Approaches and Models for Teaching Digital Ethics in Information Systems Courses – A Review of the Literature

Minna Paltiel
Melbourne Law School, The University of Melbourne, Australia
mpaltiel@student.unimelb.edu.au

Marc Cheong
School of Computing and Information Systems, The University of Melbourne, Australia

Simon Coghlan
School of Computing and Information Systems, Centre for AI and Digital Ethics (CAIDE), The University of Melbourne, Australia

Reeva Lederman
School of Computing and Information Systems, The University of Melbourne, Australia

Abstract

The value and importance of teaching Digital Ethics within Information Systems and ICT courses is widely recognized and stand-alone or integrated digital ethics units are broadly implemented across degree courses. However, how such courses should be taught and what content they should include is a little explored topic. Using a narrative literature review methodology, this paper reviews the pedagogical theories underpinning digital ethics courses discussed in the literature and outlines approaches that deploy standalone ethics units, integrated ethics teaching, and hybrid teaching approaches and the use of interdisciplinary models. The paper identifies the employment of, and emphasis given to various moral theories in digital ethics education. The paper then discusses how our findings relating to different pedagogical approaches, degree of integration of ethics teaching, the use of interdisciplinary models and use of moral theories—are related to each other. The discussion explores trends in approaches and models for teaching digital ethics highlighted in the review, and makes recommendations for further exploration and inquiry. It concludes that the effective teaching of digital ethics will likely involve a considered combination of approaches, models and techniques, which may also be tailored to the needs of different roles and industries.

Keywords: information systems, digital ethics, education, pedagogical theories, moral theories.

1 Introduction

Digital ethics examines moral questions raised by information and communication technologies (ICT), from laptops and smartphones to social media and modern artificial intelligence (AI) (Véliz, 2021). There is a growing appreciation that digital ethics education is required to address the pervasive impact of ICT in society. This view is reflected in current requirements for accreditation in engineering generally (Bradley, 2008; Martin et al., 2021) and computing and information system sciences specifically. The rising need for digital ethics education raises the question of how teachers currently teach digital ethics in tertiary settings. In this paper, we address this question by means of a narrative review and discussion of the literature on digital ethics teaching programs.

Although relatively new, digital ethics is a growing discipline that encompasses computer ethics, information ethics (Bynum, 2015), and AI ethics (Müller, 2020). It addresses issues such as data security and privacy (which are core IS concerns), algorithmic fairness and transparency, superintelligence, social media ethics, AI explainability, data law, the power of tech giants, surveillance, moral machines, and much more (Floridi et al., 2019; Véliz, 2021).

Recognition of the importance of ethics education in computing sciences (CS) and Information Systems (IS) goes back some decades, see e.g., the works of Johnson (1985), Moor (1985), and Mumford (1995). For example, ethics teaching requirements were included in the ACM/IEEE-CS *Computing Curriculum 1991* (Tucker, 1991) and have been present in the Computer Science Accreditation Board curriculum standards since 1987 (Califf & Goodwin, 2005; Huff & Martin, 1995).

Huff and Martin (1995) claimed that ICT students who plan to enter industry should understand the ethical and social consequences of their work. Growth in computerisation, AI, and automated decision making in business, services, and society generally (O'Neil, 2016) has only increased the significance of their observation. Such technology can carry substantial risks and harms. Indeed, the overarching impact of computer technology on all aspects of personal, social and public life has made computer scientists "some of the most powerful moral agents in today's world" (Skirpan et al., 2018, p. 940). Since these ubiquitous algorithms are created by humans and based on human data, they tend to reflect and embody imperfect, biased, and unfair circumstances and decision-making (Borenstein & Howard, 2021; see also Eubanks, 2018).

The need for ethics education goes beyond algorithm designers to include IS practitioners and academics. For example, IS specialists must often implement, oversee, and teach others about such algorithms. Many of the large ecosystems powering modern leisure and work – from social networking sites to cloud-based systems enabling us to work and connect during the ongoing Covid-19 pandemic – are information systems at their very core. In fact, Mason's (1986) "prominent contribution to IS ethics" in the 1980's raised issues still very much present in contemporary IS discourse: "privacy, accuracy, property, accessibility" (Hassan et al., 2018), or "PAPA". Unfortunately, as Kordzadeh and Ghasemaghaei (2021) write in a reflection upon Someh et al.'s (2019) work on algorithmic bias, "Information Systems (IS) researchers have largely fallen behind in addressing the behavioural, organisational, and social implications, antecedents, and consequences of this issue" when compared to their computational science counterparts (Kordzadeh & Ghasemaghae, 2021, p. 388).

There is thus good reason for highlighting the need for IS professionals to be better trained to identify ethical issues, reflect upon their own biases, and provide ethically sound solutions to contemporary techno-social problems. Unsurprisingly, recent scholars have again called upon ICT educators to train future professionals (and other community stakeholders) not only in the requisite technical proficiencies, but also in ethics, so that they can be better prepared for the ethical responsibilities inherent in practice (Borenstein & Howard, 2021; Grosz et al., 2019; Skirpan et al., 2018).

Many of those who currently teach ethics in IS and related disciplines, however, lack experience and training in philosophical theory and digital ethics pedagogy. Furthermore, programs and interventions for teaching ethics vary greatly in structure, methods, and pedagogical approach and in the educational tools they employ. This review aims to assist digital ethics educators in understanding the teaching options available.

Our narrative literature review focuses especially on more recent scholarly work concerning the teaching of ethics in ICT-based university courses. It uncovers several themes and trends regarding current recommendations for digital ethics teaching in the present digital age. One theme relates to the different pedagogical approaches adopted. While a very broad range of approaches was reflected in the literature, we focus on the several key pedagogical theories. As we shall see, some of the literature recommends that ethics be taught in a way that is inextricable from the typical activities of CS and IS (this combination hereinafter abbreviated CIS) practitioners. As such, integrating ethics into the basic curriculum, rather than only presenting stand-alone ethics courses, is sometimes seen as desirable. This emphasis is also seen in adopting pedagogical approaches that teach ethics through direct experience, *in situ*, and within the CIS context.

Another trend in the literature which we identify concerns the need to merge different ways of thinking in teaching, such as including social and learning sciences approaches in designing digital ethics courses and creating multidisciplinary teaching partnerships. There is also a theme about building ethical skills and virtues rather than acquiring moral knowledge sets. Finally, some papers we reviewed described different learning tools such as case studies and games.

The rest of the paper runs as follows. Section 2 describes our methodology, Section 3 presents our findings, Section 4 discusses the findings and offers recommendations for digital ethics educators and direction for further inquiry, and Section 5 concludes our analysis of digital ethics teaching.

2 Search methodology

To obtain a broad overview of the relevant literature, we conducted a narrative review, ‘a scholarly summary along with interpretation and critique’ (Greenhalgh et al., 2018, pp. 2-3) to identify papers about digital ethics teaching for students in IS and related fields. Despite a preference for systematic reviews in IS and the broader research community (such as medicine), we adopt a narrative review for two key reasons. First, systematic reviews presuppose a "specific" research question[1] , whereas a narrative review style is "better suited to addressing a topic in wider ways" (Baethge et al., 2019, p. 2). Second, a broad narrative review allows us to learn from counterparts in the various other CIS subfields and disciplines (Kordzadeh & Ghasemaghaei, 2021; Someh et al., 2019). Principles such as PAPA (Privacy, Accuracy, Property, Accessibility) (Mason, 1986), and other ethical frameworks are not exclusive to IS, but are shared across allied ICT disciplines, including AI, data science, and engineering.

The articles reviewed were found via a database search using University of Melbourne library resources, and included: *AI and Ethics, AI and Society, International Journal of Artificial Intelligence in Education, Science and Engineering Ethics, Journal of Information, Communication and Ethics in Society, Ethics and Information Technology* and *Communications of the ACM*. Conference proceedings were also accessed, including: *ACM Conference on Fairness, Accountability and Transparency (FAaccT), Conferences of the Institute of Electrical and Electronics Engineers (IEEE)* and conferences of the *Australasian Institute of Computer Ethics*. Additional searches were conducted on Google and Google Scholar.

[1]We thank the anonymous reviewers from ACIS 2022 for their valuable suggestions.

Key terms used to search in titles or abstracts of publications included combinations of *"AI/artificial intelligence"* OR *"computer science/CS"* AND *"ethics"* AND *"teaching"* OR *"education"* OR *"curriculum"* OR *"training"*. Second phase searches addressed particular models such as *"FATE"/"FAccT"* and *"Embedded"*. As a third phase, references in key publications were reviewed to identify additional relevant publications, journals, and conference proceedings pertaining to the subject of the review. Papers selected for review were generally published within the past 10 years, although some earlier papers were included as background and to provide perspective on trends and developments. In all, 35 studies were reviewed, 86% of which were published in the last decade. For completeness, an overview table of the reviewed studies is provided in Table 2, in the Appendix. Many additional papers were extracted from relevant databases, but were excluded for lack of relevance or significant contribution, based on reading the abstract and/or the article.

3 Findings

Five themes and trends concerning recommended dimensions of digital ethics teaching emerged from our literature review. Our findings reveal, first, a broad range of pedagogical approaches were adopted, but with an emphasis amongst scholars on ensuring that ethics is taught as inextricable from the activities of an ICT practitioner. This involves pedagogical approaches that teach ethics through direct experience, *in situ*, and within an ICT context (detailed in section 3.1). A second theme relates to the benefits of integrating ethics into the IS/CS curriculum as opposed to presenting stand-alone ethics courses (see section 3.2). A third theme in the literature concerns the need to merge different ways of thinking, including social and learning sciences, in the design of digital ethics courses, and to work in multidisciplinary and interdisciplinary teaching partnerships (see section 3.3). A fourth theme is a focus on building ethical skills or developing moral values and virtues rather than acquiring a particular moral knowledge set. Here there is considerable reference to the "virtue ethics" school and its focus on developing "ethical character" (see section 3.4). A final theme is the different content emphasis and range of pedagogical tools employed in the CIS ethics courses described in the literature, which reflect the key themes and trends found (see section 3.5).

3.1 Pedagogical approaches

A pedagogical theory is a "theory of educational action, or a systematic view and reflection of pedagogic practice" (Hämäläinen, 2012). In a sense, pedagogical theories operate at a higher level than educational approaches to specific subject matter and are compatible with a range of disciplines. Even so, selection of an appropriate theory can be influenced by the particular subject being taught.

The literature addressing digital ethics education refers to a broad variety and different combinations of pedagogical approaches. This appears to be partly because teaching ethics in CS and IS involves the synthesis of several disciplines, and partly because although the need for ethics education was flagged thirty years ago, researchers are still working on formulating the optimal pedagogical basis for doing this. Despite the wide range of pedagogical approaches referred to, several primary theories stand out, which we present below.

3.1.1 Constructivist/constructionist approaches

A dominant pedagogical theory referred to in the literature is constructivism, which has been called the "dominant theory of learning today" (Ben-Ari, 2001, p. 45). Constructivism as a pedagogical approach stems from the Piagetian theory of cognitive development according to which knowledge is not passively absorbed, but actively constructed by a learner in interacting with her world (Ackermann, 2001; Piaget, 1971). The learner does this through the processes of *assimilation* and *accommodation*, integrating new experience in their environment with knowledge already held, and adapting mental schemas accordingly. While Piaget developed his theory about children's cognitive development, constructivist learning theory as a pedagogy applies to child and adult learners alike. A constructivist approach underpinned several digital ethics programs reviewed in the literature (e.g., Bates et al., 2020; Lewis & Stoyanovich, 2021).

Papert's constructionist theory of learning (Papert, 1993; Ackermann, 2001) informed a number of studies reviewed (Ali et al., 2019; Hjorth, 2019; Lewis & Stoyanovich, 2021; Wise, 2020). Constructionism is based on Piaget's constructivism and holds in common with that approach the principle that learning occurs through active engagement with the environment and internalisation of experience. However, constructionism adds that this occurs particularly when "the learner is engaged in constructing a public entity…"; in other words, it involves "learning through making" (Papert & Harel, 1991, p. 1). According to constructionism, knowledge is formed and transformed in particular contexts (and through particular individuals), and via particular uses and media (Ackermann, 2001.) Constructionism has been described as "both more situated and more pragmatic" than constructivism (Ackerman, 2001, p. 5).

A number of digital ethics programs in the literature, involving learning from direct experience, used constructivism and/or constructionism as guiding principles (Briggle et al., 2016; Holmes et al., 2021; Richards et al., 2020). These programs are variously project-based (Ali et al., 2019; Hildt et al., 2019), interactive (Lewis & Stoyanovich, 2021), *in situ* (Skirpan et al., 2018) and involve design or analysis and include "deliverables" such as "nutritional labels" (Lewis & Stoyanovich, 2021). These closely related approaches encourage "situated" ethical reasoning (Grosz et al., 2019).

Many programs, even when not expressly described as based on constructivist and constructionist principles, illustrate these approaches. For example, Grosz et al.'s "Embedded EthiCS" model employs a "distributed pedagogy" which conveys the message that ethical reasoning is an integral part of computer science. It "situates" ethics learning in the context of CS activity, and aims for students to identify the ethical implications of their technological work and reason clearly about them in the process of developing and designing algorithms, systems, and codes (Grosz et al., 2019).

3.1.2 Social analysis approaches

Huff and Martin (1995) suggested that every ethical concern arising in computer science is located at a particular level of social analysis, arguing that "[o]nly analysis that accounts for at least three dimensions - technical, social, and ethical - can represent the issues as they affect computer science in practice" (Huff & Martin, 1995, p. 76). The authors referred to Project ImpactCS, begun in 1994, teaching students to identify and understand social issues and ethical issues associated with technology. The project uses a "social analysis" teaching approach in which: a technology and social issue was identified, the appropriate level of social analysis (e.g., individuals, communities, groups, organisations, cultures, institutions, etc.) is determined, and the appropriate tools, literature and methods are then applied (Huff & Martin, 1995). The authors identified the importance of developing a corresponding skill set in computer professionals. Referring to Project ImpactCS, Barnard et al. (2003) set out the social analysis skills critical to CIS ethics as: "identification and interpretation of the social contexts of a particular implementation" identification of assumptions and values embedded in a particular system; and evaluation by means of empirical data of a particular implementation of a technology" (Barnard et al., 2003, p. 269).

A more recent example is the CS ethics curriculum presented by Carter and Crocket, which is based on "proven learning, teaching, moral education, and social analysis theory", and integrates the social analysis approach with moral and ethical thinking (Carter & Crockett, 2019, section I). Carter and Crocket's work, and moral reasoning approaches, are discussed further below.

3.1.3 Approaches emphasising microethics versus macroethics

The social analysis approach, and emphasis on large social issues, is congruent with a "macroethics" approach to ethics education. A macroethics approach in engineering and information sciences is concerned with the collective social responsibility in the profession, and societal decisions made about technology. In contrast, a "microethics" approach is concerned with individual responsibility and internal relations within the profession (Herkert, 2004). Li and Fu (2012) hold that the microethics emphasis detracts from the need to deal with the social nature of technology practice and that it is fundamental to teach ethics within the applicable social, organisational, and political contexts.

In contrast, Bezuidenhout and Ratti argue for a microethics model aimed at connecting "big picture ethics" to "everyday practice" (Bezuidenhout & Ratti, 2021,p. 940). They seek to foster "moral virtue" in learners by focusing on individual decision making and action, rather than on key themes and high-level case studies found in macroethics discourse. The aim is to develop moral excellence through virtue training, which is in contrast both to the approach described above focusing on identification of social issues, and the focus on ethical skills in the neo-Kohlbergian approaches discussed below. Bezuidenhout and Ratti's virtue training is accomplished through "discrete and repetitive" practice and by packaging "daily events" such as coding, clicking on content, and engaging in chat forums into "discrete instances of ethical reflection" (Bezuidenhout & Ratti, 2021, p. 947).

3.1.4 Cognitivist or moral reasoning approaches

Training in cognitive moral reasoning is referred to in many of the courses we reviewed (Mumford et al., 2008; Richards et al., 2020; Sprague & Diaz-Sprague, 2019). These are often based on ideas of moral learning advanced by psychologist Lawrence Kohlberg (1984) or on the Neo-Kohlbergian idea that skills of moral focus, sensitivity, and action may be strengthened through practice (Mumford et al., 2008; Rest et al., 1999; Richards et al., 2020).

An illustrative example is Mumford et al.'s (2008) "sensemaking" approach which uses a moral reasoning model based on Kohlberg and Rest (Kohlberg, 1984; Mumford et al., 2008; Rest, 1986). "Sensemaking" refers to a complex cognitive process activated when people are presented with "ambiguous high-stake events" (Mumford et al., 2008, p. 316). Various mental models are applied to understand the situation as the foundation of decision making. According to this approach, ethical decision making is based on the available case-based models applied to the situation by the decision maker (Mumford et al., 2008).

This moral reasoning model is combined with a "field practices approach", focusing instruction on codes of conduct and guidelines specific to a particular field, and a "case analysis approach", in which analysis of prior cases provides a framework to "make sense" of the ethical dilemma (Mumford et al., 2008). The cognitivist and moral reasoning approaches often employ case studies as a pedagogical tool, as discussed further below.

Another interesting example is Richards et al.'s "AI enhanced" game for training in cybersecurity ethics, which is likewise based on the neo-Kohlbergian view that ethical expertise may be assessed in terms of the skills of moral focus, sensitivity and action, and that these skills may be strengthened through practice (Richards et al., 2020; see also Rest et al., 1999). The game itself, a training simulator, is guided by the "theory of situated and experiential learning" developed by articulated by James Paul Gee (Gee, 2007, cited in Richards et al., 2020), which requires "contextualised exploration, discovery, and practice" (Richards et al., 2020, p. 6).

3.1.5 Reflexive and deliberative approaches

The final pedagogical theory emerging from our review are reflexive and deliberative approaches to digital ethics teaching. The reflexive approach involves critically questioning one's own assumptions, judgements, and practice. Bezuidenhout and Ratti (2021) sought to develop "critical reflexivity" in CS students, while Barabas et al. (2020) propose a "studying up" model, borrowed from the field of anthropology, aimed at achieving more reflexive data science practices. Barabas et al.'s approach focuses on critically reflecting on dominant modes of interpretation data which reinforce factors leading to negative social outcomes. The authors argued that algorithmic fairness must be examined in the light of institutional context, oppression and control, with dominant modes of data interpretation reinforcing hierarchies and biasing the outcomes (Barabas et al., 2020). Barabas et al. emphasised the role of data scientists as agents in this process of developing frameworks and structures for evaluating algorithmic fairness, and suggested that computer scientists must embrace more reflexive practice (Barabas et al., 2020).

The literature also includes recommendations for a "deliberative" approach in CIS ethics courses. This approach enables consideration of many points of view and involves discussion and interaction with others. It emphasises the development of an ethical *culture* rather than a strong emphasis on individual morality. Plemmons et al. (2020) reported on a STEM research ethics training intervention which sought to improve participants' "reason giving" and "interpersonal communication" abilities for more ethical practice. The course was based on the proposition that if laboratory members routinely speak to each other about ethical issues arising in research practice, they are more likely to consciously select practices and procedures that are ethically defensible. This intervention aimed at fostering a more ethically sound culture in STEM laboratories (Plemmons et al., 2020).

The deliberative approach is also illustrated in several ethics courses for neurotechnology students. Farooqui et al. (2021) proposed an ethics course for neurotechnologists and neuroengineers founded on dialogue and involving monthly discussions. The course melded the reflexive and deliberative approaches in emphasising dialogue and exploration while also focusing on the relationship between the ethical principles discussed and students' own values.

Tubig and McCusker also combined ethical reflexivity and discourse in their ethics dialogue tool – Scientific Perspectives and Ethics Commitments Survey (SPECS) (Tubig & McCusker, 2021). This involved reflection through discourse in which "discussants articulate and critically examine the values and beliefs that may be governing their group activities and the discretion of its members" (Tubig & McCusker, 2021, p. 150). SPECS has a twofold premise: first, it draws on individual researchers' interest in engaging with the ethical aspects of their work, and second, "it appreciates the discursive nature of ethics and the transformative potential of discourse" (Tubig & McCusker, 2021, p.152).

3.2 Standalone ethics units versus integration with technical content

The second main theme to emerge from our narrative review concerns the contrasting strategies of presenting ethics material as "stand-alone" units versus integration of digital ethics teaching *within* the teaching of technical content. More recent literature advocates for a more integrated or embedded model (or in some cases, a hybrid approach) as opposed to standalone modules, to ensure that ethical reasoning and skills remain relevant when students and professionals are engaged in technical tasks (e.g., Skirpan, 2018; Grosz, 2019; Hildt et al., 2019; Plemmons et al., 2020; Bogina et al., 2021).

3.2.1 Integration approach

In their analysis of recent approaches to AI ethics education, Borenstein and Howard (2021) emphasised the need for computer scientists to understand ethics as *intertwined* with the technology future practitioners will design, create, or implement, and their responsibility as computer scientist to address ethical issues arising in their work.

Borenstein and Howard (2021) suggest, first, that ethical design of AI algorithms should be part of learning about algorithm design generally, and should include FATE (i.e., "fairness, accountability, transparency and ethics" in computing – also referred to as FEAT) considerations, amongst others. Second, the authors recommend that real world datasets should be used to teach ethics of data acquisition. Third, ethics courses in their view should not be taught as one off or standalone modules but should be repeatedly incorporated in different ways. The authors propose an "ethics across the curriculum" model which would include and reinforce the significance of ethics while teaching technical courses. They also emphasise that the course should be created and taught by an interdisciplinary team including lawyers, sociologists and philosophers together with scientists and engineers (Borenstein & Howard, 2021).

Additional examples of this integration model of teaching include a "distributed pedagogy" that incorporates short ethics modules interspersed throughout the core curriculum (Grosz et al., 2019), an *in situ* or contextual embedded ethics course model (Carter and Crockett, 2019; based on Huff & Martin 1995), and research laboratory-based interventions (Hildt et al., 2019; Plemmons et al., 2020).

3.2.2 Stand-alone approach

However, despite this emphasis on embedding ethics teaching in other subjects, some authors recognise benefits in stand-alone ethics units. One critique of "embedding" digital ethics in technical courses, especially in the lab-based approach (Hildt et al. 2019; Plemmons et al., 2020), is that it can be too localised. By electing a bottom-up approach linked to practice within the lab, teachers might omit to effectively teach a broader social perspective, one that involves critical attention to the socio-ethical implications of technology. On this view, the integrated approach fails to address more holistic ethical issues relevant to IS students.

3.2.3 Hybrid approaches

Some authors preferred neither a pure integrated nor a pure standalone model but rather a *hybrid* model of ethics teaching. In the hybrid approach, ethics teaching is spread "across the curriculum" (Borenstein & Howard, 2021), but for certain reasons (including logistics around course structures) are also deployed in standalone modules. Bogina et al. propose projects such as creating a "library of hands-on exercises", integrating FATE training into existing courses, and creating courses and seminars suitable for various stakeholders (including students, faculty and the public) with varying skill levels (Bogina et al., 2021, p. 23).

Skirpan et al. (2018) discuss a short, intensive "Human Centred Computing" course that is included as part of the overall computing training. However, they also recommend that ethics teaching should be done in "small doses" throughout the curriculum. A series of similarly relevant paradigms and projects could theoretically be adopted to address the other concepts covered in the modules, and to "embed" the RDS training within the technical instruction.

Bates et al. (2020), who focus on integrating FATE topics and social justice into data science curriculum, describe using stand-alone FATE modules woven through a data science course and interspersed amongst both core and elective teaching modules. The course included a "Data and Society" module as a core unit, covering themes ranging from "conceptualisations of power, structure, and agency … [to] …Ethical Reasoning" (Bates et al., 2020).

3.3 Interdisciplinary models

The interdisciplinary aspect of teaching computer ethics recommended by Borenstein and Howard (2021) is emphasised in a number of studies promoting the integration approach (see e.g., Bates et al., 2020; Grosz et al., 2019; Skirpan et al., 2018). Skirpan et al.'s integrated program incorporates guest lecturers such as a privacy lawyer, a typography and layout artist, a researcher of terms of service agreements and online harassment, and an emeritus professor who reviewed developments in technology over the past fifty years (Skirpan et al., 2018).

Bates et al., reported on the findings of a collaborative, auto-ethnography engaged in by a multidisciplinary team involved in embedding FATE/Critical Data Studies in a Data Science Masters program (Bates et al., 2020). The authors saw "data science" as encompassing an "information science" approach rather than a strictly "computer science" approach, indicating not only technical expertise, but also social, legal, and ethical understanding. Their view is that teaching CIS ethics must by its nature involve more than one discipline.

The interdisciplinary aspect of Grosz et al.'s Embedded EthiCS program was reported as a significant factor contributing to its success (Grosz et al., 2019). Here the Harvard CS and philosophy departments collaborated in delivering the Embedded EthiCS modules. The modules were codesigned by faculty from both fields and delivered by teaching assistants from the philosophy department with expertise in ethics who worked together with CS teaching staff. The authors found that the process of co-designing the ethics modules helped mitigate insecurities felt by both these faculty in teaching across disciplines.

3.4 Moral theory selection and emphasis

The fourth theme relates to which *moral theories* are taught or emphasised in digital ethics. Standard teaching practice in professional ethics, from medicine to business ethics, involves teaching of moral theories (Beauchamp & Childress, 1994; Moriarty, 2021). Such theories may include consequentialism and utilitarianism, deontology and Kantian ethics, ethics of care, and virtue ethics. Other theoretical approaches are possible, including but not limited to Christian, Confucian, Jewish and Buddhist ethics.

Amongst the digital ethics models we reviewed, some taught a "crash course" in philosophical ethics, while others referenced a theory or several theories of ethics as their framework. Many scholars recommended that to deal with the complexities of ethical decisions in the digital age, students should be trained in many modes of ethical reasoning (Goldsmith & Burton, 2017). The three main schools of ethical reasoning referred to in the literature were utilitarianism, deontology, and virtue ethics. Roughly speaking, utilitarianism is based on the principle of producing the "the greatest good for the greatest number" while deontology is a "duty-based ethics" according to which actions are ethical if they conform to a system of rules. One version of deontology, Kantian ethics, includes the famous claim that rational beings ought always to be treated as ends and never merely as means. "Virtue ethics" focuses on individual character and the development of good character traits or dispositions to think and feel in desirable ways.

For Goldsmith and Burton (2017), the dominant western ethical framework is utilitarianism, which (they claim) most AI practitioners adopt. Nonetheless, in their view virtue ethics and deontology have much to offer AI decision making. The authors suggest that all three forms of ethical reasoning should be taught in ethics in AI courses.

Jones argues that ethical issues in CS are both complex and context-specific and that it is not possible to focus on one theory, such as deontology, without regard for another values, such as the consequences of actions (Jones, 2016). Chatila and Havens, in the IEEE's Ethically Aligned Design (Chatila & Havens, 2019), also recommend that several ethical traditions be integrated into both engineering and science programs. Carter and Crockett's (2019) proposed course, described above, also aligns with the recommendation that digital ethics courses should include modules dedicated to virtue ethics, utilitarianism, and deontological ethics (involving moral codes).

Bezuidenhout and Ratti's (2021) "micro-virtue ethics" approach specifically reflects a virtue ethics framework. Bezuidenhout and Ratti (2021) refer to the current focus on ethical outcomes and algorithmic design, describing utilitarian approaches that emphasise social impact, and deontological approaches focusing on principles and guidelines. Yet the authors' prefer an approach grounded in virtue theory which aims to develop "moral virtue" through practice and perfecting intellectual and moral skills, without teaching particular ethical content. Their emphasis is not on general "fairness, accountability and privacy" in data science but on individual responsibility.

Hagendorff (2020) reviewed a series of guidelines in AI ethics and examined the extent to which the ethical principles expressed in these guidelines are implemented in AI research and application. He noted the "deep gap" between digital ethics and "concrete contexts of research, development and application" (Hagendorff, 2020, p. 111). Like Bezuidenhout and Ratti, Hagendorff endorses a microethics approach. He writes that traditional ethics guidelines for CS tend to adopt a deontological approach, setting out rules and maxims; however, he thinks that a virtue ethics approach is called for. This approach does not involve defining codes of conduct but rather involves focusing on an individual's "situation-specific deliberations", "behavioural dispositions", and the cultivation of "moral character" to facilitate ethical decision making in ICT contexts (Hagendorff, 2020, p. 112).

Richards et al.'s (2020) serious game ethics training tool may also be seen as close to a virtue ethics approach with its emphasis on psychological development in moral. Briggle et al. (2016) examine research and engineering ethics games founded on virtue ethics. They note the conclusion of the National Research Council (National Research Council and National Academy of Engineering, 2014) that the most important mechanism for ethical decision-making is "good judgment", and they call for "a return to virtue ethics and its emphasis on practical moral reasoning" (National Research Council and National Academy of Engineering, 2014, in Briggle et al., 2016, p. 241).

3.5 Course content and pedagogical tools

The fifth theme related relates to different content taught in CIS ethics training courses and a variety of pedagogical tools employed, which aligned with the underlying pedagogical approach and choice of ethical theory.

A number of the programs structured themselves around the study of elements of ethical computing as set out in instruments such as the Montreal Declaration (Montréal Declaration for a Responsible Development of Artificial Intelligence, 2018). The Montreal Declaration sets out ten principles of responsible development of AI, including, amongst others, well-being, respect for autonomy, privacy protection, democratic participation, equity, diversity and responsibility. Another example of a framework teaching "elements" or "principles" of ethical computing is the ImpactCS project (see Barnard et al., 2003) which sets out five "knowledge units" including: Responsibility of the computer professional (including a review of major ethical models); basic elements of ethical analysis; basic skills of ethical analysis; basic elements of social analysis; and basic skills of social analysis. Other programs reviewed were structured on the teaching of the FEAT principles of "fairness, ethics, accountability and transparency" (e.g. Bates et al., 2020; Bogina et al., 2021; Lewis & Stoyanovich, 2021). Other models, aimed at pre-tertiary teaching, adopted variations of these principles. An example is Greengard's proposed program for kindergarten to high school AI ethics instruction, which refers to introducing notions of "fairness, transparency/explainability, trustworthiness and accountability" (Greengard, 2020). The focus in this course content is on the acquisition of knowledge and understanding of ethical issues, or, when approached more deeply, the imparting of societal and moral values understood to guide ethical CIS practice. In line with a social analysis and macroethics approach, the content of these programs includes formulated principles relating to "large societal issues".

In contrast, the programs focusing on development of moral "traits", adopting a virtue ethics emphasis and a microethics approach, did not focus on teaching a prepared framework of ethical principles, but on practicing ethical reasoning and reflection (Bezuidenhout & Ratti, 2021; Grosz et al., 2019). Ethical issues in these courses were not taught as principles, but as challenges to be resolved. Accordingly, the content did not comprise subject matter to be relayed through instruction, but activities and exercises undertaken in the course of CIS activities, to foster ethical skills (Burton et al., 2018) and develop practical ethical competence (Grosz et al., 2019). These courses used a range of pedagogical tools to exercise students' moral reasoning, ethical reflection and ethics communication skills. We describe several examples below, highlighting the pedagogical and ethical approach they reflect.

3.5.1 Case studies

Many of the programs rely upon use of case studies which has been a commonly used pedagogical tool for ethics training for decades (see e.g. Carter & Crockett, 2019; Hildt et al., 2019; Mumford et al., 2008; Newberry, 2004; Towell, Thompson & McFadden 2004). Some programs relied primarily on readings and reflective discussions (Farooqui et al., 2021). However, case studies can be used as a trigger for role play, simulations, and interactive activities (e.g. Hildt et al., 2019; Towell, Thompson & McFadden, 2004). These techniques are associated with cognitivist theories of moral psychology such as Kohlbergian and Neo-Kohlbergian schools described above. Alternatively, they are used in conjunction with other educational tools and activities, including *in situ* activities such as designing and analysing a predictive algorithm (e.g., Carter & Crocket, 2019). Most courses used dialogue and discussion to carry through study of ethics intervention to other CS units and to ensure ethical issues were intertwined with technical ones.

A variation on the use of case studies and simulations is Burton et al.'s (2018) use of science fiction as a tool for teaching ethics in AI. Burton et al. used science fiction to allow "reframing" familiar situations and problems in unknown settings which enables leaners to recognise and approach ethical issues unencumbered by their own biases. The authors suggested that this was preferable to the use of case studies in which the dilemmas are already identified and characterised. Burton et al.'s (2018) program included a "crash course" in the three main ethical theories: deontology, virtue ethics and utilitarianism. The students were then presented with select science fiction stories and given assignments requiring them to work descriptively with the three theories.

3.5.2 *In situ* learning and using "deliverables"

Bezuidenhout and Ratti (2021) criticise the reliance on "high level" case studies for not sufficiently connecting "big picture" ethics to individual everyday practice. Instead, they propose a range of 15 minute "ethics exercises" linked to computing modules in the course. The exercises involved mind mapping, line voting, and use of sticky notes. The aim of the exercises was to create "critical reflexivity" in individuals' daily CS practice and to enlarge the "moral imagination".

Lewis and Stoyanovich (2021) sought concrete educational materials to teach CS ethics. They use "nutritional labels" (as an expression of the constructivist paradigm of "object-to-interpret-with") to assess the interpretability of AI algorithms. A nutritional label is a visual artifact, combining text and graphics "to communicate highly technical and opaque information". Nutritional levels have been shown to increase understanding of complex topics and aid in decision-making. Lewis and Stoyanovich's (2021) program required students to design nutritional labels for Automated Decisions Systems and to communicate important information about a machine learning model and hidden algorithms, in this way promoting interpretability and transparency.

3.5.3 Gaming

An interesting adaptation of simulation techniques in CIS ethics is the use of "serious games" such as Richards et al.'s (2020) serious video game for teaching ethical conflicts in cybersecurity. Rather than employing pre-scripted scenarios with predetermined solutions to ethical problems and multiple-choice responses, the game represented "complex social simulations" in which ethical problems arise and introduced "intelligent virtual agents" which can "reason about the world". The program used knowledge acquisition and representation techniques to distil the values underlying learners' ethical choices within the game and the reasoning behind them (Richards et al., 2020, p. 9).

Briggle et al. (2016) reviewed the use of games for teaching research ethics in STEM fields. Founding their evaluation on a virtue ethics approach, the authors concluded that the "immersive and interactive experiences" provided by the games were conducive to practicing and developing skills and virtues (Briggle et al., 2016, 248).

A summary of our findings is presented in Table 1. (Please refer to the Appendix for further detail on the features of each paper).

Theme	Findings
Pedagogical theories (3.1)	Constructivism/Constructionism seems to be a dominant pedagogical theory (Ben-Ari, 2001). Social analysis contrasts with concentration on specific issues within a research lab; macroethical contrasts with microethical approaches. Cognitivist/moral reasoning uses case studies to 'put students in the shoes of a moral decision maker.' Reflexive approaches promote interdisciplinary thinking (Ackerman,2001).
Standalone vs integration (3.2)	Integrating (or embedding) of ethics seem to be the preferred approach, i.e. a distributed pedagogy (Grosz et al., 2019) and 'ethics across the curriculum' model (Borenstein & Howard, 2021). FEAT studies which involve critical data studies often are useful in a self-contained (standalone) module. However, the hybrid approach may fare better e.g., in capstone projects.
Interdisciplinary models (3.3)	Engaging expertise of a variety disciplines was recommended (Skirpan et al., 2018; Borenstein & Howard, 2021). Cross-disciplinary co-design and co-teaching of CIS ethics modules was found to be most effective (Grosz et al., 2019).
Moral theory selection and emphasis (3.4)	Moral theory selection: Typical moral theories taught were utilitarianism, deontology, and virtue ethics. A trend exists towards preferring virtue ethics and development of moral character in students.
Content Emphasis and Pedagogical Tools (3.5)	Content emphasis – knowledge based and skills based: Some courses were structured on knowledge of principles of ethical IS practice such as FATE, as frameworks for ethical analysis (Bates et al., 2020; Bogina et al., 2021). Courses based on virtue ethics focused on training skills of ethical reasoning, discourse and communication, and discrete repetitive practice (Bezuidenhout & Ratti 2021). Pedagogical tools: Use of case studies to practice moral reasoning (Carter & Crockett, 2019; Hildt, 2019; Mumford et al., 2008; Newberry, 2004; Towell, Thompson & McFadden, 2004), *in situ* learning and deliverables (Lewis and Stoyanovich, 2021), gaming tools (Briggle et al., 2016; Richards et al., 2020).

Table 1. Summary findings on ICT ethics education – design criteria and observed trends

4 Discussion

In this section we discuss some implications of our findings including the diversity of pedagogical approaches, stand alone and integrated teaching models, the trend towards interdisciplinary involvement, and a range of moral theories and pedagogical tools (4.1). We then provide recommendations for further inquiry (4.2).

4.1 Implications of findings

Our review of the literature on teaching ethics in CIS courses and related fields first identified the pedagogical approaches of constructivism/constructionism, social analysis, microethics/macroethics, cognitivism or moral reasoning, and reflective/deliberative approaches. Second, we found that digital ethics was sometimes taught as a standalone unit, sometimes integrated into technical subjects, and sometimes taught as a hybrid of both. Third, we uncovered a trend towards interdisciplinary involvement both in the design of CIS ethic courses, and in teaching. Fourth, we found that the main moral theories that were taught or used in teaching were utilitarianism, deontology, and virtue ethics. Finally, we observed several different teaching tools, such as ethical case studies, *in situ* learning, and gaming. In this sub-section, we discuss such findings and identify some possible connections between them.

We start by considering the different emphases in the content of digital ethics courses, with some courses focussing on imparting knowledge and understanding ethical and societal issues associated with IS and others focusing more on training IS students to be actual moral decision-makers through repeated practice in the course of their technical work. These different foci were also associated with varied educational tools and techniques. While the pedagogical approaches, learning models and course structures adopted were very diverse, several newer studies emphasised a virtue ethics approach, with the related emphasis on teaching CS and IS students to be ethical people and decision-makers in the field through discrete repetitive, in-situ ethical training.

Furthermore, teaching digital ethics in a more distributed and embedded fashion in ICT curricula, rather than as a single, one-off unit for example, may promote the teaching not just of cognitivist or moral reasoning but also the imparting of long-lived ethical character traits and skills. This trend towards recognizing the importance of moral character in technologists may reflect the rapid changes and greater technological capabilities and increasingly pervasive scope of application of ICT in society, the growing discussion of digital technologies in media and politics, and the need for computer science professionals today to be able to engage with a range of ethical issues arising in many areas of personal and public life.

Clearly, the literature on digital ethics education refers to a broad variety of pedagogical approaches or theories. This appears to be partly because teaching ethics in computer and information sciences involves the synthesis of several disciplines, and partly due to the fact that, despite the need for ethics education being flagged thirty years ago, researchers are still working on formulating the pedagogical basis of ethics education for tertiary ICT students. As we have noted, digital ethics as a discipline is often yet to play a significant role in curricula.

In addition, there are likely to be ongoing disagreements about which pedagogical approach(es) are best. There is no good reason why, however, teachers cannot draw from a range of pedagogical theories. For example, they might use constructivist approaches that build on students' existing knowledge of a technology use while encouraging reflection on its social implications and encouraging them to reflect on their own experiences with that technology. From the perspective of teachers who are not expert in pedagogical theory, drawing on different pedagogical approaches seems entirely reasonable.

The relative newness of digital ethics education may partly explain the fact that some educators taught digital ethics as standalone units while others preferred to integrate ethics into the ICT curriculum. Knowing which, if any, is the more effective mode of teaching—standalone versus integrated versus hybrid—for ICT students will need to be investigated with empirical studies. There was, however, a general feeling amongst scholars that implementing "ethics across the curriculum" (Borenstein & Howard, 2021) could be beneficial for learning. This could be partly due to the nature of the students being taught, namely, ICT students, rather than, say, humanities students, who may have more specialised interests in ethical and social questions.

Some authors pointed out that teaching digital ethics to ICT students in dedicated philosophy departments may be relatively ineffective (see e.g., Carter & Crockett, 2019). In such cases, the ethics content is separated from the details and activities connected with ICT. In contrast, embedding ethics into ICT courses may allow students to more clearly see the relevance of ethics and to connect it more substantially to their interest and future work. Such an approach is aligned with the pedagogical approaches of constructivism and constructionism, whereby assimilating new knowledge and understanding builds on existing knowledge and understanding and is formed and shaped in particular contexts. It is also well aligned with reflexive and deliberative pedagogical approaches, in which learners reflect on and question their practices and engage in deliberation and discussion about their values concerning technology. Making ethics relevant to the technological interests and activities of students is clearly going to be important for engaging students in the journey of ethics learning as well as imparting understanding.

While it may be true that most ICT students do not have a primary interest in ethics and social issues, it is worth pointing out that there may be an increasing appetite for such learning. A mounting hunger for this knowledge and insight may be due to developments we have been emphasizing, namely the substantial and growing media, cultural, political, and social interest in the responsible use of digital technologies. In our own teaching experience, students from IS and CS are enrolling in burgeoning numbers in digital ethics subjects. Moreover, more-and-more new digital ethics courses are springing up in universities and as online options, potentially reflecting a greater desire for learning about digital ethics.

Who should teach digital ethics to IS and CS students? The benefits of bringing ethics teaching closer to the interests, values, and activities of ICT students does not necessarily imply that such teaching should be done by educators whose expertise is in ICT, though that may also be beneficial. If it is important to teach moral theory, as some scholars believe, then it may be necessary to recruit educators with expertise in ethics and philosophy. One obvious remedy here is to have digital ethics taught by experts in both ICT and ethics. This was illustrated in courses including guest lecturers from other disciplines (Skirpan et al., 2018) or partnering philosophy and computer ethics departments in designing and delivering computing ethics modules (Grosz et al., 2019). Indeed, Borenstein and Howard go a step further and suggest that digital ethics could be taught by a larger interdisciplinary team including lawyers, sociologists, and philosophers together with scientists and engineers (Borenstein & Howard, 2021). Such experts could deliver content separately, or else collaborate in a multi-disciplinary way, to craft course materials that may be more appealing and relevant to ICT students.

The selection of moral theories in digital ethics teaching warrants further comment. Earlier we mentioned the preference some have for teaching virtue ethics over other moral theories. In contrast to theories like utilitarianism and deontology, virtue ethics stresses the importance of developing stable characters that involve the right kind of action, feeling, and attitude (Crisp & Slote, 1997). However, while it may be important to society as well as to engineering disciplines to encourage good character development in future ICT engineers, it would arguably be lop-sided to teach only virtue ethics theory to students. Other moral theories, certainly utilitarianism and deontology but also care ethics, are equally well recognized in moral philosophy and many kinds of applied ethics. These theories also challenge aspects of virtue ethics; teaching them alongside that theory would provide for a more balanced understanding of how to approach digital ethical problems.

Furthermore, one may ask whether there could be some pedagogical benefits to widening the scope to include other moral approaches. This could include, for example, introducing students to Eastern ethics, eco-feminism, and/or other less traditional frameworks (Singer, 2013). There are also various Indigenous ethics approaches in different countries. Consider, for example, the approach of Indigenous Australians to the effects of technology on country (Graham, 1999). The choice of such moral frameworks may be determined by the nature of the student body as well as educators' expertise. For example, students from some Asian backgrounds may already have some knowledge of (say) Buddhist and Confucian ethics, while students from some African backgrounds may have knowledge of ethical approaches such as Ubuntu. One could easily imagine that moral approaches which mesh well with students' cultural backgrounds will be especially useful in stimulating learning about digital ethics, and may also align well with, for example, constructivist/constructionist, social analysis, and reflexive/deliberative pedagogical models preferred by some educators.

4.2 Questions for further inquiry

Our literature review indicates significant scope for further inquiry. We note here two questions arising from our findings. The first of these regards CIS ethics education for professionals who have completed their university training and are working in the field. While the issue was raised in two of the studies reviewed (Hildt et al., 2019; Plemmons et al., 2020), both these projects involved student researchers working within university laboratories. Our review reveals a relative gap in understanding of the need and nature of digital ethics education outside of the university environment.

In the work by Plemmons et al. (2020) a randomised trial of research ethics training aimed at making ethics discourse routine practice in university STEM labs was conducted. The authors suggest that a lack of communication about ethical considerations in these environments resulted from researchers' equating ethics with regulatory compliance, and that this flowed from a disconnect between Responsible Conduct of Research (RCR) training and the reality of the laboratory. The authors used a customised training script to provide opportunities for members to discuss present and future relevant behaviour. The pre- and post- program surveys indicated that following the intervention, researchers were more engaged in discussions about the ethical implications of their work. As stated above, however, the scope of the ethical training, similarly to RCR principles, related to practice within the laboratory, and did not extend to broader ethical implications of the technology being produced by the work. In this sense, as in the case of Hildt et al. (2019), the breadth of the focus would be insufficient to address the ethical issues relevant to CIS professionals working in industry or students planning to move out of the university environment. Both studies proposed programs for STEM students and researchers (who in their roles act both as students and educators), and not specifically CIS students. It may be that in CIS, the broader social and ethical implications of the technology demand a broader approach to CIS ethics. In such a case, ethics training focused specifically on ethical conduct within a university laboratory, while important, is limited.

Thus, there is potential to consider an example of teaching beyond universities. Hildt et al. (2019) proposed an intervention which moved "out of the classroom" to integrate ethics education into a laboratory work environment. The authors emphasised a "bottom up" approach of working with ethical issues arising in the workplace, rather than using instructor-presented dilemmas. Interestingly, the authors noted a distinction between the broader "bottom up" meaning of ethics in the research context expressed by student participants - focusing on communication, power imbalances and social issues - and the narrower "top down" focus held by faculty – focusing on issues of plagiarism, data quality and code of conduct requirements. This finding raises the question of whether ethics training may need to be tailored differently for professionals in different roles and different stages of their career.

A second question is whether there is a need for specifically designed ethics training programs for students and professionals designing and analysing CS technology employed in particular fields. In recent years, a number of studies have addressed the need for training in different professions using CS technology, not just in technological proficiency, but in the knowledge and skills required to address ethical issues arising with the use of the technology in that field. For example, in the field of medicine, scholars have recently examined the need for specifically teaching "health AI" or "medical AI" ethics, along with the required technical skill, in medical education courses (see Civaner et al., 2022; Katznelson & Gerke, 2021; Quinn & Coghlan, 2021; Weidener & Fischer, 2023). Similarly, studies have considered the need to train teachers in the uses and ethical implications of digitalised teaching tools (Buchanan, 2019) and digital journalism (García-Avilés, 2014), and the need to train social workers in the ethical issues inherent in use of online technologies used in their profession (Goldingay & Boddy, 2016). These studies, however, consider training in dealing with ethical considerations arising when employing technology designed for use in particular professions. The question we raise here is whether specialised ethics training is needed in educating CIS students and professionals involved in the design of medical, teaching, journalistic, social work technologies. As many professions become increasingly reliant on specialized AI tools, this question gains in significance.

5 Conclusion

In our narrative literature review, we have progressed the understanding of ethics in CIS education, by identifying several pedagogical approaches to teaching digital ethics, differences in the degrees to which ethics components are integrated into ICT courses, various moral theories in ethics teaching practices, and various pedagogical tools. When designing digital ethics teaching for IS students and other students in ICT fields, educators might profitably reflect on these different dimensions of ethics teaching and the interconnections between them that we discussed above. It is likely good teaching in digital ethics will involve using a range of theoretical techniques and teaching and learning approaches rather than just one or two.

The necessity of teaching digital ethics to future ICT practitioners is increasingly accepted. To improve digital ethics education, further work in this space will be important. This would include a more systematic review and analysis of the relevant literature and exploring in more depth the roles, linkages, and suitability of various pedagogical theories, levels of embeddedness of ethics material, and the philosophical nature of the subject matter in digital ethics teaching. There is also benefit in extending the scope of inquiry to also examine ongoing ethics education of ICT practitioners engaged in different roles and industries, and exploring a possible need for specialised ethics training in designing ICT technology for use within particular professions such as medicine, journalism, and teaching.

References

Ackermann, E. (2001). Piaget's Constructivism, Papert's Constructionism: What's the Difference. http://learning.media.mit.edu/content/publications/EA.Piaget%20_%20 Papert.pdf

Ali, S., Williams, R., Payne, B., Park, H., & Breazeal, C. (2019). Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. *International Workshop on Education in Artificiation Intelligence K-12 (EDUAI '19*), USA. https://robots.media.mit.edu/wp-content/uploads/sites/7/2019/08/Constructionism__Ethics__and_Creativity.pdf

Baethge, C., Goldbeck-Wood, S. & Mertens, S. (2019). SANRA—a scale for the quality assessment of narrative review articles. *Rebidwells Integr Peer Rev, 4*(5), 1–7. https://doi.org/10.1186/s41073-019-0064-8

Barabas, C., Doyle, C., Rubinovitz, J. B., & Dinakar, K. (2020). Studying up: Reorienting the Study of Algorithmic Fairness around Issues of Power. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20*, New York, NY, USA, 167–176. https://doi.org/10.1145/3351095.3372859

Barnard, A., de Ridder, C., Pretorius, L., & Cohen, E. (2003). Integrating Computer Ethics into the Computing Curriculum: A Framework for Implementation. *Proceedings of the 2003 InSITE Conference, 3*, 265–279. https://doi.org/10.28945/2619

Bates, J., Cameron, D., Checco, A., Clough, P., Hopfgartner, F., Mazumdar, S., Sbaffi, L., Stordy, P., & de la Vega de León, A. (2020). Integrating FATE/Critical Data Studies into Data Science Curricula: Where Are We Going and How Do We Get There? *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20*, New York, NY, USA, 425–435. https://doi.org/10.1145/3351095.3372832

Beauchamp, T. L. & Childress, J. F. (1994). *Principles of Biomedical Ethics*. Oxford, UK: Oxford University Press.

Ben-Ari, M. (2001). Constructivism in Computer Science Education. *Journal of Computers in Mathematics and Science Teaching, 20*(1), 45–73. https://www.learntechlib.org/p/8505

Bezuidenhout, L., & Ratti, E. (2021). What Does It Mean to Embed Ethics in Data Science? An Integrative Approach Based on Microethics and Virtues. *AI & Society, 36*(3), 939–953. https://doi.org/10.1007/s00146-020-01112-w

Bidwell, N. (2010). Ubuntu in the network: humanness in social capital in rural Africa. *Interactions*, *17*(2), 68–71. https://doi.org/10.1145/1699775.1699791

Bogina, V., Hartman, A., Kuflik, T., & Shulner-Tal, A. (2021). Educating Software and AI Stakeholders about Algorithmic Fairness, Accountability, Transparency and Ethics. *International Journal of Artificial Intelligence in Education*, *3*, 808-833. https://doi.org/10.1007/s40593-021-00248-0

Borenstein, J., & Howard, A. (2021). Emerging Challenges in AI and the Need for AI Ethics Education. *AI and Ethics*, *1*(1), 61–65. https://doi.org/10.1007/s43681-020-00002-7

Bradley, A. (2008). Accreditation Criteria Guidelines. No. G02, Engineers Australia Accreditation Board. https://www.engineersaustralia.org.au/sites/default/files/content-files/2016-12/G02_Accreditation_Criteria_Guidelines.pdf

Briggle, A., Holbrook, J. B., Oppong, J., Hoffmann, J., Larsen, E. K., & Pluscht, P. (2016). Research Ethics Education in the STEM Disciplines: The Promises and Challenges of a Gaming Approach. *Science and Engineering Ethics*, *22*(1), 237–250. https://doi.org/10.1007/s11948-015-9624-6

Buchanan, R. (2019). Digital Ethical Dilemmas in Teaching. In Peters, M.A. (Ed.), *Encyclopedia of Teacher Education*. Singapore: Springer nature Singapore. http://link.springer.com/10.1007/978-981-13-1179-6_150-1

Burton, E., Goldsmith, J., & Mattei, N. (2018) How to Teach Computer Ethics through Science Fiction.' *Communications of the ACM*, *61*(8), 54-84.

Bynum, T. (2015). Computer and Information Ethics. In Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy (Summer 2018 Edition)*. https://stanford.library.sydney.edu.au/archives/sum2019/entries/ethics-computer/

Califf, M. E., & Goodwin, M. (2005). Effective Incorporation of Ethics into Courses That Focus on Programming. *Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education - SIGCSE '05*. New York, NY, USA, 347–351. https://doi.org/10.1145/1047344.1047464

Carter, L., & Crockett, C. (2019). An Ethics Curriculum for CS with Flexibility and Continuity. *2019 IEEE Frontiers in Education Conference (FIE)*. IEEE, Cincinnati, OH, USA, October 16-19, 2019. https://doi.org/10.1109/fie43999.2019.9028356

Chatila, R., & Havens, J. C. (2019). The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *Robotics and Well-Being*. International Series on Intelligent Systems, Control and Automation: Science and Engineering, *95*, 11–16. https://doi.org/10.1007/978-3-030-12524-0_2

Civaner M.M., Uncu, Y., Bulut, F., Giounous Chalil, E., & Tatli, A. (2022). Artificial Intelligence in Medical Education: A Cross-Sectional Needs Assessment. *BMC Medical Education*, *22*(1), 772–780.

Crisp, R., & Slote, M. A. (Eds.) (1997). *Virtue Ethics*. Oxford: Oxford University Press.

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* New York, NY, USA: St. Martin's Publishing Group.

Farooqui, J., Dawod, S., & Sarma, D. (2021). Towards a Neuroethical Ethos: A Case Study in Reframing Neuroethics Education for Engineers and Researchers. *The Neuroethics Blog, Emory Center for Ethics*. Accessed January 24, 2022. http://www.theneuroethicsblog.com/2021/02/towards-neuroethical-ethos-case-study.html

Floridi, L., Cath, C., & Taddeo, M. (2019). Digital Ethics: Its Nature and Scope. *The 2018 Yearbook of the Digital Ethics Lab*, 9–17. https://doi.org/10.1007/978-3-030-17152-0_2

García-Avilés, J.A. (2014). Online Newsrooms as Communities of Practice: Exploring Digital Journalists' Applied Ethics. *Journal of Mass Media Ethics, 29*(4), 258-272.

Goldingay, S. & Boddy, J. (2017). Preparing Social Work Graduates for Digital Practice: Ethical Pedagogies for Effective Learning. *Australian Social Work, 70*(2), 209–220.

Goldsmith, J., & Burton, E. (2017). Why Teaching Ethics to AI Practitioners Is Important. *Proceedings of the 31st AAAI Conference on Artificial Intelligence,* 4836-4840. https://doi.org/10.1609/aaai.v31i1.11139

Graham, M. (1999). Some Thoughts about the Philosophical Underpinnings of Aboriginal Worldviews. *Worldviews: Environment, Culture, Religion, 3*(2), 105–118. https://doi.org/10.1163/156853599x00090

Greenhalgh T., Thorne S., & Malterud K (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation, 48*(6). https://doi: 10.1111/eci.12931

Grosz, B. J., Grant, D. G., Vredenburgh, K., Behrends, J., Hu, L., Simmons, A., & Waldo, J. (2019). Embedded EthiCS. *Communications of the ACM, 62*(8), 54–61. https://doi.org/10.1145/3330794

Hagendorff, T.(2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines, 30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hämäläinen, J. (2012). Pedagogic Theories. *Oxford Bibliographies Online Datasets*. Oxford, UK: Oxford University Press (OUP). https://doi.org/10.1093/obo/9780199791231-0015

Hassan, N. R., Mingers, J., & Stahl, B. (2018). Philosophy and Information Systems: Where Are We and Where Should We Go. *European Journal of Information Systems: An Official Journal of the Operational Research Society, 27*(3), 263–277. https://doi.org/10.1080/0960085x.2018.1470776

Herkert, J. (2004). Microethics, Macroethics, and Professional Engineering Societies. *Emerging Technologies and Ethical Issues in Engineering: Papers from a Workshop*. National Academy of Engineering. Washington, DC, USA: The National Academies Press. https://doi.org/10.17226/11083

Hildt, E., Laas, K., Miller, C., Taylor, S., & Brey, E. M. (2019). Empowering Graduate Students to Address Ethics in Research Environments. *Cambridge Quarterly of Healthcare Ethics: CQ: The International Journal of Healthcare Ethics Committees, 28*(3), 542-550. https://doi.org/10.1017/S096318011900046X

Huff, C., & Martin, C. D. (1995). Computing Consequences. *Communications of the ACM, 38*(12), 75–84. https://doi.org/10.1145/219663.219687

Johnson, D. G. (1985). *Computer Ethics*, Englewood Cliffs, NJ, USA: Prentice-Hall.

Jones, S. (2016). Doing the Right Thing: Computer Ethics Pedagogy Revisited. *Journal of Information Communication and Ethics in Society, 14*(1), 33–48. https://doi.org/10.1108/jices-07-2014-0033

Katznelson, G. & Gerke, S. (2021). The Need for Health AI Ethics in Medical School Education. *Advances in Health Sciences Education, 26*(4), 1447–1458.

Kohlberg, L. (1984). *The Psychology of Moral Development: The Nature and Validity of Moral Stages*. New York, NY, USA: Harper & Row.

Kordzadeh, N. & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems, 31*(3), 388–409. https://doi.org/10.1080/0960085X.2021.1927212

Lewis, A., & Stoyanovich, J. (2021). Teaching Responsible Data Science: Charting New Pedagogical Territory. *International Journal of Artificial Intelligence in Education*, 32, 783-807. https://doi.org/10.1007/s40593-021-00241-7

Li, J., & Fu, S. (2012). A Systematic Approach to Engineering Ethics Education. *Science and Engineering Ethics, 18*(2), 339–349. https://doi.org/10.1007/s11948-010-9249-8

Martin, D. A., Conlon, E., & Bowe, B. (2021). A Multi-Level Review of Engineering Ethics Education: Towards a Socio-Technical Orientation of Engineering Education for Ethics. *Science and Engineering Ethics, 27*(5), Article 60. https://doi.org/10.1007/s11948-021-00333-6.

Mason, R. O. (1986). Four Ethical Issues of the Information Age. *MIS Quarterly, 10*(1), 5–12. https://doi.org/10.2307/248873

Montréal Declaration for a Responsible Development of Artificial Intelligence (2018). https://montrealdeclaration-responsibleai.com/

Moor, J. H. (1985). What Is Computer Ethics? *Metaphilosophy, 16*(4), 266–275. https://doi.org/10.1111/j.1467-9973.1985.tb00173.x

Moriarty, J. (2021). Business Ethics. In Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2021/entries/ethics-business/

Müller, V. C. (2020). Ethics of Artificial Intelligence and Robotics. In Zalta, E. N. (Ed.), *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2020/entries/ethics-ai/

Mumford, E. (1995). *Effective Systems Design and Requirements Analysis: The ETHICS Approach*. London, UK: Macmillan.

Mumford, M. D., Connelly, S., Brown, R. P., Murphy, S. T., Hill, J. H., Antes, A. L., Waples, E. P., & Devenport, L. D. (2008). A Sensemaking Approach to Ethics Training for Scientists: Preliminary Evidence of Training Effectiveness. *Ethics & Behavior, 18*(4), 315-339. https://doi.org/10.1080/10508420802487815

National Research Council and National Academy of Engineering. 2014. *Emerging and Readily Available Technologies and National Security: A Framework for Addressing Ethical, Legal, and Societal Issues*. Washington, DC, USA: The National Academies Press. https://doi.org/10.17226/18512

Newberry, B. (2004). The Dilemma of Ethics in Engineering Education. *Science and Engineering Ethics*, *10*(2), 343–351. https://doi.org/10.1007/s11948-004-0030-8

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY, USA: Broadway Books.

Papert, S. (1993). *Mindstorms*, London, UK: Basic Books.

Papert, S., & Harel, I. (1991). Situating Constructionism. In Papert, S., & Harel, I. (Eds.), *Constructionism*. New York, NY, USA: Ablex Publishing Corporation.

Piaget, J. (1971). *Psychology and Epistemology: Towards a Theory of Knowledge*. London, UK: Penguin Books.

Plemmons, D. K., Baranski, E. N., Harp, K., Lo, D. D., Soderberg, C. K., Errington, T. M., Nosek, B. A., & Esterling, K. M. (2020). A Randomized Trial of a Lab-Embedded Discourse Intervention to Improve Research Ethics. *Proceedings of the National Academy of Sciences of the United States of America, 117*(3), 1389–1394. https://doi.org/10.1073/pnas.1917848117

Quinn, T.P. & Coghlan, S. (2021) Readying Medical Students for Medical AI: The Need to Embed AI Ethics Education (No arXiv:2109.02866, arXiv, 7 September 2021). http://arxiv.org/abs/2109.02866

Rest, J. R., Darcia, N. ez, Thoma, S. J., Bebeau, M. J.. (1999). *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*. Milton Park, UK: Psychology Press (Routledge).

Richards, D., Formosa, P., Ryan, M., Hitchens, M., & McEwan, M. (2020). A Proposed AI-Enhanced Serious Game for Cybersecurity Ethics Training. *Proceedings of the 9th Conference of the Australian Institute of Computer Ethics (AiCE 2020)*, 1–9. https://researchers.mq.edu.au/en/publications/a-proposed-ai-enhanced-serious-game-for-cybersecurity-ethics-trai

Singer, P. (Ed.). (2013). *A Companion to Ethics* (1st ed.), Blackwell Companions to Philosophy. London, UK: Blackwell.

Skirpan, M., Beard, N., Bhaduri, S., Fiesler, C., & Yeh, T. (2018). Ethics Education in Context. *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, New York, NY, USA, 940–945. https://doi.org/10.1145/3159450.3159573

Someh, I., Davern, M., Breidbach, C. F., & Shanks, G. (2019). Ethical Issues in Big Data Analytics: A Stakeholder Perspective. *Communications of the Association for Information Systems*, *44*, Article 34, 718–747. https://doi.org/10.17705/1CAIS.04434

Sprague, A. P., & Diaz-Sprague, R. (2019). Tying Ethics to Teamwork Training in a Minimodule. *Proceedings of the 8th Computer Science Education Research Conference*, *CSERC '19*, New York, NY, USA, 120–122. https://doi.org/10.1145/3375258.3375275

Towell, E., Thompson, J. B., & McFadden, K. L. (2004). Introducing and Developing Professional Standards in the Information Systems Curriculum. *Ethics and Information Technology 6*(4), 291–299. https://doi.org/10.1007/s10676-005-5985-x

Tubig, P., & McCusker, D. (2021). Fostering the Trustworthiness of Researchers: SPECS and the Role of Ethical Reflexivity in Novel Neurotechnology Research. *Research Ethics, 17*(2), 143-161. https://doi.org/10.1177/1747016120952500

Tucker, A. B. (1991). Computing Curricula 1991. *Communications of the ACM, 34*(6), 68–84. https://doi.org/10.1145/103701.103710.

Véliz, C. (ed.). (2021). *The Oxford Handbook of Digital Ethics*, Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198857815.001.0001

Weidener, L. & Fischer, M. (2023). Artificial Intelligence Teaching as Part of Medical Education: Qualitative Analysis of Expert Interviews. *JMIR Medical Education, 9,* Article e46428. DOI: 10.2196/46428

Wong, P. H. (2012). Dao, harmony and personhood: Towards a Confucian ethics of technology. *Philosophy & Technology*, 25, 6786.

Appendix

Table 2. An overview table of relevant papers discussed, and key themes/ideas discussed within

Legend

Cons: Constructivist/constructionist approaches

Soc: Social analysis approaches

M/M: Microethics/macroethics approaches (including evaluation of tradeoffs between the two)

Cog: Cognitivism/moral reasoning

R/D: Reflexive/deliberative approaches

EC/FEAT: Ethical computing elements/Fairness Ethics Accountability and Transparency

TSRD: Ethical traits, skills, reasoning and discourse

Case: Case studies

Deliv: Deliverables/concrete educational materials for *in situ* learning

Gam: Gaming/gamification approaches

**: Asterisks in "integration/hybrid approaches" indicates an evaluation of tradeoffs with the 'standalone approaches' technique

Note that the table is non-exhaustive; papers may overlap in their coverage.

Paper	Pedagogical Approaches					Integration/ Hybrid Approaches	Inter-disciplinary Approaches	Moral Theory Selection/ Combination	Course Content		Pedagogical Tools		
	Cons	Soc	M/M	Cog	R/D				EC/ FEAT	TSRD	Case	Deliv	Gam
Ali et al., 2019	✓												
Barabas et al., 2020					✓								
Barnard et al., 2003		✓							✓				
Bates et al., 2020	✓					✓	✓		✓				
Ben-Ari, 2001	✓												
Bezuidenhout and Ratti, 2021			✓		✓			✓		✓		✓	
Bogina et al., 2021						✓			✓				
Borenstein & Howard, 2021						✓	✓						
Briggle et al., 2016	✓							✓					✓
Burton et al., 2018								✓			✓		
Burton, Goldsmith & Mattei, 2018										✓			
Carter & Crockett, 2019		✓				✓		✓			✓		
Chatila & Havens, 2019								✓					
Farooqui et al., 2021					✓						✓		
Goldsmith & Burton, 2017								✓					
Greengard, 2020									✓				
Grosz et al., 2019	✓					✓	✓			✓			
Hagendorff, 2020			✓					✓					
Herkert, 2004			✓										
Hildt et al., 2019	✓					✓**					✓		
Hjorth, 2019	✓												
Holmes et al., 2021	✓												
Huff & Martin, 1995		✓				✓							

Paper	Pedagogical Approaches					Integration/ Hybrid Approaches	Inter-disciplinary Approaches	Moral Theory Selection/ Combination	Course Content		Pedagogical Tools		
	Cons	Soc	M/M	Cog	R/D				EC/ FEAT	TSRD	Case	Deliv	Gam
Jones, 2016								✓					
Lewis & Stoyanovich, 2021	✓								✓			✓	
Li and Fu, 2012			✓										
Mumford et al., 2008;				✓							✓		
Newberry, 2004											✓		
Plemmons et al., 2020					✓	✓**							
Richards et al., 2020	✓			✓				✓					✓
Skirpan et al., 2018	✓					✓	✓						
Sprague & Diaz-Sprague, 2019				✓									
Towell, Thompson & McFadden 2004											✓		
Tubig & McCusker, 2021					✓								
Wise, 2020	✓												

doi: https://doi.org/10.3127/ajis.v27i0.4517