12-7-2022

Teaching Digital Ethics in Information Systems

Minna Paltiel
Melbourne Law School, University of Melbourne, minna.paltiel@unimelb.edu.au

Marc Cheong
School of Computing and Information Systems, University of Melbourne, marc.cheong@unimelb.edu.au

Simon Coghlan
School of Computing and Information Systems, University of Melbourne, simon.coghlan@unimelb.edu.au

Reeva Lederman
School of Computing and Information Systems, University of Melbourne, reeva.lederman@unimelb.edu.au

Teaching Digital Ethics in Information Systems

Full research paper

Minna Paltiel
Melbourne Law School
The University of Melbourne
Parkville, VIC, Australia
Email: minna.paltiel@unimelb.edu.au

Marc Cheong
School of Computing and Information Systems, Faculty of Engineering and IT
The University of Melbourne
Parkville, VIC, Australia
Email: marc.cheong@unimelb.edu.au

Simon Coghlan
School of Computing and Information Systems, Faculty of Engineering and IT
Centre for AI and Digital Ethics (CAIDE)
The University of Melbourne
Parkville, VIC, Australia
Email: simon.coghlan@unimelb.edu.au

Reeva Lederman
School of Computing and Information Systems, Faculty of Engineering and IT
The University of Melbourne
Parkville, VIC, Australia
Email: reeva.lederman@unimelb.edu.au

Abstract

This paper reviews and discusses the little-explored topic of digital ethics education in Information Systems and related fields. The importance of teaching digital ethics to students studying information and communication technologies (ICT) is increasingly recognised. However, it is unclear how ethics should best be taught to these future ICT practitioners. Using a narrative literature review methodology to explore existing scholarly work, this paper identifies five pedagogical theories related to digital ethics education. Additionally, the paper outlines approaches that deploy standalone ethics units, integrated ethics teaching, and hybrid teaching approaches, and identifies the employment of and emphasis given to various moral theories in digital ethics education. The paper then discusses how these three sets of findings—namely, different pedagogical approaches, degree of integration of ethics teaching, and use of moral theories—are related to each other. It provides educators with information and reflections to consider when designing digital ethics teaching.

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

1 Introduction

Digital ethics is an applied ethics discipline that explores moral questions raised by information and communication technologies (ICT), from laptops and smartphones to social media and modern artificial intelligence (AI) (Véliz 2021). The need for teaching digital ethics in Information Systems (IS) and related fields like computer science (CS) has been recognised for some time (Johnson 1985). For example, ethics teaching requirements were included in the ACM/IEEE-CS *Computing Curriculum 1991* (Tucker 1991) and in the Computer Science Accreditation Board curriculum standards since 1987 (Califf and Goodwin 2005; Huff and Martin 1995). Today, people increasingly believe that some sort of digital ethics education is required to address the pervasive impact of ICT in society—a belief reflected in current requirements for accreditation in engineering generally (Bradley 2008; Martin et al. 2021) and computing and information system sciences specifically.

But how is and should digital ethics be taught? Digital ethics comprises computer ethics, information ethics (Bynum 2015), and AI ethics (Müller 2020). It is a relatively new discipline, not yet well established in university curricula. Further, many who teach ethics in IS and related disciplines lack experience and training in philosophical theory and digital ethics pedagogy. Programs and interventions for teaching ethics vary greatly in structure, methods, pedagogical approach, and in the educational tools they employ. In this paper, we review the literature on teaching ethics in ICT-based university courses, focusing on more recent scholarly work.

Our review uncovers several themes and trends that concern recommendations for digital ethics teaching. An emphasis exists on ensuring that ethics is taught as inextricable from the activities of ICT practitioners. As such, integrating ethics into the curriculum rather than presenting stand-alone ethics courses, is seen as desirable. This emphasis is also seen in adopting pedagogical approaches that teach ethics through direct experience, *in situ*, and within the ICT context. Another trend in the literature concerns the need to merge different ways of thinking, to include social and learning sciences approaches in designing digital ethics courses and to create multidisciplinary partnerships. A final theme concerns building ethical skills and virtues rather than knowledge sets. The paper runs as follows. Section 2 outlines the nature and importance of digital ethics, Section 3 describes our methodology, Section 4 presents our findings, Section 5 discusses them and offers points for digital ethics educators to consider, and Section 6 concludes our analysis of teaching digital ethics in IS and similar disciplines.

2 Digital ethics and its importance

Digital ethics is a growing field which addresses many issues, including data security and privacy which are core IS concerns, as well as algorithmic fairness and transparency, superintelligence, social media ethics, AI explainability, data law, the power of tech giants, surveillance, moral machines, and more (Floridi et al. 2019; Véliz 2021).

Recognition of the importance of ethics education in the computing sciences goes back some decades (Moor 1985). 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 (O'Neil 2016) in business, services, and society generally have only increased the significance of their observation. Such technology can carry substantial risks. Modern algorithms, as Borenstein and Howard (2021) observe, may be inaccurate, biased, and unfair, reflecting human fallibility and prejudice (Eubanks 2018).

The apparent need for ethics education is not limited to algorithm designers such as computer scientists and engineers, but also applies to IS practitioners and academics. 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 (1986)'s "prominent contribution to IS ethics" in the 1980's concerned issues still very much present in contemporary IS discourse: "privacy, accuracy, property, accessibility" (Hassan et al. 2018), or 'PAPA'.

These developments further support the claim that ICT professionals must be 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 educators to train future professionals (and other community stakeholders) not only in requisite technical proficiencies, but also in ethics (Borenstein and Howard 2021; Grosz et al. 2019; Skirpan et al. 2018).

3 Search methodology

To obtain a broad overview of the relevant literature, we conducted a narrative review to identify papers about digital ethics teaching for students in IS and related fields. 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,* and *Ethics and Information Technology*. *C*onference proceedings were also accessed, including: *Conferences on Fairness, Accountability and Transparency, Conferences of the Institute of Electrical and Electronics Engineers (IEEE)*, conferences of the *Australasian Institute of Computer Ethics* and *Communications of the ACM,* a monthly journal of the Association for Computing Machinery (US). Additional searches were conducted on Google and Google Scholar.

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'/'FAcT'* 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.

4 Findings

From our literature review, a number of themes and trends emerge as recommended dimensions of teaching to prepare for the pervasive impact and subsequent responsibility carried by ICT scientists in the digital age. Our findings reveal, first, an emphasis amongst scholars on ensuring that ethics is taught as inextricable from the activities of an ICT practitioner, by adopting pedagogical approaches that teach ethics through direct experience, *in situ,* and within an ICT context (S4.1). The second feature we found relates to the need for integrating ethics into the IS/CS curriculum as opposed to presenting stand-alone ethics courses (S4.2). A third theme is a focus on building ethical skills or developing moral values and virtues rather than acquiring a particular knowledge set. Here there is considerable reference to the 'virtue ethics' school and its focus on developing 'ethical character' (S4.3). A final theme in the literature is 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 partnerships (S4.4).

4.1 Pedagogical theories

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. Despite the wide range of pedagogical approaches referred to, several primary theories stand out, which we discuss in turn.

4.1.1 Constructivist/constructionist approaches

A dominant pedagogical theory referred to in the literature is *constructivism* (Ben-Ari 2001), according to which knowledge is not passively absorbed but actively constructed by a learner in interacting with her world (Ackermann 2001; Piaget 1971). This occurs through processes of 'assimilation' and 'accommodation' and through integrating experiences and adapting mental schemas accordingly. A constructivist approach underpinned a number of the digital ethics programs reviewed in the literature (Bates et al. 2020; Lewis and Stoyanovich 2021).

Seymour Papert's *constructionist* theory of learning (Ackermann 2001; Papert 1993) is based on Piaget's constructivism and shares with it the principle that learning occurs through active engagement with the environment and internalisation of experience. However, constructionism also says that this process occurs particularly when 'the learner is engaged in constructing a public entity'—that is, it occurs by 'learning through making' (Papert and Harel 1991). According to constructionism, knowledge is formed and transformed in particular contexts (and through particular individuals) and via particular uses and media (Ackermann 2001). A number of courses in the literature used constructivism and/or constructionism as guiding principles for digital ethics programs which involve learning from direct experience (Briggle et al. 2016; Holmes et al. 2021; Richards et al. 2020). These programs are also

project based (Ali et al. 2019; Hildt et al. 2019), interactive (Lewis and Stoyanovich 2021), and *in situ* (Skirpan et al. 2018), involving design or analysis and including 'deliverables' such as 'nutritional labels' (Lewis and Stoyanovich 2021). These closely related approaches encourage 'situated' ethical reasoning (Grosz et al. 2019).

4.1.2 Social analysis approaches

Huff and Martin suggested that only analysis that accounts for at least three dimensions—technical, social, and ethical—can 'represent the issues as they affect computer science in practice' (Huff and Martin 1995). These authors referred to *Project ImpactCS*, begun in 1994, which described a 'social analysis' teaching approach in which a technology, an ethical or social issue is identified, the appropriate level of social analysis (e.g., individuals, communities or groups, organisations, cultures, institutions etc.) is determined, and appropriate tools, literature and methods are then applied. They identified the importance of developing a corresponding skill set in computer professionals, namely: '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' (Huff and Martin 1995; see also Barnard et al. 2003).

A more recent example is the CS ethics curriculum advanced 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 and Crockett 2019). Carter and Crocket's work, and moral reasoning approaches, are discussed further below.

4.1.3 Approaches emphasising microethics versus macroethics

The social analysis approach, and emphasis on large social issues, is congruent with a 'macroethics' approach to ethics education.. The 'macroethics' approach in engineering and information sciences is concerned with 'the collective, social responsibility' of the professional and 'societal decisions about technology', while the 'microethics' approach is concerned with individual responsibility and internal relations within the profession (Herkert 2004). Li and Fu (2012) hold that the micro-ethics 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 and Ratti 2021). They sought to foster 'moral virtue' by focusing on individual decision making and action, rather than on key themes and high-level case studies found in 'macroethics' discourse. The aim was to develop 'moral excellence' through virtue training, which is in contrast both to the approach described above focusing on identification of social issues, and to 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 and Ratti 2021).

4.1.4 Cognitivist or moral reasoning approaches

Moral cognition and education were referred to in many of the courses we reviewed (Mumford et al. 2008; Richards et al. 2020; Sprague and Diaz-Sprague 2019), These were 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 involves a complex cognitive process activated when people are presented with 'ambiguous high-stake events'. Ethical decision making is based on the available case-based models applied to the situation with a 'field practices approach' (codes of conduct and guidelines specific to a particular field) to 'mak[e] sense' of the ethical dilemma (Mumford et al. 2008).

The use of case studies for ethics teaching has been a commonly used pedagogical tool for a number of decades (Carter and Crockett 2019; Hildt et al. 2019; Mumford et al. 2008; Newberry 2004; Towell et al. 2004). Some of the programs in the literature relied primarily on readings and reflective discussions (Farooqui et al. 2021). However, typically the case studies are used as a trigger for role play, simulations, and interactive activities (e.g., Towell et al. 2004; Hildt et al. 2019). These techniques have been associated with cognitivist theories of moral psychology like the Kohlbergian and Neo-Kohlbergian schools.

4.1.5 Reflexive and deliberative approaches

The final pedagogical theories in the literature on digital ethics education for ICT students are a reflexive and deliberative approach.

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

The related, deliberative approach courses refers to enabling discussion and interaction with others with different viewpoints. It emphasises the development of an ethical 'culture' rather than individual morality. Plemmons et al reported on a STEM research ethics training intervention which sought to improve participants' 'reason giving' and 'interpersonal communication' abilities for more ethical practice (Plemmons et al. 2020). It held 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 (Plemmons et al. 2020).

The deliberative approach is also illustrated in several ethics courses for neurotechnology students. Farooqui et al (Farooqui et al. 2021) proposed an ethics course founded on monthly dialogues. It effectively melds reflexive and deliberative approaches by emphasising dialogue and reflection 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, called *Scientific Perspectives and Ethics Commitments Survey (SPECS)* (Tubig and McCusker 2021). This involved mutual reflection in which 'discussants articulate and critically examine the values and beliefs that may be governing their group activities and the discretion of its members'. SPECS draws on individual researchers' interest in engaging with ethical aspects of their work and it recognises 'the discursive nature of ethics and the transformative potential of discourse' (Tubig and McCusker 2021).

4.2 Standalone ethics units versus integration with technical content

The second main finding from our narrative review concerns the contrasting strategies of having ethics material as 'stand-alone' units versus the 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).

4.2.1 Integration approach

Borenstein and Howard emphasised the need to understand ethics as *intertwined* with the technology that future practitioners will design, create, or implement. They argue that digital ethics should thus not be taught as one-off modules, but should instead be *incorporated* in an 'ethics across the curriculum' fashion (Borenstein and Howard 2021). Examples of this integration model of teaching include a 'distributed pedagogy' that incorporates short ethics modules throughout the core curriculum (Grosz et al. 2019), an *in situ* or contextual embedded ethics course model (Carter and Crockett 2019; based on Huff and Martin 1995) and research laboratory-based interventions (Hildt et al. 2019; Plemmons et al. 2020).

4.2.2 Stand-alone approach

Nonetheless, some authors recognised 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.

Authors such as Bates et al (2020) who focus on fairness, accountability, transparency, and ethics in computing[1] (hereinafter, FATE), employ the standalone unit approach. Bates et al (2020), who reported on the process of embedding FATE/Critical Data Studies in the Sheffield Information School's Data Science Masters Programme, taught a 'Data and Society' module which contains themes ranging from "conceptualisations of power, structure, and agency … [to] …Ethical Reasoning" (Bates et al. 2020).

4.2.3 Hybrid approaches

Another author team who dealt with FATE education, Bogina et al. (2021), employ neither a pure integrated nor a pure standalone model but rather a *hybrid* model of ethics teaching. On their hybrid approach, ethics teaching needs to be spread 'across the curriculum' (Borenstein and Howard 2021), but it may for certain reasons (including logistics) need also be deployed in standalone modules. Bogina et al.'s work is in progress; however, some of the proposed projects included 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).

Skirpan et al (2018) propose a short, intensive 'Human Centred Computing' course, 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.

4.3 Moral theory selection and emphasis

The final key finding relates to which moral theories were taught or emphasised in digital ethics. Standard teaching practice in professional ethics, from medicine to business ethics, involves teaching of moral theories (Beauchamp and 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, Judaic, 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. (Others did not refer directly to the study of 'ethics' at all.) 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 and Burton 2017). The three main schools of ethical reasoning referred to in the literature were utilitarianism, deontology, and virtue ethics. Utilitarianism is based on the principle of the 'the greatest good for the greatest number'. 'Deontology' refers to 'law-based ethics', according to which actions are ethical if they conform to a system of rules. One version of deontology, Kantian ethics, notably includes a 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, the dominant western ethical framework is utilitarianism, which (they say) most AI practitioners adopt. But in their view, virtue ethics and deontology have much to offer AI decision making. They suggested that all three forms of ethical reasoning should be developed in ethics in AI courses (Goldsmith and Burton 2017). Likewise, Jones argued that ethical issues in CS are both complex and context-specific, and it is not possible to focus on one aspect, such as deontological duties, without regard for another values, such as consequences of actions (Jones 2016). Chatila and Havens, in the IEEE's *Ethically Aligned Design* (Chatila and Havens 2019), also recommended that several ethical traditions should be integrated into both engineering and science programs and general A/IS awareness. Carter and Crockett's proposed course, described above, also aligned with the recommendation that digital ethics courses should include modules dedicated to virtue ethics, utilitarianism, and deontological ethics (involving moral codes) (Carter and Crockett 2019).

Alternatively, Bezuidenhout and Ratti's 'micro-virtue ethics' approach specifically reflects a virtue ethics framework (Bezuidenhout and Ratti 2021). Bezuidenhout and Ratti referred to the current focus on ethical outcomes and algorithmic design, describing utilitarian approaches that emphasise social

[1] See e.g. ACM https://facctconference.org/ and the National Science Foundation https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf19016, and the University of Illinois https://ischool.illinois.edu/research/areas/FATE#:~:text=FATE%20(fairness%2C%20accountability%2C%20transparency,ethics)%20%7C%20School%20of%20Information%20Sciences

impact, and deontological approaches focusing on development of 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 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). Like Bezuidenhout and Ratti, Hagendorff endorsed a 'microethics' approach. He wrote that traditional ethics guidelines for CS adopt the deontological approach, setting out rules and maxims, and yet 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).

Richards' serious game ethics training tool may also be seen as close to a virtue ethics approach, with its emphasis on psychological development in moral thinking (Richards et al. 2020). Briggle et al developed a research and engineering ethics game founded on virtue ethics (Briggle et al. 2016). They noted the conclusion of the National Research Council (Committee on Ethical and Societal Implications of Advances in Militarily Significant Technologies that are Rapidly Changing and Increasingly Globally Accessible et al. 2014) that the most important mechanism for ethical decision-making is 'good judgment' (2014, in Briggle et al., 2016), and called for 'a return to virtue ethics and its emphasis on practical moral reasoning'.

5 Discussion

Our review of the literature on teaching ethics to IS and related students identified the pedagogical approaches of constructivism/constructionism, social analysis, micro/macro ethics, cognitivism or moral reasoning, and reflective/deliberative approaches. We also found that digital ethics was sometimes taught as a standalone unit, sometimes integrated into technical subjects, and sometimes taught as a hybrid of both. Finally, we found that the main moral theories that were taught or used in teaching were utilitarianism, deontology, and virtue ethics. Notably, several authors advocated for emphasising virtue ethics. In this section, we shall discuss these three sets of findings and identify some possible connections between them, before advancing some recommendations for educators.

The literature addressing digital ethics education refers to a broad variety—and to different combinations—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. On top of that, there is likely to be ongoing disagreements about which pedagogical approach(es) are best.

The relative newness of digital ethics education may also 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 and 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 and Crockett 2019). In such cases, the ethics content is separated from the details of, and the activities connected with, ICT. In contrast, embedding ethics into ICT courses may allow students to better see the relevance of ethics and to connect it more substantially to their interests and future work. Such an approach is aligned with the pedagogical approach of constructivism/constructionism, whereby assimilating new knowledge and understanding builds on existing knowledge and understanding. 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.

Teaching digital ethics not in a single, one-off unit but rather in a more distributed and embedded fashion in ICT curricula may also promote the teaching not just of cognitivist or moral reasoning, but

also the imparting of ethical character traits and skills. The latter form of education is a feature that a virtue ethics approach could readily embrace. 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 and Slote 1997). If that is part of the aim of digital ethics education for ICT students, then a more integrated and extended program of ethics teaching in curricula seems desirable.

Table 1 summarises the findings (i.e., a meta-analysis) of the literature, based on the three broad themes covered in the narrative review.

Criterion	Findings
Pedagogical theories	Constructivism/Constructionism seems to be a dominant pedagogical theory. Social analysis can be limited in cases such as limiting scope to a research lab, for instance. Cognitivist/moral reasoning uses case studies to 'put students in the shoes of a moral decision marker.' Reflexive approaches promote interdisciplinary thinking.
Standalone vs integration	Integrating (or embedding) of ethics seem to be the preferred approach, i.e. a distributed (Grosz et al. 2019) and 'ethics across the curriculum' model (Borenstein and Howard 2021). FEAT studies which involve critical data studies often are useful in a self-contained (standalone) module. However, the hybrid approach will fare better e.g., in capstone projects.
Moral theory selection and emphasis	Typical moral theories taught were utilitarianism, deontology, and virtue ethics. A trend exists towards preferring virtue ethics and development of moral character in students.

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

The apparent benefits of bringing ethics teaching closer to the interests, values, and activities of ICT students does not automatically imply that such teaching should be done by educators whose expertise is in ICT. If it is important to teach moral theory, as some writers in our review 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 both 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 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 and 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 emphasis digital ethics courses place on moral theories of utilitarianism, deontology, and virtue ethics is not surprising, as these are widely recognised and core normative philosophical theories. Nonetheless, one may ask whether there could be some pedagogical benefits of widening the scope to include other moral approaches. This could include, for example, introducing students to Buddhist or Confucian ethics, ethics of care, eco-feminism, and/or other less traditional frameworks (Singer 2013). There are 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. One could imagine that moral approaches which mesh well with students' cultural backgrounds may be especially useful in stimulating learning about digital ethics, and may also align well with the constructivist/constructionist and reflexive/deliberative pedagogical models preferred by some educators.

6 Conclusion

Our narrative literature review identified several pedagogical approaches to teaching digital ethics, differences in the degrees to which ethics components are integrated into ICT courses, and various moral theories in ethics teaching practices. 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. 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. It would also include 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.

7 References

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

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

Barabas, C., Doyle, C., Rubinovitz, J. B., and Dinakar, K. 2020. "Studying up: Reorienting the Study of Algorithmic Fairness around Issues of Power," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, New York, NY, USA: Association for Computing Machinery, January 27, pp. 167–176. (https://doi.org/10.1145/3351095.3372859).

Barnard, A., de Ridder, C., Pretorius, L., and Cohen, E. 2003. "Integrating Computer Ethics into the Computing Curriculum: A Framework for Implementation," in *Proceedings of the 2003 InSITE Conference*, Informing Science Institute. (https://doi.org/10.28945/2619).

Bates, J., Cameron, D., Checco, A., Clough, P., Hopfgartner, F., Mazumdar, S., Sbaffi, L., Stordy, P., and 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?," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, New York, NY, USA: Association for Computing Machinery, January 27, pp. 425–435. (https://doi.org/10.1145/3351095.3372832).

Beauchamp, T. L., and Childress, J. F. 1994. *Principles of Biomedical Ethics*, Oxford University Press.

Ben-Ari, M. 2001. "Constructivism in Computer Science Education," *Journal of Computers in Mathematics and Science Teaching* (20:1), Norfolk, VA: Association for the Advancement of Computing in Education (AACE), pp. 45–73. (https://www.learntechlib.org/p/8505).

Bezuidenhout, L., and 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), pp. 939–953. (https://doi.org/10.1007/s00146-020-01112-w).

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

Borenstein, J., and Howard, A. 2021. "Emerging Challenges in AI and the Need for AI Ethics Education," *AI and Ethics* (1:1), Springer Science and Business Media LLC, pp. 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., and Pluscht, P. 2016. "Research Ethics Education in the STEM Disciplines: The Promises and Challenges of a Gaming Approach," *Science and Engineering Ethics* (22:1), pp. 237–250. (https://doi.org/10.1007/s11948-015-9624-6).

Bynum, T. 2015. "Computer and Information Ethics," *The Stanford Encyclopedia of Philosophy (Summer 2019 Edition)*, (E. N. Zalta, ed.). (https://stanford.library.sydney.edu.au/archives/sum2019/entries/ethics-computer/).

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

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

Chatila, R., and Havens, J. C. 2019. "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems," in *Robotics and Well-Being*, International Series on Intelligent Systems, Control and

Automation: Science and Engineering, Cham: Springer International Publishing, pp. 11–16. (https://doi.org/10.1007/978-3-030-12524-0_2).

Committee on Ethical and Societal Implications of Advances in Militarily Significant Technologies that are Rapidly Changing and Increasingly Globally Accessible, Computer Science and Telecommunications Board, Center for Engineering, Ethics, and Society Advisory Group, Board on Life Sciences, Committee on Science, Technology, and Law, Center for Engineering, Ethics, and Society, National Research Council, and National Academy of Engineering. 2014. *Emerging and Readily Available Technologies and National Security*, (H. S. Lin and W. F. Ballhaus, eds.), Washington, D.C., DC: National Academies Press. (https://doi.org/10.17226/18512).

Crisp, R., and 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*, St. Martin's Publishing Group.

Farooqui, J., Dawod, S., and 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*, , February 2. (http://www.theneuroethicsblog.com/2021/02/towards-neuroethical-ethos-case-study.html, accessed January 24, 2022).

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

Goldsmith, J., and Burton, E. 2017. "Why Teaching Ethics to AI Practitioners Is Important," *Proceedings of the ... AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence* (31:1), Association for the Advancement of Artificial Intelligence (AAAI). (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), Brill, pp. 105–118. (https://doi.org/10.1163/156853599x00090).

Grosz, B. J., Grant, D. G., Vredenburgh, K., Behrends, J., Hu, L., Simmons, A., and Waldo, J. 2019. "Embedded EthiCS," *Communications of the ACM* (62:8), Association for Computing Machinery (ACM), pp. 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), pp. 99–120. (https://doi.org/10.1007/s11023-020-09517-8).

Hämäläinen, J. 2012. "Pedagogic Theories," *Oxford Bibliographies Online Datasets*, Oxford University Press (OUP), March 23. (https://doi.org/10.1093/obo/9780199791231-0015).

Hassan, N. R., Mingers, J., and 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), Informa UK Limited, pp. 263–277. (https://doi.org/10.1080/0960085x.2018.1470776).

Herkert, J. 2004. "Microethics, Macroethics, and Professional Engineering Societies," in *Emerging Technologies and Ethical Issues in Engineering: Papers from a Workshop*, National Academy of Engineering (ed.), Washington, DC: The National Academies Press.

Hildt, E., Laas, K., Miller, C., Taylor, S., and 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), pp. 542–550. (https://doi.org/10.1017/S096318011900046X).

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., and Koedinger, K. R. 2021. "Ethics of AI in Education: Towards a Community-Wide Framework," *International Journal of Artificial Intelligence in Education*, Springer Science and Business Media LLC. (https://doi.org/10.1007/s40593-021-00239-1).

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

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

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

Kohlberg, L. 1984. *The Psychology of Moral Development: The Nature and Validity of Moral Stages*, Harper & Row.

Lewis, A., and Stoyanovich, J. 2021. "Teaching Responsible Data Science: Charting New Pedagogical Territory," *International Journal of Artificial Intelligence in Education*, pp. 1–25. (https://doi.org/10.1007/s40593-021-00241-7).

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

Martin, D. A., Conlon, E., and 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), Springer Science and Business Media LLC, p. 60. (https://doi.org/10.1007/s11948-021-00333-6).

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

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

Moriarty, J. 2021. "Business Ethics," *The Stanford Encyclopedia of Philosophy*, (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University. (https://plato.stanford.edu/archives/fall2021/entries/ethics-business/).

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

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

Newberry, B. 2004. "The Dilemma of Ethics in Engineering Education," *Science and Engineering Ethics* (10:2), Springer Science and Business Media LLC, pp. 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*, Broadway Books.

Papert, S. 1993. *Mindstorms*, London, England: Basic Books.

Papert, S., and Harel, I. 1991. "Situating Constructionism," in *Constructionism*, S. Papert and I. Harel (eds.), Ablex Publishing Corporation.

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

Plemmons, D. K., Baranski, E. N., Harp, K., Lo, D. D., Soderberg, C. K., Errington, T. M., Nosek, B. A., and 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), pp. 1389–1394. (https://doi.org/10.1073/pnas.1917848117).

Rest, J. R., Thoma, S. J., Bebeau, M. J., and Others. 1999. *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*, Psychology Press.

Richards, D., Formosa, P., Ryan, M., Hitchens, M., and McEwan, M. 2020. "A Proposed AI-Enhanced Serious Game for Cybersecurity Ethics Training," in *Proceedings of the 9th Conference of the Australian Institute of Computer Ethics (AiCE 2020)*. (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, EPUB, London, England: Blackwell.

Skirpan, M., Beard, N., Bhaduri, S., Fiesler, C., and Yeh, T. 2018. "Ethics Education in Context," in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, New York, NY, USA: ACM, February 21. (https://doi.org/10.1145/3159450.3159573).

Sprague, A. P., and Diaz-Sprague, R. 2019. "Tying Ethics to Teamwork Training in a Minimodule," in *Proceedings of the 8th Computer Science Education Research Conference*, CSERC '19, New York, NY, USA: Association for Computing Machinery, November 18, pp. 120–122. (https://doi.org/10.1145/3375258.3375275).

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

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

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

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