The moralization of technologies is the deliberate development of technologies in order to shape moral action and decision-making. Instead of moralizing only other people, humans should and could also moralize their material environment, including the technologies designed and adopted. One of the biggest challenges of this debate today is whether it is possible to moralize technologies in a democratic way.

The case for AI, as argued by (Gurkaynak, Yilmaz, & Haksever, 2016), presents an exceptionally optimistic view of AI, highlighting those researchers that have predicted positive outcomes for the use of AI, and they claim that "*fear mongering continues to hinder AI development*". They worry that regulations will stifle the success of AI, and suggest that future AI researchers will see efforts to regulate AI as the 'dark age' of human advancement. They discuss the problem with the lack of consensus in terms of a robust definition of AI, and put forward the idea that it is impossible to regulate something that cannot be defined, they also mention other challenges that regulators will face; liability gaps, control and transparency problems. They therefore reiterate that it is "*very early to begin thinking about regulating AIs or AI studies, particularly if such regulations may hinder developments that could prove essential for human existence*". They also point out that in terms of liability AI systems are made up of a number of computer programs, some of which might have been written many years before an AI element was introduced, therefore it would be unfair to hold the developers of those programs liable for results caused by the AI system. They do underscore that Irish and UK law have explicit legislation stating that computer-generated works are owned by the person who created the programs. They conclude that until there is a better understanding of the potential of AI, it is best to nurture it.

5.2.2. Building AI Systems

(Dignum, 2018) looks at an approach to the design of AI systems that incorporate human values and ethics. The author proposes the use of the Accountability, Responsibility and Transparency (ART) design principles for an enhanced development process of AI systems. She highlights that there is almost no agreement as to what the future of AI will bring, with predictions ranging from utopia to dystopia. The research uses an adapted version of the Delphi Method (iteratively surveying experts) to identify areas of agreement, and found that respondents reached consensus on the following points:

1. Future AI will decrease the number of mechanical jobs in the short term
2. It will create new, most likely very specialized jobs
3. It will therefore have a large impact on the nature of European jobs
4. Governments will need to revise their education system to make sure their future workforce can work with AI

The author also recommends that three pillars are needed for Responsible AI: (i) All of society must take responsibility for impact of AI (this should be addressed in the education process, and specifically in the Computer Science students), (ii) Models and algorithms of ethics and values need to be developed, and (iii) it is necessary to understand how a range of cultures work with and live with AI.

(Pistono & Yampolskiy, 2016) take a contrary approach and suggest that instead of trying to design a safe AI system with good in-built ethics, an interesting alternative is to design a malevolent AI to understand the potential implications of such a development. The authors develop some general guidelines for developing such an AI, including two key steps, don't create a global AI regulatory body, and make sure all AI code is closed source. Furthermore, they speculate if a malevolent AI were to operate on the cloud, it would be difficult to detect, and it could move from system-to-system, as well as potentially have the ability to predict when it is in danger and to move to another server. Such a system can use falsified communications to overthrow governments and launch missiles; alternatively, it could slowly, and incrementally, influence legislative policy decisions and judicial decisions. The author concludes that if such a malevolent AI is possible, then it is the obligation of researchers to publish any instances of AI projects that negative outcomes occurred, and to share code to help the community understand why these things happen, and how to prevent them.

This discussion can be inserted in the debate about moral artificial agents, the idea that machines can, in some sense, be ethical agents responsible for their actions, or autonomous moral agents (Van Wynsberghe & Robbins, 2019). This idea is related to the approach called machine ethics, that is the ethics for machines, where machines are meant as subjects (Anderson & Anderson, 2011). The basic idea of machine ethics is now finding its way into current robotics, where the assumption that these machines are artificial moral agents in any substantial sense is usually not made (Winfield, Katina, Pitt, & Evers, 2019).

5.2.3. AI Legal Liability

(Čerka, Grigienė, & Sirbikytė, 2015) explore the liability for damages caused by AI, and highlight the point that even without specific AI laws Article 12 of *United Nations Convention on the Use of Electronic Communications in International Contracts* applies. This states that a person (whether a natural person or a legal entity) on whose behalf a computer was programmed should ultimately be responsible for any message generated by the machine. In this regard, AI programs can be seen in the same way as a hammer or spanner, and with this view of AI-as-Tool with no independent volition of its own in some cases vicarious and strict liability is applicable for AI actions. From this point of view, the AI could be treated in the same way as a slave in Ancient Rome and the *Respondeat Superior* liability theory applies; or

as a child, and the *Vicarious Liability Doctrine* applies. Alternatively, if AI is seen as fully autonomous, AI-as-Person, then AI systems must be aware of their actions, and liable for their actions. The authors conclude that at the moment AI is not recognised as a legal person, and therefore the AI-as-Tool theory is applied, and therefore strict liability rules govern the behaviour of AI, and that liability applies to the developers, users and owners of the AI systems.

5.2.4. AI Policy

(Catch, Wachter, Mittlestadt, Taddeo, & Floridi, 2018) compare three government reports on how to prepare for the future of AI, these reports were from the American government, the European Parliament, and the UK House of Commons, all released in 2016. They looked at three key criteria:

1. The development of a 'good AI society'
2. The roles of the government, the private sector and researchers
3. The shortcomings in these reports

The authors begin by pointing out that AI has had many false dawns since its inception in the 1950s, but with the advent of four key factors (better statistical models, very large datasets, cheap computational power, and pervasiveness of technology in own lives), AI is making significant progress. They highlight the fact that at the moment AI research is being driven by the private sector, and thus there is a deficit of social and political accountability and long-term planning, which must be addressed. All of the reports shared the view that three key elements of good AI are transparency, accountability, and a 'positive impact' on the economy and society.

(Winfield & Jirotka, 2018) describe a roadmap to AI (and robotics) ethics governance, that combines a number of elements, including standards, regulation, Responsible Research and Innovation (RRI), and public engagement. The authors argue that this roadmap is vital to ensure there is public trust in AI, otherwise the economic and societal benefits of AI will not be fully realized. They view ethics as a subset of a broader framework of responsible research and innovation, which incorporates a verification-and-validation stage to link the standards to the real-world. They suggest that a regulatory body is needed to ensure transparency and to build public trust (and provide a list of principles concerning ethics in AI and robotics from 1950 to 2017). They also recommend looking at this from the point-of-view of AI safety, and looking for inspiration to safety-critical systems development. They recommend that AI systems in the future should not be based on Artificial Neural Networks (ANNs), as they do not have the explainability of other AI techniques, and therefore cannot be fully checked and validated. They conclude with the idea that all

organizations working with AI should:

1. Publish an ethical code of conduct, so that everyone in the organization understands what is expected of them
2. Provide ethics and RRI training for everyone, without exception
3. Practice responsible innovation, including the engagement of wider stakeholders within a framework of anticipatory governance
4. Be transparent about ethical governance
5. Really value ethical governance

(Erdélyi & Goldsmith, 2018) also propose the development of an international AI regulatory body, highlighting that there are a number of both academic, and joint public and private sector venues, that support governments in promoting AI Research and Development, but an independent body is needed to help improve policymaker's expertise on matters related to AI. The authors point out that AI can have international impacts, therefore an international perspective is needed, especially to avoid the conflicts surrounding differing domestic approaches to legislation. They propose the International Artificial Intelligence Organisation (IAIO) that would unite the views of the public sector, industry, and academia. They point out that as AI is a rapidly changing field, themes to consider include whether or not countries should have a binding commitment to the IAIO or a flexible cooperation arrangement; for certain applications such as weaponized AI, countries may need flexibility in terms delegation and autonomy; information sharing between countries can either be collective or closed; moving from more simple administrative functions to more complex centralized administration; and finally whether the focus should be on management of routine matters or focus on crisis issues.

(Cath, 2018) looks at the governance of AI, specifically focusing on the ethical, legal-regulatory and technical challenges. Given the potential harm that AI could do in a diverse range of areas including granting parole, diagnosing patients and managing financial transactions, the author suggests that a multidisciplinary approach is the key to success. She highlights three areas that should be focused on:

1. Ethical governance: Looking at the most important AI Ethics issues, such as fairness, transparency and privacy
2. Explainability and interpretability: these two concepts could be seen as possible mechanisms to increase algorithmic fairness, transparency and accountability
3. Ethical auditing: for highly complex algorithmic systems, accountability mechanisms cannot solely rely on interpretability. Auditing mechanisms are proposed as possible solutions

The author outlines the global debate that is occurring in the field of AI Ethics, whether something other than regulatory and ethical approaches are needed, and if not, she questions if the existing regulatory and ethical frameworks are sufficient. She also notes that academics are being criticized for taking complex social concepts like “fairness” and “discrimination” and attempting to reduce them to “simple statistics”, and because of this, they are misleading policy-makers on the ease of measuring these outcomes. The paper concludes with the notion that it is vitally important that everyone has a voice in how these systems are deployed, and how their data is used to build these systems.

In (Floridi, et al., 2018) an ethical framework for a good AI society is introduced. This article introduces the opportunities and risks of AI for society and presents five ethical principles: beneficence, non-maleficence, autonomy, justice, and explainability. They are derived from the four traditional principles of bioethics and from a survey of existing sets of principles produced by various reputable, multi-stakeholder organisations and initiatives. Finally it offers concrete recommendations to assess, to develop, to incentivise, and to support good AI.

5.2.5. Ethics Washing

(Wagner, 2018) argues that corporations are using the term “ethics'' as a smokescreen to avoid further regulation, and it simply allows these organizations to continue existing self-regulatory initiatives. The paper cites an example of an employee from a large multinational AI organization publicly claiming their actions were ethically justifiable although they had broken the law, leading the author to reflect whether it is possible to take an action that is both ethical and illegal. The difference between ethical and legal issues is very relevant in the ethical debate. Ethics and the law do not fully coincide, and it is perfectly plausible to have actions that are legally permissible, but that are considered immoral by somebody. Think for example to the case of abortion. Accordingly, the difference between moral responsibility and legal liability is very important to stress, even if the two concepts sometimes overlap (Van de Poel & Royakkers, 2011). Even at an international level, the private sector is increasingly portraying the world’s governments and their regulatory instruments as the reason why ethical regulation cannot be successfully implemented. The paper goes on to critique the EU position on AI Ethics who published draft guidelines based on their report on the Ethics of Artificial Intelligence (EGE, 2018). The principles developed do include important rights, such as human dignity, but also introduce completely unrelated principles such as sustainability while also entirely leaving out other aspects such as freedom of assembly or cultural rights. The report also suggests that self-regulation for organizations is the default manner of ensuring ethical adherence, and only if that fails should new regulations be explored. The author points out that approaches to software development like Value based Design, Privacy by Design, Legal Protection by Design, and Ethical Design (Cavoukian, 2009), (Balkan, 2017) align to the need to combine the legal and ethical aspects

of AI system development, and he concludes that a crucial aspect that must be brought to the forefront is transparency on all levels.

5.2.6. Value-Sensitive Design

In the last years the idea of active responsibility in the development of technologies in general, and of AI technologies in particular, has emerged. This means not only preventing the negative effects of a technology, but also realizing some positive effects (Bovens, 1988). One way of implementing active responsibility is Value-Sensitive design, where moral considerations and values are used as requirements for the design of technologies (Friedman, Kahn, Borning, Zhang, & Galletta, 2006), (Van den Hoven, 2007), (Friedman & Hendry, 2019). This approach aims at integrating three kinds of investigations: empirical investigations that take into account contexts and experiences of people affected by technological design; conceptual investigations that consider the values at stake and their possible trade-offs; technical investigations that analyze the relationship between design and values.

When Value-Sensitive design is applied to AI technologies, the issues related to the choice of incorporating moral values into these sophisticated technologies become more serious. The idea of incorporating positive values, such as for example in the case of so-called Beneficial AI, is not without risk. In particular, there is a variety of negative reactions to the AI technologies that are created to steer human behaviour (also when they are for the good) (Verbeek, 2011). A possible fear is that human freedom is threatened and that democracy is exchanged for technocracy. The idea that not humans but technologies are in control is tightly connected to the perception of a reduction of autonomy as a threat to human dignity. There is also the risk that, when moral decisions are delegated to machines, humans can become lazy in moral decision-making or even incapable of it. What has been stressed is that technologies differ from laws in limiting human freedom because they are not the result of a democratic process. So, as already mentioned, one of the open challenge is to find a democratic way to moralize technologies.

5.2.7. Medical applications of AI

Artificial intelligence (natural language processing, machine learning, robotics, intelligent prosthetics) has been increasingly used in almost any field of medicine (Buch, Ahmed, & Maruthappu, 2018), (Righby, 2019). It is used in medical education and healthcare and diagnostics, clinical decision making, personalized medicine, biomedical research and drug development, tele-health, radiology, electronic health records, imaging/image analysis, cognitive enhancements, and more.

Medical applications of AI have very sensitive ethical aspects (Righby, 2019) - from big data with privacy, personal integrity, and fairness issues typical of data-driven medicine, to the "AI for good" with the ambition not only to solve problems but also actively improve human condition. One illustrative example of promises and challenges are prosthetic devices that are being used in patients with dementia to assist memory encoding and retrieval, which at the same time can change memories or other cognitive functions. Among the central questions is the responsibility when intelligent programs make decisions, give diagnosis or choose treatment. Moreover, among ethical aspects are requirements for transparency and explainability of intelligent algorithms used in decision making in medicine and healthcare.

5.2.8. Summary

AI has emerged as one of the central issues in Digital Ethics. While many researchers and technology experts are excited by AI's potential, many others are unsettled by it. Authors balance the positive effects of AI (self-driving cars leading to better safety, digital assistants, robots for heavy physical work; and powerful algorithms to gain helpful and important insights from large amounts data) against the negatives (automation leading to job losses, rising inequalities attributed to AI haves and have nots, bias, and threats to privacy).

When developing educational content on Digital Ethics with a focus on AI, educators should especially highlight the potential arms of AI based technologies starting from the most common and most controversial. Content should be contextualized within legal frameworks since it is crucial in order to understand a few of the most commonly concepts associated with AI such as *"behavioural economics", "persuasive computing", "Big Nudging" and "Hyper Nudging"* and the concept of malevolent AI.

Value 10

Develop teaching content that highlights both the potential benefits and potential harms of AI. To do this look at specific topics such as AI safety, intended and unintended consequences of systems, fairness, accountability and transparency of AI systems, AI bias, responsible AI and regulatory issues.

Value 11

Develop teaching content that discusses ethical issues that arise from the use of AI, for example, self-driving vehicles, and automated military weapons.

Value 12

Develop teaching content that discusses existing legal frameworks that apply to AI (including Article 12 of United Nations Convention on the Use of Electronic Communications in International Contracts), and the links of these legal frameworks with Digital Ethics.

Value 13

Develop teaching content that will explain the dangers of concepts such as "behavioural economics", "persuasive computing", "Big Nudging" and "Hyper Nudging".

Value 14

Develop teaching content that will explore the notion of designing a malevolent AI to understand the implications of such a creation.

5.3. Pervasive Computing

The terms "pervasive computing," "ubiquitous computing", "ambient intelligence," and "the Internet of Things" refer to technological visions that share one basic idea: to make computing resources available anytime and anywhere, freeing the user from the constraint of interacting with ICT devices explicitly via keyboards and screens. This is possible by seamlessly embedding computational devices in everyday objects and equipping them with sensors that enable them to collect data without the user's active intervention or even awareness. This vision has partly become a reality during the last two decades through the continued miniaturization of ICT devices, the use of positioning systems making devices aware of their location, and the growth of networks for wireless or mobile communication.

Ethics is a central topic in pervasive, mobile and ubiquitous computing and the question of adopting an ethical approach to pervasive computing extends beyond the research domain, affecting practitioners as well (Davies, 2013) (Kranzberg, 2019). One of the central tenets of pervasive computing is "Understanding and Changing Behavior," a topic that clearly has significant ethical considerations (Berdichevsky & Neunschwander, 1999). It has far-reaching implications, for example surveillance technologies, effects on privacy, virtual and augmented reality applications and technological paternalism (Hilty, 2015), (Macnish, 2017), (Zuboff, 2019).

5.3.1. Surveillance

The ethics of surveillance considers the moral aspects of how surveillance is employed (Macnish, 2017). Common questions posed include: Is it a value-neutral activity which may

be used for good or ill? or is it always problematic? and if so why? What are the benefits and harms of surveillance? Who is entitled to carry out surveillance? When and under what circumstances? One of the core arguments against surveillance is that it poses a threat to privacy, which is of value to the individual and to society (Zuboff, 2019). This raises a number of questions about privacy, what it is and to what extent and why it is valuable. The discussion about automatic identification started with RFID (Oertel, Wölk, Hilty, & Köhler, 2005), which is, however, less powerful than newer technologies of face recognition or device fingerprinting. In a world of ubiquitous automatic identification, the amount of personal data generated and circulated is expected to increase dramatically. Facial recognition technology is a major focus of study in Digital Ethics. It has been applied in diverse areas such as catching criminals, finding missing people, biometric scanning in medicine, advertising and shopping. However there are concerns about the technology too. For example, there are concerns about consent issues in storing and collecting the biometric data. In addition the accuracy of systems and in particular the poorer accuracy among some racial groups.

5.3.2. Privacy

Privacy, according to (Rasmussen, Beardon, & Munari, 2001) is an integral part of pervasive computing and is defined as an individual condition of life characterized by exclusion from publicness. In the context of computing, privacy is usually interpreted as "informational privacy," which is a state characterized "by controlling whether and how personal data can be gathered, stored, processed or selectively disseminated". Also constitutional or decisional privacy, that is the freedom to make one's own decisions without interference by others in regard to matters seen as intimate and personal, is needed to be taken into account when the pervasiveness of computing affects human behaviours and the same idea of free will (Van den Hoven, Blaauw, Pieter, & Wartnier, 2020). There is a clear conflict between privacy and pervasive computing technologies, particularly those technologies that deal with sensing and storage (Jacobs & Abowd, 2003). As an ethical issue in computing, information privacy is usually discussed as being threatened by computing infrastructures that facilitate the dissemination and use of personal data. The resulting requirement to protect individual privacy against data misuse entered many laws and international agreements under different terms, some of them focusing on the defensive aspect, such as "data protection," others emphasizing individual autonomy, such as "informational self-determination". Threats against informational self-determination were mainly perceived as originating from governments. Later data protection in the Internet age was discussed in connection with data security and encryption, and the focus increasingly turned to the private sector. For example, the use of cookies, the creation (and sale) of profiles about individuals' financial behaviour, and the private sector's interest in geographic data were discussed in the context of data protection in 2001 (Macnish, 2017). It is interesting to note the work of Miltgen (Miltgen & Peyrat-Guilla, 2014) that found significant differences among European citizens in their attitudes to privacy and in particular to how they disclose and protect their personal

data. The authors found that a geographical north–south divide appears for the importance of responsibility for personal data management. Moreover, people regard disclosure differently in the south (as a choice) and east (as forced) of Europe. Younger people express more positive attitudes toward data management, feel more responsible, and are more confident in their ability to prevent possible data misuse.

5.3.3. Virtual Reality

Communicating through virtual reality can be challenging because many natural aspects of communication may become unclear, for example, with whom we are communicating, who is following the communication (Berleur, Hercheui, & Hilty, 2010). Virtual or augmented reality techniques are likely to be used in a context connected to physical reality, such as remote medical diagnosis or surgery. There is a risk that communicative acts in such environments are more ambiguous than in a natural environment, which can cause damage, or that decisions are delegated to the technology in a way that affects the autonomy of the humans involved (both doctor and patient). On the other hand, augmented reality is expected to improve the precision of interventions and the availability of information during operations (Haluza & Jungwirth, 2018). Similar arguments may apply in other safety-critical domains.

5.3.4. Technology Paternalism

In terms of pervasive computing and technology paternalism, (Spiekermann & Pallas, 2005) question how people can maintain control in environments that are supposed to be totally automated. In the general discourse about the ethics of computing, paternalism is discussed mainly in two domains: security and e-health, e.g. active implants and other remote methods of personal health monitoring (McCullagh, Beattie, & Nugent, 2010), (Wickramasinghe, Troshani, & Goldberg, 2012). Technology paternalism, however, is considered an inherent tendency in pervasive systems, in particular when machine-learning techniques are applied to infer the user's intentions. Another important aspect of technological paternalism discussed in the pervasive context is the use of tracking and tracing devices in dependency relationships. On the one hand, tracking can enhance the safety and security of the tracked persons, in particular patients, children, or employees. On the other hand, tracking represents a serious threat to the self-determination of the tracked individual. Who should be given the right to track and trace whom for what purpose?

5.3.5. Smart Cities

Smart Cities and the ethical issues that arise from them are a topic of many papers in the literature. For example, (Callaghan, Clarke, & Chin, 2009) are concerned with technology and privacy related to intelligent buildings and environments. While promising huge benefits, this technology raises new and significant dangers for users and their privacy. The

authors call for regulation of intelligent buildings and environments. In (Sholla S. , 2018) the authors discuss the ethical considerations of a smart city. A Docile Smart City Architecture to address the question of ethics of a smart city is introduced. This architecture has five layers - physical layer, network layer, data analytics layer, transparency layer and business layer incorporates social and ethical dimension and business layer provides end user services. By employing a transparency layer, sociological requirements of a smart city are addressed, social acceptance facilitated and the economic value of smart city increased. The transparency layer is concerned with ethics, culture and law.

Other authors have noted the dream of efficiencies smart homes could bring and then at the "privacy nightmare" that this could introduce (Albrechtslund, 2007). Adaptive environments have the potential to empower people at home and at work is in the near future. However, inhabitants of the home and the employees at work rather than solely the architectural and technological possibilities need to be priority.

5.3.6. Summary

Pervasive computing aims to create services that respond directly to their user and environment, with greatly reduced explicit human guidance. Ethical guidance is required when considering the design and implementation of complex, integrated, multiple systems embedded within the social infrastructure in a way that their use is often invisible. For example, (Godara, 2008) states that pervasive technologies should be under the following headings – privacy, equal access, cultural, religious and linguistic diversity, informed consent and the normative ethics that guide the ICT professional.

When creating educational content on pervasive technologies, Computer Science students should be made aware on how far-reaching the implications of pervasive computing are. Below we tried to synthetize the main perspectives the subject should be addressed from.

Value 15

Develop teaching content that will explore pervasive technologies from the points-of-view of privacy, equal access, diversity, informed consent and normative ethics.

Value 16

Develop teaching content that will explore pervasive technologies and the Internet of Things from the perspective of them as being invisible technologies, which need to be transparent and explain to users how they are collecting, aggregating and interpreting their data.

Value 17

Develop teaching content that will explore pervasive technologies from the perspective of policies and regulation, and the rights of individuals to retain control over their data, and understand the control mechanisms that govern these processes.

5.4. Social Media

Social media is communication through websites and other online platforms (e.g., Facebook, Twitter, Instagram, and LinkedIn) that are used by large groups of people to share information, develop social and professional contacts, and promote business. In recent years these new social media technologies began to transform the social, political and informational practices of individuals and institutions across the globe, inviting a philosophical response from the community of applied ethicists and philosophers of technology (Vallor, 2016).

Given the high rate of social media use by the public, organizations are compelled to engage with key audiences through these outlets. Social media engagement requires organizations to actively participate with public groups, and this highly-interactive exchange raises a new set of ethical concerns for communications (Bonsón, Royo, & Ratkai, 2015), (Williams, Burnap, & Sloan, 2017). Facebook, the largest social media platform was the first social media channel to be put under the spotlight. Its platform features an increasingly effective model for targeting specific groups, globally, with the right advertising message.

5.4.1. Free Speech on Social Media Platforms

Free speech is a human right, and social media is its facilitator. Social media facilitates free speech, unfettered but for the policing of abuse, and censoring of posts which violate societal norms. Commentators are divided over whether the value of a free speech arena is compromised by the ease with which bots and trolls are able to manipulate the system (Leerssen, 2015). Despite criticism over fake news and the current advertising and influence scandals, the digital giants are wary of actions that may open them up to accusations of bias. Traditional media outlets are already clear on their responsibilities. Reuters aims for “independence, integrity and freedom from bias.” So far, the influence of social media has been overlooked by regulators and providers. A central question is to what extent existing media ethics is suitable for todays and tomorrow’s media that is immediate, interactive and “always on” – a journalism of amateurs and professionals. Most of the principles were developed over the past century, originating in the construction of professional, objective ethics for mass commercial newspapers in the late 19th century.

5.4.2. Privacy on Social Media Platforms

Another key ethical concern for social media is privacy. Some fundamental practices of concern include: the potential availability of users' data to third parties for the purposes of commercial marketing, data mining, research, surveillance or law enforcement; the capacity of facial-recognition software to automatically identify persons in uploaded photos; the ability of third-party applications to collect and publish user data without their permission or awareness; the frequent use by social media companies of automatic 'opt-in' privacy controls and the lack of informed consent; the use of 'cookies' to track online user activities after they have left a social network; the potential use of location-based social networking for stalking or other illicit monitoring of users' physical movements; the sharing of user information or patterns of activity with government entities; the unaware involvement of users into experimentation; and, last but not least, the potential of social media companies to encourage users to adopt voluntary but imprudent, ill-informed or unethical information sharing practices, either with respect to sharing their own personal data or sharing data related to other persons and entities (Flick, 2015), (Vallor, 2016).

These new actors in the information environment create particular problems with respect to privacy norms, for example, since it is the ability to access information freely shared by others that makes social media uniquely attractive and useful, and given that users often minimize or fail to fully understand the implications of sharing information on social media, it can be found that contrary to traditional views of information privacy, giving users greater control over their information-sharing practices may actually lead to decreased privacy for themselves or others (Hull, Lipford, & Latulipe, 2011), (Bakardjieva & Gaden, 2012), (Hull, 2015).

5.4.3. Cyberharassment on Social Media Platforms

Another emerging ethical concern on social media platforms is the increasingly political character of cyberharassment. For victims of cyberthreats, traditional law enforcement bodies offer scant protection, as these agencies are often ill-equipped or unmotivated to police the blurry boundary between virtual and physical harms (Björn Ross, Ross, Rist, & Carbonell, 2017). Cyberharassment has been shown to be disproportionately aimed at specific groups, for example, the United Nations Broadband Commission Working Group on Gender suggests that 73% of women worldwide have been exposed to or have experienced some form of online violence (UN Broadband Commission for Digital Development, 2015). In the EU-28, 18 percent of women have experienced a form of serious Internet violence at ages as young as 15. This corresponds to about 9 million women (UN Broadband Commission for Digital Development, 2015). The WWW Foundation has found that law enforcement agencies and the courts are failing to take appropriate actions for cyber harassment against women in 74% of 86 countries surveyed (World Wide Web Foundation, 2015). The sheer volume of cyberharassment experienced by women has severe social and

economic implications for women's status on the Internet. These include time, emotional bandwidth, financial resources including legal fees, online protection services, and missed wages. Research carried out in 2018 by the International Women's Media Foundation and Troll Busters found that nearly two thirds of female journalists surveyed said they have experienced online harassment. 40% of respondents said they avoided reporting on certain stories as a result of experiencing such abuse (Ferrier, 2018). As such cyberharassment has a profound impact on free speech and advocacy. This is a problem that needs to be addressed if social media is to remain an open and empowering space for women and girls, and by extension, for boys and men.

5.4.4. Summary

The phenomenal rise and large scale growth of social media platforms happened on the back of multiple emerging technologies, not individually but together. Cloud-enabled big data, mobile technology and increasingly machine learning helped deliver influence through social media, all made possible by the Internet and the World Wide Web. Behaviour modification facilitated by social media whether through targeted advertising, influencing democratic processes or cyberharassment are coming increasingly under the spotlight, however legislation to regulate social media platforms has been slow to emerge.

To cope with this legislative vacuum, it becomes more important to raise awareness in students on how to consider the ethical consequences in the development of ICT technologies through the use of consequence scanning, a methodology that allows software designers and developers to consider the potential consequences - intended and unintended – of new technologies (Brown, 2019). This should be taken into consideration when developing educational content on Digital Ethics.

Value 18

Develop teaching content that will use social media case studies to explore a range of key digital ethical themes, such as data, informed consent, targeted advertising, AI and privacy.

Value 19

Develop teaching content that will explore the relationship between social media and the behaviour of a society.

Value 20

Develop tools and materials to highlight the importance of consequence scanning.

Value 21

Develop teaching content that will discuss how trolls and bots are being used by malevolent actors to manipulate social media systems. Also look at the impact that cyberharassment has on free speech and advocacy, particularly considering issues around gender and online abuse.

5.5. Governance and Legislation, including GDPR

Digital governance is the practice of establishing and implementing policies, procedures and standards for the proper development, use and management of the infosphere. For example, through digital governance, a government agency or a company may (i) determine and control processes and methods used by data stewards and data custodians in order to improve the data quality, reliability, access, security and availability of its services; and (ii) devise effective procedures for decision-making and for the identification of accountabilities with respect to data-related processes. Digital regulation concerns the relevant legislation to regulate the behaviour of the relevant agents in the infosphere.

A comprehensive overview of digital governance is given in (Floridi, 2018) which is summarized in the paragraphs below. Digital governance may comprise guidelines and recommendations that overlap with digital regulation, but are not identical to it. Not every aspect of digital regulation is a matter of digital governance and not every aspect of digital governance is a matter of digital regulation. In this case, a good example is provided by the General Data Protection Regulation (GDPR, more on the GDPR presently). Compliance is the crucial relation through which digital regulation shapes digital governance.

5.5.1. GDPR Overview

GDPR (General Data Protection Regulation) is a collection of regulations intended to protect the data of citizens within the European Union. The regulations were brought forward by the Council of the European Union, European Parliament and European Commission with the goal of providing EU citizens with a greater level of control over their personal data. It came into effect on May 25th 2018.

GDPR legislation replaces the Data Protection Directive 95/46/EC which was first created during the 1990s and had struggled to keep pace with rapid technological changes. The Data Protection Directive required that personal data should be fairly and lawfully collected for a valid purpose; accurate, relevant, and up-to date; not excessive in relation to its purpose; not retained for longer than needed for the purpose; collected with the knowledge and consent of the individual or otherwise on a legal basis; not communicated to third

parties except under specified conditions that might include consent; kept under secure conditions; and accessible to the individual for amendment or challenge.

GDPR is designed to harmonize data privacy laws across Europe, to protect and empower all EU citizens' data privacy, independently of geographical location, and to improve the way organizations across the EU approach data privacy. The ethical principle behind GDPR is that the processing of personal data ought to be lawful, fair and transparent and should meet the reasonable expectations of the individuals concerned. The principle is also a guide to robust ethical corporate behaviour generally - in a way that is familiar for the scientific research field where the robustness of the data protection system to be implemented is often a criteria in the context of funding application (Chassang, 2017). The ultimate goal of GDPR is to create legal certainty and sustainability of the data protection measures in a technological neutral approach. GDPR still applies to the data controllers and processors acting in the public and private sectors for profitable and not-profitable purposes. It still differentiates between two kinds of personal data by strictly regulating the processing of special categories of data (the so-called 'sensitive personal data' such as health data, genetic data and biometric data) because of their potential risks regarding the rights and freedoms of the data subject. It still also considers scientific research activities as a specific context of personal data processing where the equilibrium between individual freedom and the freedom of research triggers particular challenges and ethical issues, thus necessitating appropriate rules allowing both personal data processing and sharing in the pursuit of the public interest. GDPR adopts a new general risk-based approach intended to facilitate the case-by-case identification of data protection issues and the related necessary data protection measures to be respected.

Organisations that come within GDPR's jurisdiction need to be able to demonstrate they have a lawful basis for processing data – by obtaining the consent of the individual, fulfilling the terms of a contract or meeting the legitimate interest of the organization. They also need to be able to explain in clear terms what the processing is about, including the logic behind any automated decision-making. Where the processing activity presents a high risk to individuals or society, then a data protection impact assessment should be carried out and appropriate measures put in place to help mitigate the risk.

GDPR gives EU residents more control over their data and includes a right-to-erasure portion that allows people to request that companies delete their details in some cases. Besides ensuring that data gets collected legally, the law obliges the relevant entities to protect that information and safeguard it from misuse. GDPR also requires companies to anonymize their data, unless identifying information is crucial to its worthiness.

As a rule of thumb, quite apart from any legislative requirements, but key to ethical considerations, data should be sourced and shared responsibly. If organizations are unaware of the provenance of data and unsure whether data is properly protected when shared with third parties, the risk of data breaches rises.

On a more practical level, those organisations operating under the requirements of GDPR should carry out a data protection impact assessment for high-risk data processing, which could include supplementary questions on outcomes for customers and society. The data protection officer (a mandatory appointment for public bodies and for certain types of data processing activity) could offer advice on ethical processing, as could a stakeholder group, which would provide direction on the approach to be taken by the organisation. Professionals working with data need to take out identifying details before processing the information. Similarly, the businesses employing them should allow for training to occur or verify their workers know how to handle big data and how to manage documents, looking particularly at storage, access and auditing to avoid ethical violations and significant fines.

5.5.2. Ethics and the GDPR

An important perspective on GDPR is that it does recognize the contribution of technology to economic and social progress, but holds that technology should be developed in a responsible manner and, in particular, that individuals should have control over their personal data (Hijmans & Raab, 2018). For example, GDPR's emphasis on respect for the 'fundamental rights and freedoms' of natural persons is ubiquitous and GDPR specifies the fundamental right of data protection in a number of detailed rights of the data subject. The notion of fairness plays a key role in this respect. Fairness is an ethical dimension that is central to legal requirements for data protection under international and EU law, as well as national law. Article 8 of GDPR requires fair processing, whilst Article 5(1)(a) elaborates this and associates it with transparency.

The value of fairness is directly implicated in Article 22 of GDPR, which provides that individuals 'shall have the right not to be subject to a decision based solely on automated processing', save for exceptions. The same Article provides that – where an exception to the main rule applies – an individual has nevertheless a claim to obtain human intervention. Article 22 reflects the view that important decisions for individuals should be made by humans, not by mathematical models. If the way that data is processed is considered ethically questionable, results in unfair outcomes for individuals, or has an adverse effect on society, then it may infringe GDPR's fairness principle. And while notions of fairness can be difficult to frame, guidance is emerging. The European Commission issued a set of ethics guidelines for trustworthy AI earlier this year (AI HLEG, 2019). Key elements of the guidelines include transparency (processing decisions should be explainable), diversity and

non-discrimination (unfair bias must be avoided), and accountability (design processes should be assessed and auditable).

GDPR goes further into the realm of ethics, emphasizing principles that were not previously so prominent. These include transparency (or openness) and accountability, both of which address the ethical and practical relationship between controllers and processors of personal data and individuals, and reflect the OECD Principles (Hijmans & Raab, 2018). GDPR draws attention to the general societal interest in the protection of individuals' rights, as it does for example in Article 57(1)(b), which says that a supervisory authority must "promote public awareness and understanding of the risks, rules, safeguards and rights in relation to processing". There are practical reasons for raising the level of public understanding about information processing and rights, but there is also an ethical dimension: awareness-raising could contribute to the shaping of societal conditions and an 'ecosystem' for privacy and data protection that would be considered 'good' in the information age.

Thinking ethically about the processing of personal data often causes dilemmas because judgment may be involved in situations where the line between 'should' and 'should not' is not clear, thus GDPR always requires judgement. Data protection law cannot – and should not – be merely technically applied according to the Court of Justice of the EU (CJEU) (Hijmans & Raab, 2018), authorities are required to 'balance' individuals' rights with the economic interest of the free flow of data.

Moreover, GDPR itself contains a number of components that may require an ethical judgement, when applied. Article 24 of GDPR, the general provision on the obligations of the controller, introduces a risk-based approach. It specifies that the controller must take into account 'the risks of varying likelihood and severity for the rights and freedoms of natural persons'. Recital 75 describes specific risks and harms to rights and freedoms of individuals that deserve protection. Its application may involve an ethical judgement that includes balancing these risks of data use with the benefits of data use in a developing information society, but this judgement is not straightforward (Hijmans & Raab, Ethical Dimensions of the GDPR, 2018).

A key concept in GDPR is the principle of accountability as laid down in the aforementioned Article 24 of GDPR. This article requires data controllers to implement the necessary measures to ensure compliance and to be able to demonstrate that they have taken these measures. Both components are part of the legal responsibility of the controller, whereas the second component includes a procedural requirement to report, thereby demonstrating or giving an account of what they have done.

There is also a link to responsible innovation (or responsible research and innovation). Where, as it is often claimed, the law cannot keep up with new technology, it may be a task for business to ensure responsible innovation (Jakobsen, Fløysand, & Overton, 2019). To be responsible, innovation should be coupled with justified concerns for privacy, taking into account future uses of technologies and their possible privacy implications as well as consequences for other human rights.

5.5.3. Summary

GDPR is about the use and protection of personal data. Compared to the previous data protection legislation, it widens the scope of what is considered to be personal information and enforces much stricter policies on what information businesses are allowed to collect on individuals, along with putting more pressure on them to keep this information secure. The regulation aims to put the individual back in control of their data. GDPR is designed to protect personal data and sensitive personal data such as political views, medical details, passport or identification document scans. Modern businesses are built on data, with companies regularly compiling information on their customers' buying habits, browsing history, and even financial data. GDPR is designed to restrict the ease with which businesses and organizations can collect data for one purpose, and then continue using it long into the future.

Educational content for Computer Science students should include notions on governance and legislation, with special regards to the topics covered by GDPR.

Value 22

Develop teaching content that explains the overall purpose of GDPR, as well as highlighting some of the key Articles.

Value 23

Develop teaching content that looks at the broader issues around managing and storing personal data.

Value 24

Develop teaching content that explores document management, looking particularly at storage, access and auditing.

6. Conclusions from Part 1

The literature review conducted on pertinent Digital Ethics concerns and challenges highlighted many key areas for the development of Digital Ethics curricula, teaching and learning content for Computer Science students. An important insight is the significance of data and data ethics which are a cornerstone of many of the emerging technologies and applications including Artificial Intelligence, Pervasive Computing and Social Media platforms. Governance and regulation also has a key emphasis on data and data ethics, including the GDPR. Interestingly, differences were found in geographical perceptions of Digital Ethics, in particular regarding the attitudes towards it. Data on the perceptions of Digital Ethics differ across Europe. A further key insight in how a confluence of emerging technologies, e.g. sensors and IoT, cloud-enabled big data, and mobile platforms are enabling the fast creation of wide ranging applications, often at a pace that is quicker than allows for detailed analysis of the ethical implications of these applications. Hence we advocate the widespread use of consequence scanning (Brown, 2019), when developing new technologies and the importance of raising awareness in students, of teaching how to consider the ethical consequences in the development of ICT technologies by using imagination. Our literature review has resulted in a set of values that we will use to direct the development of teaching content for delivering Digital Ethics to Computer Science students. These values can be found in Appendix A of this report.

Part 2

Insights from academics, industry specialists and citizens

1. Introduction

Following on from the literature review we conducted and presented in Part 1 of this report, we conducted sessions with three groups of key stakeholders - academics, industry specialists and citizens to capture and understand their concerns about ICT and associated ethical issues. We were interested in their concerns and also if these concerns were covered by and reflected by the current academic literature on Digital Ethics. We found broad overlap between topics highlighted by the groups of stakeholders, but also some additional concerns swerve discussed by the stakeholders which went beyond what we found in the literature review. These new insights lead to an additional set of values that can be used by educators as guidelines to develop and deliver teaching content on Digital Ethics. We present the findings and the derived values for the three groups in this section of the report.

2. Methodology

2.1. Data Collection

Academics and industry specialists were interviewed as part of our first multiplier event, citizens were interviewed on a separate occasion, and the focus of the conversation was “The Teaching of Ethics to Computer Science Students”.

Three specific questions were developed to garner the participants’ specific concerns and insights into ethics aspects of ICT:

1. What ethical concerns do you have about new technologies?
2. What skills or training should people have to protect themselves in the online world?
3. What ethical training should be given from persons designing and developing technology, and who do you think should give that training?

While the third question directly investigates the topic of teaching ethics to Computer Science students, the first two questions were formulated in order to indirectly extract a number of topics and skills that should possibly be covered by the teaching material Ethics4EU will develop.

Participants were arranged seated at a round table, with the researcher exploring a combination of predetermined topics related to the main focus, as well as topics that arose

as part of the focus group session. The session lasted approximately one hour. This was done in an informal and relaxed manner, with the use of some predetermined questions, and some that emerged as part of the discussion[1].

Photos from the event (Dublin 21st November 2019)

As is typical with a focus group, the experiences and views of one participant lead to another participant building upon and amplifying that particular topic in a cascading effect, the researcher allowed this to occur on several occasions, but was careful to avoid groupthink by checking for agreement with participants who did not speak as part of the chain. The researcher also challenged any issues that there appeared to be in complete agreement by presenting a contrary opinion, to ensure that all views were considered and to avoid social desirability bias.

The session began with a question to obtain verbal consent from all participants to ensure they were willing to participate in the focus group, and additionally their consent was sought to allow an aggregate analysis of their views to be reported, as well as some limited quotations from individual participants. By agreement with the participants, the key points of the discussion were written down using pen-and-paper, but the participants did not unanimously agree to an audio recording, as they felt it might hamper and impair the discussion, therefore it wasn't done.

2.2. Data Analysis

Participants responses to the three questions were analysed using a thematic analysis, which is an approach for identifying themes or patterns of meanings with qualitative data. The key step in thematic analysis is coding the data, which involves attaching labels (or codes) to phrases or sentences of analytic interest. In this research a modified version of the coding process based on Gorden (1992) was followed.

[1] In a manner similar to a semi-structured interview.

1. *Define the coding categories*: When the focus groups were completed, the researchers familiarized themselves with the data by reading the notes many times. The researchers looked for patterns and themes across the data. Initially a *colour coding* approach was used to identify the main themes emerging from the notes. This involves highlighting different parts of the notes in different colours to represent initial themes. This gave the researchers the ability to look at emerging themes "at a glance", and to explore the balance of text that relates to each theme.

2. *Assign code labels to the categories*: From this first step, a preliminary tentative set of text codes were created to describe the computing ethics topics emerging from the notes. These codes replaced the coloured text. Examples of the initial codes included: *digital-literacy*, *where-law-overlaps-ethics*, *older-people-concerns.*

3. *Classify relevant information into the categories*: Following this initial process, the notes were re-read and the initial codes were attached to all relevant text. It was found that in some cases these codes were too general, e.g. one of the early themes was: *data-ethics*, which was later deemed to be too general; and in other cases the code were too specific, e.g. *acm-professional-code-of-ethics*.

4. *Refining the codes*: Following the identification of codes that were not fully suitable, those that were too general were further refined, e.g. *data-ethics* became *data-ethics-privacy*, *data-ethics-reliability*, *data-ethics-retention*, and *data-ethics-misuse*, and those codes that were too specific were merged, e.g. *acm-professional-code-of-ethics*, *employee-responsibilities* and *organizational-specific-guidelines* became *professional-ethics*. Some other codes were renamed, e.g. *relevant-European-laws* became *importance-of-GDPR*. This was an iterative process, and took place over a period of three weeks to complete.

5. *Test the reliability of the coding*: The reliability of the coding was tested by asking an independent reviewer to code one of the notes without having access to our coding process (this is called an *independent-coder* method). There was a strong overlap between the two coding processes, thereby validating the approach.

The final codes are the key themes of the notes, which are described in the subsections below. Each theme is highlighted in bold. Responses from the different groups are presented comparatively for each question. Where there were multiple groups of participants, we have combined the responses. Demographic data on the participants to the focus groups can be found in Appendix B of this report.

As most of the participants didn't have any specific expertise related to Digital Ethics, they often expressed concern or gave us their opinion on unrelated topics (e.g. legal issues); mentioned generic examples; used improper or simplistic terminology (e.g.

interchangeability of "moral" and "ethical"); reported debatable generalizations. While these are underlined as the participants' opinions and thoughts in the text, we believed it was of special interest for the aims of the project to report on all the topics discussed.

3. Results

Participants' responses to the three questions were analysed using a thematic analysis. The analysis for each question is presented in the next subsections.

3.1. Question 1: What Ethical Concerns do you have about new technologies?

3.1.1. Industry Responses

There was agreement amongst the participants that one of the major areas of concern was the ***on-going automation of activities that were previously undertaken by human beings*** (this concern was further exacerbated for the group when the automation is achieved through Machine Learning). Participants discussed what to do with the people when their jobs are being replaced by machines - "*we're developing technology that is taking jobs away from people, and although they can reskill, it is not clear that enough new jobs will be created to replace those that are going to be lost*".

One of the most discussed considerations was the challenge of ***bias in automated decision-making systems***, and this issue was looked at from the interrelated perspectives of bias in datasets, and bias in machine learning algorithms. When discussing bias in datasets participants highlighted the potential dangers of using open datasets, which may not have been analysed for completeness, and "*may exclude particular populations, for example, those on the margins*" and yet the conclusions derived from the analysis of these datasets may be presented as fact. There was also a good deal of discussion of potential historical patterns of bias in datasets (including issues around gender and race), and how to prevent those historical issues from being propagated. Suggestions for solutions included *"exploring patterns for bias in data, looking at statistical variances, examining who owns or controls the data and how the datasets were created"*. Participants suggested that this type of analysis "*should look at composite biases that are more difficult to detect, for example, hiring women over a certain age in employment practices"*. It was also suggested that the "*GDPR guidelines are useful for exploring bias"*.

There were discussions around ***environmental considerations in the collection of that data***, particularly in the context of IoT (the Internet of Things). Another participant mentioned the role of data centers contributing to the environmental impact.

One organisation that was highlighted for their excellence in ethics was German Software Company, SAP SE, who incorporates a great deal of ethics in their graduate training

programmes. A representative from SAP described how their training "p*rovided three examples of where there had been ethical compliance breaches, and one example looked at a senior female colleague working in Israel and Korea and how she was treated differently in different countries*." This discussion led onto a conversation on ***regional and cultural differences in ethical standards.***

Participants discussed ***IT professional standards and their relation to ethical standards***, in particular if IT professional standards can ensure that ethical standards are adhered to. Some participants felt it was important to note that very often employees don't necessarily have control over their work and their use of data and code, and may not be aware of where their work will be used or how it can be repurposed and that there is no clear guidance in IT professional standards in such cases.

Finally all participants agreed that students should be educated ***about legal frameworks with relevant ethical aspects*** including Confidentiality Agreements, Non-Disclosure Agreements and Intellectual Property.

3.1.2. Academic Responses

The academic participants began by discussing their concerns about ***data and datasets*** – one participant remarking "*I worry about the amount of personal information that is being kept, I think we are keeping a lot more data than we need*." Both groups felt that a lot of organisations seem to be collecting as much data as possible with no clear purpose, other than feeling there is something of value in the data. They also expressed their concerns about ***completeness and representativeness of datasets*** (particularly for open source datasets), and the ***potential bias that an AI system might embody*** in its decision-making if trained on these datasets. This led to a discussion on the ownership of decisions in an online context, for example, if a search engine is suggesting search phrases, and suggesting sites to visit, who is really making the decisions – one participant said "*where is the line between me and the application?*" Another concern raised was the dangers of ***technological or digital colonisation in developing countries*** and a lack of control of their data on the part of citizens in the developing world.

Another concern was that datasets might exclude certain people for ***privacy*** reasons because they might be identifiable due to their specific characteristics, and therefore could be unrepresented in the overall datasets. This conversation about ***marginalisation*** led to a discussion on how some voice recognition technologies, for example, one participant remarked how voice assistants such as Alexa, might not be as effective for people with speech impediments or strong regional accents.

On the theme of ***privacy***, participants expressed concern over the possibility of governments or private organisations combining various data "*personal, legal and financial information"* about an individual and the impact that might have. One example cited by a participant was

the South Korean government limiting the number of (and size of) gambling transactions that any one person can take part in per day when gambling online. Even though there was general agreement that too much gambling is a bad thing, nonetheless the notion of the government being able to restrict an individual's liberty was considered objectionable, and they felt that education and help is preferable to controlling them. Participants remarked that a combination of control and educational measures have been used by governments to reduce the number of people who smoke.

Both groups highlighted the importance of cybersecurity as being an ethical imperative in the context of the significant amount of personal data being collected on individuals, as well as the numerous high profile data leaks that have occurred.

Both groups discussed research projects and the importance of ***ethics for research projects***, particularly when funded by public monies. They discussed the importance of transparency in research projects and how this should be communicated to the public. For example participants discussed how it is important to make clear "*whether enticements or incentives were permitted.*" They discussed the importance of research communication more generally and how this is an ethical issue, for example how research findings are presented to the general public who may not be familiar with the general area or the specific details. Finally both groups were agreed on the importance of teaching students computing ethics in the areas of plagiarism, copyright and the honest presentation of results.

3.1.3. Citizen Responses

Many of the concerns of the citizen participants related to ***data***. Some common themes were around ***collection and retention of data*** and ***misuse of data***, in particular how data collected on individuals may be used and misused. Data gathering with the purpose of creating profiles, for example voter profiles with the purpose of influencing democratic elections were mentioned frequently by the participants.

Privacy was another topic discussed by all participants, in particular data appropriation on behalf of privately owned businesses, which do not reward their users for the acquisition of such data and use it to maximize their profits. Others raised concerns about ***third parties gaining access to data without the original users consent***. Concerns were raised about applications (e.g. Facebook, Siri) that can listen to conversations through digital devices and use that information for targeted advertising. The balance between commercial gain and social benefit was discussed.

Respondents cited a number of ***concerns related to social media*** including the ***normalisation of unacceptable behaviours***, inappropriate exposure to information and content***, widespread dissemination of misinformation***, the fuelling addictive tendencies and preying on vulnerable individuals. Social media platforms can have negative affect on individuals psyche leading to mental health issues and reduce genuine human interactions and empathy. Particular concerns were raised about the influence of social media on

teenagers and young people, for example cyberbullying and a lack of awareness on the part of young people about the longevity of data online. One participant said "*people who come into the public eye are likely to have their social media postings from decades earlier where they were perhaps a younger and less-informed person, examined for any failures to be used against them*". There were also concerns about misinformation online and the sharing of 'fake news' via social media, and the detrimental impact that can have on younger, most impressionable people, one participant noted "*As a parent I wonder how safe my teenagers are online, they lack the experience and skills to distinguish reputable sources from fake news*".

The participants also felt that ***surveillance of individuals, in particular using facial recognition technology*** was of great concern, where surveillance companies are potentially storing images and information about people who are unaware they are captured on camera, for example from walking on the street, going shopping or entering other commercial buildings. One participant notes that there cultural differences in how surveillance technologies are being used "*Surveillance is creeping up in prevalence but thus far I believe it is being used for legitimate and good purposes, for solving crimes, preventing crimes, and finding missing persons, in Ireland at least. I wouldn't be quite so sure about other countries and other one-party states*."

Based on the responses to Question 1 (What Ethical Concerns do you have about new technologies?), we have produced a number of Values (25-34 below) for the development of Digital Ethics content for Computer Science students.

Value 25

Develop content that explains the costs of development, the costs of adding in ethics checks, and the potential costs of ethical violations being subsequently discovered after the product has been completed.

Value 26

Develop content that teaches students about bias in data and bias in decision-making systems, and discusses statistical techniques in order to facilitate the identification and mitigation of those biases, including Differential Privacy.

Value 27

Develop content that teaches students to put people (users) at the centre of all developments, and to learn about empathy, and to learn about informed consent.

Value 28

Develop content that teaches students how to explain both technical terminology and to explain the functionality of models in clear and understandable language, particularly in areas such as Artificial Intelligence. Emphasize the important role of communication with stakeholders (people who are affected by the technology).

Value 29

Develop content that teaches students about the range of ways data is collected both online (transactions, cookies, social media) and offline (Internet of Things, Digital Assistants), and ways of handling sensitive data.

Value 30

Develop content that explains GDPR and its implications. Also, look at other legal aspects of data management, including confidentiality agreements, and give the students a general appreciation of how the law works.

Value 31

Develop content that encourages students to explore the differences between the meaning of "legal" and "ethical".

Value 32

Develop content that highlights the environmental impact of technology.

Value 33

Develop teaching content to explore the dangers of technological colonization, and the subtle and obvious pressures of funding in third-world countries.

Value 34

Develop content that explains different ethical models (e.g. deontological, utilitarianism, etc.).

3.2. Questions 2: What skills or training should people have to protect themselves in an online world?

3.2.1. Industry Responses

The group analysed this question from the point-of-view of their work as software designers and developers. There was a discussion centred around the notion of ***consequence scanning,*** where designers and developers try to predict the consequences of developing the software they are asked to create. For example, designers and developers should consider what the software should and should not do, the worst possible negative consequence of the software and how the software would work if repurposed for another system.

Generally the participants reflected that a lot of employees in organisations don't get to see the "big picture", and therefore don't have the opportunity to evaluate the ethical implications of the processes that they are involved in. On the other hand, managers who have the bigger picture, but don't know the exact detail of how systems have been developed might also be missing out of some of the ethical implications of how the work of different designers and developers impact each other from a moral point of view.

Another consideration that was discussed was the ***dangers of using off-the-shelf code***, particularly when used by naïve or novice designers and developers, who may not have thought of the complete ethical implications of using that code, or may not have full information on how the off-the-shelf code works, and therefore have no awareness of the potential ethical issues. The conversation moved to the important role that educators must play in exploring these issues.

The participants reflected that there is a need for ***more diversity in the IT profession***; importantly as trainers, as designers, as developers, and as testers, so that "*they can ask the ethical questions that others don't think of*."

3.2.2. Academic Responses

The groups looked at this question from the point-of-view of teaching students how to design and develop a computer system. The groups, began by discussing considerations that students should have before designing and developing a computer system. These considerations could broadly be characterised as ***consequence scanning***. For example, what is the best outcome of this development? What is the worst unintentional outcome that could happen? How would I mitigate the worst outcome if it happened?

There was a general agreement of the importance of **"*always keeping a human in the loop*",** and particularly to ensure that there is consultation with persons with significant domain

knowledge as they will understand in what context the technology will be used. There was also agreement that we have to encourage designers and developers to think more reflectively and "*consider what the system they are developing is really for, and is really about*". One participant suggested a possible scenario where a developer was asked to create software following a specification, and it became evident that the system as a whole was designed to make the software addictive, even though it wasn't evident to the individual developers, what should they do?

There was also general agreement that ***explainability*** is vital in all ***automated decision-making*** processes. This explainability concept refers to both the ability to understand terminology such as "features" and "weights", but also that each individual decision that the system takes can be explained. As part of this conversation, participants questioned whether or not it is appropriate to develop systems using partially correct data. Another participant mentioned the use of software libraries like LIME for Python which help explain how machine learning systems make specific predictions.

The groups agreed that there is a need for designers and developers to be aware of the law as it pertains to them, and ***where the law overlaps with ethics***. Although there was general agreement that often ethical principles can be of a higher standard than the law, but the groups wondered if there is one set of ethical principles that should be followed by everyone. Further to this, there are different laws in different countries, and there may even be different ethical standards in different regions that developers should be aware of. The topic of outsourcing was discussed, where systems can be developed in one region but used in another and different ethical standards can apply in the different regions.

The groups also discussed the nature of ethical standards, and wondered where can you find (and find out about) standards and ***how ethical standards can be enforced***. They also questioned ***whether ethical standards can keep up with the rapid developments of software***. An issue discussed was the impact of unethical behaviour which impacts commercial activities, but more importantly impacts people and consumers. One participant highlighted ***opaque Terms & Conditions*** that many users sign up to without reading, and agreeing to things they either don't read or don't fully understand.

Another key issue discussed was ***accessibility*** and the importance of ensuring that as wide a range of people as possible can use the software being developed. One participant commented *"Sometimes the client might not be aware of, or concerned with, accessibility consideration, but does that mean the developers shouldn't consider it?"*

Finally there was some discussion on how equipped academics are to teach this type of content, and ***what sort of training or teaching content is required by academics*** so that they can become confident in teaching this topic. Both groups felt that such content should be publicly available to private and public organisations.

3.3.3. Citizen Responses

Participants in this group took a broad view of the question. The issue of the ***longevity of digital information*** was raised as a concern, particularly social media posts which may be innocuous in the context in which they were created, but could be potentially misunderstood or misrepresented without that context, and could be detrimental in the future. As one participant phrased it: "*For all groups they need to be made aware, understand implications of posting information to a world that is never deleted, follows them around forever, for example, years old tweets coming back to haunt people and impact on work opportunities*"

Participants felt that people should ***understand how data can be obtained*** by others and be taught about their digital security and online platforms including privacy settings. As one participant said: "*People are unaware they are 'the product' in many cases. I think the phone companies and social media companies need to be much more transparent when people open accounts about ethical issues around obtaining their data*" Another participant gave the example of digital assistants like Siri or Alexa "*always listening to conversations in the home and using the information for marketing purposes*". At a more fundamental level ***digital literacy*** was considered extremely important, one participant remarked "*People should know the basics of digital literacy, cookies come to mind. I don't fully understand these, yet every website asks to allow them be used*."

Both groups agreed that parents need specific training to help them navigate the online world and to understand the implications of having a digital presence. They suggested that training should include using parental controls (including monitoring tools) and other ways of securing devices (in hardware and software), knowing some of the key social media applications, how to deal with cyberbullying, and how to talk to their children "*about the positives and negatives of social networks, and ensuring they keep the channels open to enable children to discuss issues or bumps they encounter in their cyber journeys"*. Another participant suggested that parents and children "*should be taught about risks, to have an awareness of strangers online and false accounts or information".* They also felt it would be extremely helpful to learn about the addictive nature of social media applications and smartphones, and advice on how to limit their children's use of these technologies. They stated it would be helpful to know what is legal and illegal in terms of sharing and downloading of audio and video files.

Both groups also agreed that ***older people need training to help them navigate the online world***, particularly if it could be tailored for their interests, including lifestyle application (health, banking, shopping) and privacy settings on their devices and applications. Most participants felt that training about scams and fraud would also help, as well as general personal data protection online. All participants felt that some older people may not want to (or be able to) access digital services, and therefore offline services (government services, libraries, postal services) should be maintained for this age group if that is their preference. One participant commented that "*the move to online services is regrettable particularly*

since some older people may not have a laptop or smartphone, may not have an Internet connection (their area may not have coverage), or may not be comfortable using online banking applications, and maybe therefore be far more vulnerable to phone scams".

Participants also mentioned that "*voluntary organizations, clubs, societies also need training on the do's and don'ts of social media and use of members' data"*. They also felt these groups should be trained on how to recognize and report inappropriate content. Finally they felt that "*all elements of security awareness from spamming, phishing, identifying secure sites, should be taught, with real life examples*".

Based on the responses to Question 2 (What skills or training should people have to protect themselves in an online world?), we have created a set of values (35-44 below) to guide developers on the creation of Digital Ethics content.

Value 35

Develop content that teaches Software Methodologies that incorporate Consequence Scanning.

Value 36

Develop content that teaches students diversity, and the need for it in datasets, in interface design, and in software teams.

Value 37

Develop ethical scenarios to prepare students for potential workplace dilemmas. Also develop scenarios that place similar ethical dilemmas in different contexts to help the students explore if their ethical perspectives are consistent.

Value 38

Develop content to encourage students to be sufficiently confident in themselves to question the ethics of their workplace.

Value 39

Develop content around the potential issues of datasets, including the amount of data being stored, the representativeness of the data, the security of the data, and the dangers of partial data and of biased data.

Value 40

Develop a series of generic ethical frameworks that cover categories of ethical dilemmas that will provide future-proofing for new and upcoming technologies.

Value 41

Develop content that teaches people the importance of "Terms & Conditions" and develop content that helps them understand what the terms mean.

Value 42

Develop content that prompts learners to ask relevant questions at each state of the data management life cycle.

Value 43

Develop case studies based around the general area of social media, looking at things such as how many individual technologies, designed and developed for disparate purposes can be repurposed and combined to develop different products and services that tap into unmet customer needs. Social media case studies offer the opportunity to show how multiple pertinent Digital Ethics issues, e.g. data, informed consent, AI and privacy need to be considered in singular technology applications.

Value 44

Develop tools and materials for consequence scanning for social media and related applications. Such applications have shown the influence of technology on the behaviour of a society, not simply that of individuals. This highlights the importance of consequence scanning, considering the potential consequences – intended and unintended of new technologies.

3.3. Questions 3: What ethical training should be given to persons designing and developing technology and who do you think should give that training?

3.3.1. Industry Responses

The first topic that the participants discussed in some detail was ***the importance of GDPR –*** (the General Data Protection Regulation GDPR 2016/679), and the importance of ensuring that students know how to perform data protection impact assessments, be able to handle data securely, and realise the importance of handling sensitive data before graduation. There was also agreement that a work placement during a university course can be very

beneficial in terms of learning the importance of Digital Ethics, and can teach the students lessons that may be more difficult to teach in the classroom.

The discussion then turned to what other skills need to be and participants suggested "*how to develop empathy*" and "*respecting social norms*" as fundamental skills. Additionally there was agreement on the importance of giving students that ability to deal with situations where they are asked to do something unethical, including "*the tools to ask further questions in situations where there appears to be unethical activities occurring to gain a deeper understanding of the situation*". Participants also noted that another way to help students develop an appreciation of Digital Ethics is to help them understand the reason why ethics are being breached.

Participants felt that the most effective way to successfully teach ethics would be to ***incorporate ethics it into existing modules*** rather than creating a dedicated separate module, and it was suggested that something as simple as writing a small reflective piece on how data protection or computing ethics applies to a specific module might be a way of raising awareness, and to consider having someone external to the module review the content of these pieces.

Finally everyone agreed that ethics is a broad societal issue, it is up to everyone to help develop their personal, and societal, understanding on ethical standards.

3.3.2. Academic Responses

The groups stressed the importance of ***communication, teamwork, and most especially a sense of personal responsibility as key skills*** that need to be taught as part of computing ethics courses. One participant mentioned the importance for graduates to understand the concepts of *"informed consent and voluntary consent"*. All participants agreed that "*confidence is also a vital skill in graduates, including having the confidence to ask difficult questions of colleagues and of management*", as well as having the confidence to speak up when unethical issues arise, and having the confidence to adapt to a changing environment. A related discussion involved how to equip graduates with the ability to think critically about their work once working in a company, and to appreciate that there may be conflict between ethics and profit.

Participants also felt that it was vitally important that graduates be equipped with a good understanding of data science giving the recent advances in the topic. Participants felt is important for students *"to explore the power and dangers of aggregate data, including a discussion of how to make an individuals' data private using techniques such as differential privacy"*. Another participant commented that students should be equipped with "*an understanding of both bias and representativeness of datasets"*.

The groups discussed ***the relationship between ethics and the law***, and how knowledge of GDPR legislation is very important for all graduates. From there the discussion moved onto the potential conflict between legal issues and ethical issues, and what choices the

graduates should make under those circumstances. One participant commented *"it is important that every organisation should have a clear code of ethics, and promote that code, and promote ethical thinking in their organisation".*

3.3.3. Citizen Responses

The participants felt that the key skills for people designing and developing technology is "*sympathy and empathy to think about the people who will be using the technology*", and that it is *"important to realise that not everyone is a technology wizard*". They felt that using new technology can be difficult for many end users and that many of them may not even think about or understand the ethical implications of technology. The participants also noted that even if they fully understand how to use the system "*that doesn't mean I fully understand how it works, so there might be a whole layer of ethical issues that are not visible to me".*

The groups felt that ***ethical training about "laws, codes, and policies"*** are also very important, and both groups mentioned that the people designing and developing technology must also have the confidence and courage to ask questions of their organisations, and ask questions of themselves to ensure that the highest ethical standards are being adhered to.

The majority of participants thought that Universities should be responsible for the ethical training of persons designing and developing technology. Some did suggest the responsibility lies with employers, and a few thought it should be part of a continuing professional development and others thought it should be an individual's personal responsibility.

Based on the responses to Questions 3 (What ethical training should be given to persons designing and developing technology and who do you think should give that training?), we created the following set of values (45-50) to help guide the development of educational content for teaching Digital Ethics.

Value 45

Develop content that encourages students to understand how data is used by applications and how to make that use and the value of their data transparent to end users.

Value 46

Develop social media case studies looking at multiple pertinent Digital Ethics issues, e.g. information, misinformation and 'fake news', and digital literacy surveillance, the influence of social media on teenagers and young people.

Value 47

Develop social media case studies looking at the influence of social media on teenagers and young people (particularly mental health), and the exclusion of older people.

Value 48

Develop teaching content to explore diversity as a catalyst to better technology design.

Value 49

Develop teaching content to highlight the importance of transparent technologies that explains to users how they are collecting, aggregating and interpreting their data. The content should also highlight the importance of policies and regulations, particularly those that protect the right of individuals to retain control over their data, and other general control mechanisms that should be made explicit to users.

Value 50

Develop teaching content to explain and discuss the widespread use of surveillance technologies, including surveillance Smart Speakers and cameras equipped with facial recognition.

4. Conclusions for Part 2

Discussion with industry, academia and citizen stakeholders were very fruitful and yielded wide insights into the digital ethical concerns of these stakeholders. The topics highlighted during the focus group were generally in keeping those topics uncovered in literature review, for example, industry and academic experts also highlighted the importance of data as a key topic in Digital Ethics. Industry and academic participants also highlighted concerns about the accuracy, completeness and representativeness of datasets used to develop applications which is also a topic with good coverage in the literature review. Industry and academic participants also highlighted other issues that were not uncovered in the literature review such as how it is important to "always keep a human in the loop", particularly to ensure that there is someone with significant domain knowledge kept in the loop, as they will understand in what context the technology will be used. There was strong agreement about teaching Computer Science students technical skills within different ethical contexts (e.g.

data handling), case studies highlighting ethical aspects across the software lifecycle and relevant legislation which is addressed by relevant literature. Going beyond the literature, the focus groups highlighted the importance of softer skills such as empathy and how to respect social norms.

Citizens Digital Ethics concerns also chimed with the results of our literature, for example there were concerns about the impact of technology on surveillance, privacy, and vulnerable groups such as children and older people. In particular they were concerned with the potential negative impacts of technology such as social media on teenagers and young people, in particular their mental health and the exclusion of older people by increasing technological advances. Going beyond our literature review, citizens expressed concerns at low levels of digital literacy among the general public and how that may make them susceptible to fraud and misinformation online.

All groups (industry, academia and citizens) agreed that universities have a key role to play in the education, teaching and training in Digital Ethics. The analysis of data gathered from the three groups resulted in a set of values that will be used to direct the development of teaching content for delivering Digital Ethics to Computer Science students. These values can be found in Appendix A of this report.

Bibliography

ACM. (2018). *ACM Code of Ethics and Professional Conduct*. Retrieved from https://www.acm.org/code-of-ethics

ALLEA. (2013). *Ethics Education in Science.* All European Academies Permanent Working Group on Science and Ethics. Retrieved 25-03-2021, from https://www.allea.org/wp-content/uploads/2015/07/Statement_Ethics_Edu_web_final_2013_10_10.pdf

Al-Saggaf, Y., Burmeister, O., & Schwartz, M. (2017). Qualifications and ethics education: the views of ICT professionals. *Australasian Journal of Information Systems*.

Anderson, M., & Anderson, S. (Eds.). (2011). *Machine Ethics.* Cambridge: Cambridge University Press.

Barocas, S., & Nissenbaum, H. (2014). Bid Data's End Run around Anonymity and Consent. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), *Privacy, Big Data, and the Public Good*(pp. 44-75).

Bovens, M. (1988). *The Quest for Responsibility. Accountability and Citizenship in Complex Organisations.* Cambridge: Cambridge University Press.

boyd, d., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society, 15*(5), pp. 662-679.

Brey, P. (2012). Anticipatory Ethics for Emerging Technologies. *Nanoethics, 6*, pp. 1-13.

Brown, S. (2019). *Consequence Scanning Manual Version 1.* London: Doteveryone.

Buch, V., Ahmed, I., & Maruthappu, M. (2018). Artificial Intelligence in Medicine: Current Trends and Future Possibilities. *British Journal of General Practice, 68*(668), pp. 143-144.

Bunge, M. (1975). Towards a Technoethics. *Philosophic Exchange, 6*(1), pp. 69-79.

Bynum, T. (1999). The Development of Computer Ethics as a Philosophical Field of Study. *Australian Journal of Professional and Applied Ethics, 1*(1), pp. 1-29.

Bynum, T. (2000). The Foundation of Computer Ethics. *Computers and Society, 30*(2), pp. 6-13.

Bynum, T. (2006). Flourishing Ethics. *Ethics and Information Technology, 8*, pp. 157-173.

Bynum, T. (2018, Summer Edition). *Computer and Information Ethics*. (E. Zalta, Editor) Retrieved from Stanford Encyclopedia of Phylosophy: https://plato.stanford.edu/archives/sum2018/entries/ethics-computer/

Capurro, R. (2009). *The Global Impact of ICT on Society and the Environment.* Retrieved 28-07-2020, from Digital Ethics: http://www.capurro.de/korea.html

Chassang, G. (2017). The Impact of the EU General Data Protection Regulation on Scientific Research. *Ecancermedicalscience, 11:709*.

Colman, F., Bühlmann, V., O'Donnell, A., & van der Tuin, I. (2018). *Ethics of Coding: A Report on the Algorithmic Condition.* Brussels: Publications Office of the European Union.

EGAI. (2019). *Ethics Guidelines for Trustworthy AI.* High-Level Expert Group on AI. Publications Office of the European Union. Retrieved 28-07-2020, from https://ai.bsa.org/wp-content/uploads/2019/09/AIHLEG_EthicsGuidelinesforTrustworthyAI-ENpdf.pdf

EGE. (2012). *Ethics of Information and Communication Technologies.* European Group on Ethics in Science and New Technologies. Brussels: Bureau of European Policy Advisors.

EGE. (2018). *Artificial Intelligence, Robotics and "Autonomous" Systems.* European Group on Ethics in Science and New Technologies. Luxembourg: Publications Office of the European Union. Retrieved 28-07-2020, from http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

European Union. (2016). *The Precautionary Principle: Definitions, Applications and Governance.* European Parliament, Directorate-General for Parliamentary Research Service. Members' Research Service. Retrieved 28-07-2020, from https://www.europarl.europa.eu/RegData/etudes/IDAN/2015/573876/EPRS_IDA(2015)573876_EN.pdf

Faramelli, N. J. (1971). *Technoethics: Christian Mission in an Age of Technology.* Friendship Press.

Fesmire, S. (2003). *John Dewey and Moral Imagination: Pragmatism in Ethics.* Indiana University Press.

Flanagan, M., Howe, D., & Nissenbaum, H. (2008). Embodying Values in Technology: Theory and Practice. (J. van den Hoven, & J. Weckert, Eds.) *Information Technology and Moral Philosophy*, 322-353.

Flick, C. (2015). Informed consent and the Facebook emotional manipulation study. *Research Ethics*, 2016: 12(1), pp. 14-28.

Floridi, L., & Taddeo, M. (2016). What is Data Ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2083).

Friedman, B., & Hendry, D. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination.* MIT Press.

Friedman, B., Kahn, P., Borning, A., Zhang, P., & Galletta, D. (2006). Value Sensitive Design and Information System. In E. E. Technologies. M. E. Sharpe.

Frodeman, R. (Ed.). (2017). *The Oxford Handbook of Interdisciplinarity* (2nd ed.). Oxford University Press.

Gero, J. S. (1975, December). Ethics in Computer-Aided Design: a Polemic. *ACM SIGDA Newsletter(5.4)*, 9-14.

Gottardello, D., & Pàmies, M. (2019). Business School Professors' Perception of Ethics in Education in Europe. (MDPI, Ed.) *Sustainability, 11*(3), pp. 608-627.

Gotterbarn, D. (1991). Computer Ethics: Responsibility Regained. *National Forum, 71*(3), pp. 26-31.

Hauptman, R. (1988). *Ethical Challenges in Librarianship.* Phoenix: Oryx Press.

Hijmans, H., & Raab, C. (2018). Ethical Dimensions of the GDPR. In M. Cole, & F. Boehm (Eds.), *Commentary on the General Data Protection Regulation.* Cheltenham: Edward Elgar.

HiPEAC. (2019). *HiPEAC Vision 2019.* High Performance and Embedded Architecture and Compilation. HiPEAC. Retrieved 28-07-2020, from https://www.hipeac.net/vision/2019/

Hirsh Hadorn, G., Hoffmann-Riem, H., Biber-Klemm, S., Grossenbacher-Mansuy, W., Joye, D., Pohl, C., . . . Zemp, E. (Eds.). (2008). *Handbook of Transdisciplinary Research.* Berlin: Springer.

Johnson, D. (1985). *Computer Ethics.* New York: Prentice-Hall.

Lessing, L. (1999). *Code, and Other Laws of Cyberspace.* Basic Books.

Maner, W. (1980). *Starter Kit in Computer Ethics.* New York: Helvetia Press and the National Information and Resource Center for Teaching Philosophy.

Maner, W. (1996). Unique Ethical Problems in Information Technology. *Science and Engineering Ethics, 2*(2), pp. 137-154.

Moor, J. (1985). What is Computer Ethics? *Metaphilosophy, 16*(4), pp. 266-75.

Müller, V. C. (2020, Fall Edition). *Ethics of Artificial Intelligence and Robotics*. (Z. E. N., Editor) Retrieved from Stanford Encyclopedia of Phylosophy: https://plato.stanford.edu/archives/fall2020/entries/ethics-ai/

Polonsky, M., Brito, P., Pinto, J., & Higgs, K. (2001). Consumer ethics in the European Union: A comparison of northern and southern views. *Journal of Business Ethics, 31*(2), pp. 117-130.

O'Keefe, K., & O'Brien, D. (2018). *Ethical Data and Information Management: Concepts, Tools and Methods*. Kogan Page: London.

Rachels, J. (1975). Why Privacy is Important? *Philosophy and Public Affairs, 4*(4), pp. 323-333.

Regan, P. M. (2015). Privacy and the common good: Revisited. In *Social Dimensions of Privacy*, pp.50-70. Cambridge: Cambridge University Press.

Righby, M. J. (2019). Ethical Dimensions of Using Artificial Intelligence in Health Care. *AMA Journal of Ethics, 21*(2), pp. 121-124.

Searle, J. R. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences, 3*(3), pp. 417-457.

Swiss Academies of Arts and Sciences. (n.d.). *Methods for coproducing knowledge*. Retrieved 11-07-2020, from http://www.transdisciplinarity.ch/en/td-net/Methoden.html

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2019). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.* First edition: IEEE.

Thornley, C., Murnane, S., McLoughlin, S., Carcary, M., Doherty, E., & Veling, L. (2018). The Role of Ethics in Developing Professionalism Within the Global ICT Community. *International Journal of Human Capital and Information Technology Professionals, 9*(4).

Van de Poel, I., & Royakkers, L. (2011). *Ethics, Technology and Engineering: An Introduction.* Wiley-Blackwell.

Van den Hoven, J. (2007). ICT and Value Sensitive Design. In P. Goujon, S. Lavelle, P. Duquenoy, & K. Kimppa (Ed.), *The Information Society: Innovation, Legitimacy, Ethics and Democracy In honor of Professor Jacques Berleur. 233*, pp. 67-72. IFIP ACT, Springer.

Van den Hoven, J., Blaauw, M., Pieter, W., & Wartnier, M. (2020, Summer Edition). *Privacy and Information Technology*. (E. N. Zalta, Editor) Retrieved from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/archives/sum2020/entries/it-privacy/>

Van Wynsberghe, A., & Robbins, S. (2019). Critiquing the Reasons for Making Artificial Moral Agents. *Science and Engineering Ethics, 25*(3), pp. 719-735.

Verbeek, P. P. (2011). *Moralizing Technologies: Understanding and Designing the Morality of Things.* The University of Chicago Press.

Von Shonberg, R. (2013). A Vision of Responsible Innovation. In R. Owen, M. Heintz, & J. Bessant (Eds.), *Responsible Innovation* (pp. 51-74). London: John Wiley.

Ware, W. H. (1973). *Records, Computers and the Rights of Citizens.* RAND.

Winfield, A., Katina, M., Pitt, J., & Evers, V. (Eds.). (2019). Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems. *107*(3), pp. 501-632.

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus, 109*(1), pp. 121-136.

World Commision on the Ethics of Scientific Knowledge and Technology. (2006). *The Precautionary Principle.* UNESCO. United Nations Educational, Scientific and Cultural Organization. Retrieved 28-07-2020, from https://unesdoc.unesco.org/ark:/48223/pf0000139578

Zuboff, S. (2019). *The Age of Surveillance Capitalism.* Public Affairs.

Appendix A. Complete Set of Values for Developing Educational Content for Teaching Digital Ethics

Based on the literature reviewed in Part 1 and the analysis of the focus groups in Part 2 of the report, we have developed a set of 50 values for teaching Digital Ethics. These values are dispersed throughout the report at relevant parts and we list all them here in Appendix A. These values provide direction to educators for the development and delivery of teaching content for Digital Ethics by elucidating important aspects highlighted in the literature and underscored by members of our focus groups.

How to use this Appendix?

The 50 values are here classified using a thematic analysis to produce 8 categories of Digital Ethics topics:

1. Foundations of Digital Ethics
2. Digital Ethics Values
3. Data Ethics
4. AI Ethics
5. Ethics for Pervasive Computing, Privacy and Surveillance
6. Social Media Ethics
7. The relationship between Digital Ethics, Digital Regulations and Digital Governance
8. Professional Ethics

An educator who wishes to integrate the teaching of these topics with Digital Ethics content might want to go to the relevant section of this Appendix and check if their content aligns with the related values as they collate important aspects highlighted in the literature and underscored by members of our focus groups.

Foundations of Digital Ethics
Value 1: *Develop teaching content that explains the history and evolution of Digital Ethics, the debate about the different approaches, the difference between ethics and morality and that ethics is not a manual with answers: it reflects on questions and arguments concerning the moral choices people can make. Ethics is a process for searching for the right kind of morality, also in the case of Digital Ethics.*
Value 2: *Create timelines for the key events in the history of Digital Ethics and introduce and discuss with students some reference to the main ethical frameworks (e.g. utilitarianism, virtue ethics, etc.) developed by ethics in general.*
Value 34: *Develop content that explains different ethical models (e.g. deontological, utilitarianism, etc.).*

Digital Ethics Values
Value 4: *Develop teaching content that explains Europe's fundamental values of fundamental rights, democracy and the rule of law. Also create content that explains ethics are not static concepts, but rather change as other things change, such as technologies, conceptualizations and assemblies; and moral perspectives vary somewhat geographically.*
Value 5: *Develop teaching content that explains digital identity, the right to privacy, and personal data safety.*
Data Ethics
Value 3: *Develop teaching content that provides an overview of the ethical challenges related to machine learning, data science, algorithmic decision-making and other related AI issues.*
Value 6: *Develop teaching content that prompts learners to ask relevant questions at each stage of the data management life cycle, looking at the Dataethics.eu questionnaire as a teaching tool.*
Value 7: *Develop teaching content that explores the power of, and the ethical challenges associated with, aggregate data.*
Value 8: *Develop teaching content that explores the additional ethical considerations of using cloud architectures to store data.*
Value 9: *Develop teaching content that highlights specific examples of sensitive data, such as health data, biomedical data, data about children, and air travel data.*
Value 23: *Develop teaching content that looks at the broader issues around managing and storing personal data.*
Value 24: *Develop teaching content that explores document management, looking particularly at storage, access and auditing.*
Value 29: *Develop content that teaches students about the range of ways data is collected both online (transactions, cookies, social media) and offline (Internet of Things, Digital Assistants), and ways of handling sensitive data.*
Value 39: *Develop content around the potential issues of datasets, including the amount of data being stored, the representativeness of the data, the security of the data, and the dangers of partial data and of biased data.*

Value 42: *Develop content that prompts learners to ask relevant questions at each state of the data management life cycle.*
Value 45: *Develop content that encourages students to understand how data is used by applications and how to make that use and the value of their data transparent to end users.*
Value 49: *Develop teaching content to highlight the importance of transparent technologies that explains to users how they are collecting, aggregating and interpreting their data. The content should also highlight the importance of policies and regulations, particularly those that protect the right of individuals to retain control over their data, and other general control mechanisms that should be made explicit to users.*
AI Ethics
Value 10: *Develop teaching content that highlights both the potential benefits and potential harms of AI. To do this look at specific topics such as AI safety, intended and unintended consequences of systems, fairness, accountability and transparency of AI systems, AI bias, responsible AI and regulatory issues.*
Value 11: *Develop teaching content that discusses ethical issues that arise from the use of AI, for example, self-driving vehicles, and automated military weapons.*
Value 14: *Develop teaching content that will explore the notion of designing a malevolent AI to understand the implications of such a creation.*
Value 26: *Develop content that teaches students about bias in data and bias in decision-making systems, and discusses statistical techniques in order to facilitate the identification and mitigation of those biases, including Differential Privacy.*
The relationship between Digital Ethics, Digital Regulations and Digital Governance
Value 12: *Develop teaching content that discusses existing legal frameworks that apply to AI (including Article 12 of United Nations Convention on the Use of Electronic Communications in International Contracts), and the links of these legal frameworks with Digital Ethics.*
Value 22: *Develop teaching content that explains the overall purpose of GDPR, as well as highlighting some of the key Articles.*
Value 30: *Develop content that explains GDPR and its implications. Also, look at other legal aspects of data management, including confidentiality agreements, and give the students a general appreciation of how the law works.*

Value 31: *Develop content that encourages students to explore the differences between the meaning of "legal" and "ethical".*
Value 41: *Develop content that teaches people the importance of "Terms & Conditions" and develop content that helps them understand what the terms mean.*
Ethics for Pervasive Computing, Privacy and Surveillance
Value 13: *Develop teaching content that will explain the dangers of concepts such as "behavioural economics", "persuasive computing", "Big Nudging" and "Hyper Nudging".*
Value 15: *Develop teaching content that will explore pervasive technologies from the points-of-view of privacy, equal access, diversity, informed consent and normative ethics.*
Value 16: *Develop teaching content that will explore pervasive technologies and the Internet of Things from the perspective of them as being invisible technologies, which need to be transparent and explain to users how they are collecting, aggregating and interpreting their data.*
Value 17: *Develop teaching content that will explore pervasive technologies from the perspective of policies and regulation, and the rights of individuals to retain control over their data, and understand the control mechanisms that govern these processes.*
Value 50: *Develop teaching content to explain and discuss the widespread use of surveillance technologies, including surveillance Smart Speakers and cameras equipped with facial recognition.*
Social Media Ethics
Value 18: *Develop teaching content that will use social media case studies to explore a range of key digital ethical themes, such as data, informed consent, targeted advertising, AI and privacy.*
Value 19: *Develop teaching content that will explore the relationship between social media and the behaviour of a society.*
Value 20: *Develop tools and materials to highlight the importance of consequence scanning.*
Value 21: *Develop teaching content that will discuss how trolls and bots are being used by malevolent actors to manipulate social media systems. Also look at the impact that cyberharassment has on free speech and advocacy, particularly considering issues around gender and online abuse.*

Value 43: *Develop case studies based around the general area of social media, looking at things such as how many individual technologies, designed and developed for disparate purposes can be repurposed and combined to develop different products and services that tap into unmet customer needs. Social media case studies offer the opportunity to show how multiple pertinent Digital Ethics issues, e.g. data, informed consent, AI and privacy need to be considered in singular technology applications.*
Value 44: *Develop tools and materials for consequence scanning for social media and related applications. Such applications have shown the influence of technology on the behaviour of a society, not simply that of individuals. This highlights the importance of consequence scanning, considering the potential consequences – intended and unintended of new technologies.*
Value 46: *Develop social media case studies looking at multiple pertinent Digital Ethics issues, e.g. information, misinformation and 'fake news', and digital literacy surveillance, the influence of social media on teenagers and young people.*
Value 47: *Develop social media case studies looking at the influence of social media on teenagers and young people (particularly mental health), and the exclusion of older people.*
Professional Digital Ethics
Value 25: *Develop content that explains the costs of development, the costs of adding in ethics checks, and the potential costs of ethical violations being subsequently discovered after the product has been completed.*
Value 27: *Develop content that teaches students to put people (users) at the centre of all developments, and to learn about empathy, and to learn about informed consent.*
Value 28: *Develop content that teaches students how to explain both technical terminology and to explain the functionality of models in clear and understandable language, particularly in areas such as Artificial Intelligence. Emphasize the important role of communication with stakeholders (people who are affected by the technology).*
Value 32: *Develop content that highlights the environmental impact of technology.*
Value 33: *Develop teaching content to explore the dangers of technological colonization, and the subtle and obvious pressures of funding in third-world countries.*
Value 35: *Develop content that teaches Software Methodologies that incorporate Consequence Scanning.*

Value 36: *Develop content that teaches students diversity, and the need for it in datasets, in interface design, and in software teams.*
Value 37: *Develop ethical scenarios to prepare students for potential workplace dilemmas. Also develop scenarios that place similar ethical dilemmas in different contexts to help the students explore if their ethical perspectives are consistent.*
Value 38: *Develop content to encourage students to be sufficiently confident in themselves to question the ethics of their workplace.*
Value 40: *Develop a series of generic ethical frameworks that cover categories of ethical dilemmas that will provide future-proofing for new and upcoming technologies.*
Value 48: *Develop teaching content to explore diversity as a catalyst to better technology design.*

Appendix B. Demographic Data of the Participants to the Focus Groups

1. Industry participants

Industry specialists were interviewed in TU Dublin, Aungier Street, Dublin, Ireland, on Thursday 21st November 2019. These specialists were recruited based on their expressed interest in the topic, and proximity to the location, and work in a range of IT organizations (both small and large). The industry specialists consisted of 10 participants. The majority of the participants (80%) were aged 30-49, and 20% were aged 50-69.

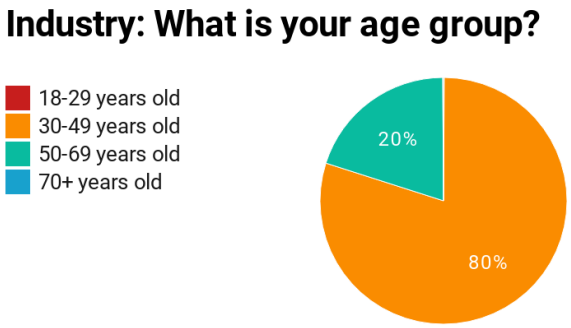


Figure 1: Age of participants

Regarding gender, 50% of the participants were female and 40% were male with 10% preferring not to say. This is not representative of the ICT profession as a whole where females represent only 17.2% of all ICT specialists employed in the EU (Eurostat, 2016). However, we believe the sample is enriched by its increased diversity and gender balance.

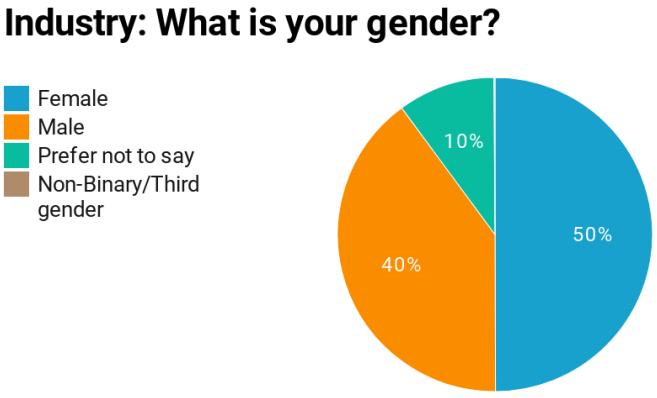


Figure 2: Gender of participants

The highest education levels of the participants is shown in Figure 3 below. 30% had a Bachelor's degree and 60% had a Master's degree, and 10% had a PhD.

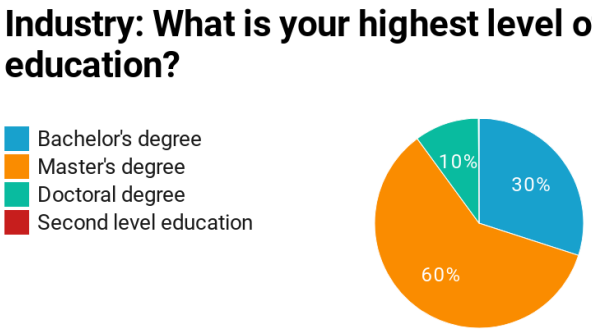


Figure 3: Highest education level of participants

2. Academic participants

Academic specialists were interviewed in TU Dublin, Aungier Street, Dublin, Ireland, on Thursday 21st November 2019. These specialists were recruited based on their expressed interest in the topic, and proximity to the location, and work in a range of local academic institutes (both small and large). The academic specialists consisted of two groups of participants (12 participants in Group A and 11 in Group B). The majority of the participants (65%) were aged 30-49, 22% were aged 18-29, and 13% were aged 50-69.

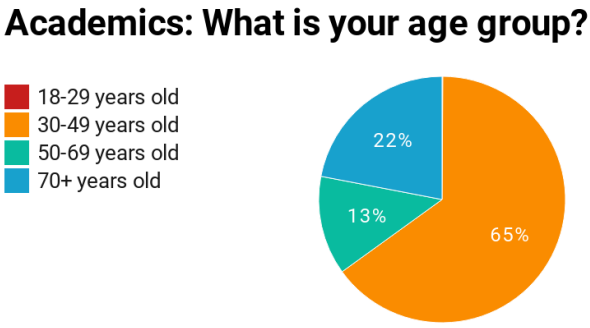


Figure 1: Age of participants

Regarding gender, 48% of the participants were female and 43% were male, with 9% preferring not to say. As with the industry participants, this high proportion of female academics is not representative of academic Computer Science, however, we note that our sample is inclusive and gender balanced.

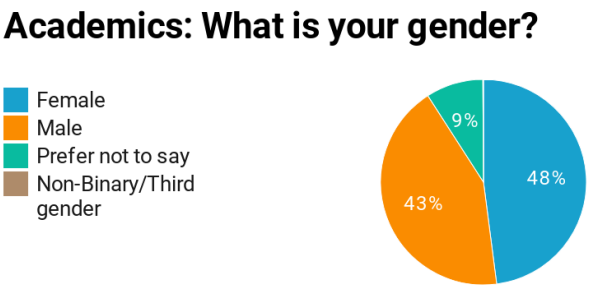


Figure 2: Gender of participants

The highest education levels of the participants is shown in Figure 3 below. 9% had a Bachelor's degree, 39% had a Master's degree, and 52% had a PhD.

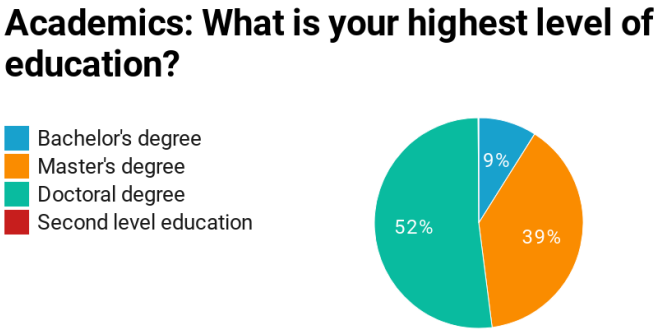


Figure 3: Highest education level of participants

3. Citizen participants

Citizens were interviewed in TU Dublin, Aungier Street, Dublin, Ireland, on Tuesday 3rd December and Wednesday 11th December 2019. Citizen participants were recruited following a snowball sampling method (Morgan, 2008). Some participants were initially contacted by the researchers and were asked to spread the word to other citizens who were interested in participating in the focus groups. The citizens were not experts in Digital Ethics and the data presented below represents their opinions on Digital Ethics topics. The citizens consisted of two groups of participants (12 participants in Group A and 10 in Group B). The majority of the citizens (82%) were aged 30-49, 14% were aged 18-29 and 4% were aged between 50-69.

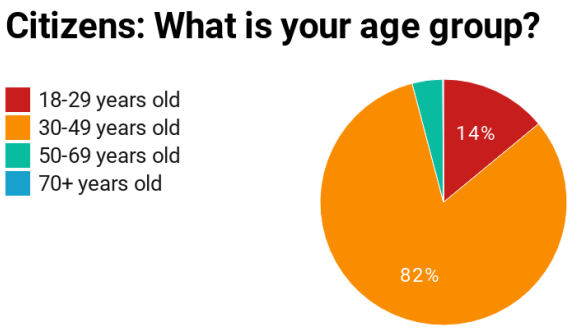


Figure 1: Age of participants

Regarding gender, 64% of the participants were female and 32 were male with 4% preferring not to say.

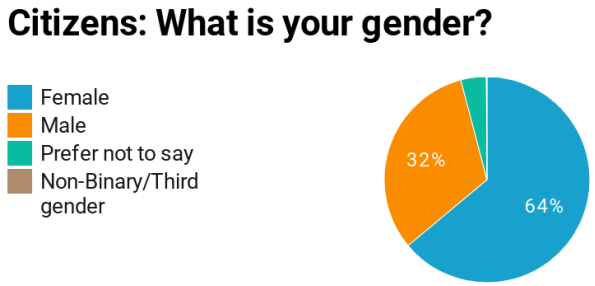


Figure 2: Gender of participants

The highest education levels of the participants is shown in Figure 3 below. 18% had second level education, 50% had a Bachelor's degree and 32% had a Master's degree.

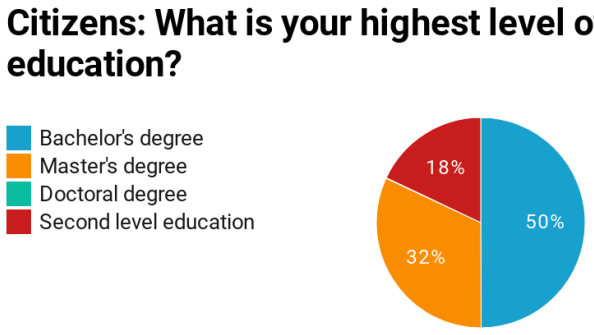


Figure 3: Highest education level of participants

The professions of the participants are shown in Table 1 below. If there was more than one participant in a given profession, the number is shown in brackets after the profession. Seventeen different professions are represented among the participants.

Managers	**Healthcare**
• Category Management • IT Manager (2) • Marketing Manager • Operations Manager • Project Manager (2)	• Dentist • Carer • General Practitioner (2) • Naturopath
Education	**Others**
• Primary School Teacher (2) • Researcher • Student	• Freelance writer • Quality Engineer • Recruitment Services • Sales (2) • Secretary

Table 1: Professions of participants