Preliminary experiences teaching ethical reasoning methods for ICT

Xinyu Mao¹ and Sergio España²

Abstract: Students currently graduating from programmes on information and computing sciences face a challenging professional landscape. Society is increasingly aware of the social and environmental dilemmas and trade-offs that ICT development and usage entail. Ethical reasoning methods for ICT allow applying moral principles and frameworks to evaluate and guide the development and use of technology in a morally responsible and socially beneficial manner. We have been teaching some of these methods to master students for several years. In this paper, we describe how we do that, reflect on our experience, and deliver some advice to improve the learning experience. In the short term, we intend to facilitate the inclusion of ethical reasoning methods for ICT in information and computing science curricula. In the long term, we hope this will increase the industrial adoption of such methods, resulting on beneficial effects in society and the environment.

Keywords: Ethics; Ethical reasoning; Information and communication technology; ICT for sustainability; Teaching

1 Introduction

Information and communication technology (ICT) academics and professionals are increasingly concerned by ethical dilemmas and trade-offs surrounding their projects [EB16]. Overall, ICT is now recognise as having strong ties with sustainability [HA15]. This heightened concern is rooted in an intrinsic motivation and understanding of the economic, social, and environmental impacts of ICT, as well as external factors such as pressures from project stakeholders, societal demands, and governmental requirements; e.g. see [BBP16; CN11; MAC14; Mu08]. Ethical reasoning methods in ICT are designed to aid in analysing these moral and sustainability impacts and trade-offs, supporting decision-making processes related to the design or use of ICT [Es23b]. However, there is a notable lack of evidence showing widespread adoption of these methods in the industry, and even in academia. This could be due, in part, to their limited exposure beyond niche academic circles or scientific communities. For several years, we have been training master’s students in various ethical

¹ Utrecht University, Department of Information and Computing Sciences, Prinsentplein 5, 3584 CC Utrecht, the Netherlands, x.mao@uu.nl
² Utrecht University, Department of Information and Computing Sciences, Prinsentplein 5, 3584 CC Utrecht, the Netherlands, and Valencian Research Institute for Artificial Intelligence, Camino de Vera, s/n 46022 Valencia, Spain, espana@uu.nl

This work is licensed under Creative Commons Attribution 4.0 International License http://creativecommons.org/licenses/by/4.0/. https://doi.org/10.18420/modellierung2024-ws-017
Preliminary experiences teaching ethical reasoning methods for ICT with the goal of increasing their visibility and adoption. This paper shares our experiences and aims to facilitate the integration of these methods into information and computing science curricula, to better prepare future professionals for the ethical challenges in the field of ICT.

The paper is structured as follows. Section 2 defines the concept of ethical reasoning method for ICT and reviews existing methods. Section 3 presents our preliminary experiences teaching such methods to master students; that is, the context of our course, the course components related to ethical reasoning methods, the results delivered by the students, and their opinion on these methods. Section 4 discusses the main findings, limitations, and future work.

2 Ethical reasoning methods for ICT

In earlier research, we have conceptualised the notion of ethical reasoning methods for ICT [Es23b]. First of all, they are methods; this means that they prescribe a specific way of thinking, consisting of directions and rules, structured in a systematic way in activities with corresponding products [Br96]. Their purpose is aiding the user(s) of the method in ethical reasoning, understood as the ability to identify, assess, and develop ethical arguments from a variety of ethical positions [Un]. Lastly, they need to have been created specifically for use in ICT-related situations, or (in cases where the method comes from a different discipline) there must be empirical evidence of their applicability to the ICT domain.

In [Es23b], we also reviewed the literature to find ten ethical reasoning methods that we have characterised and metamodelled. While they all have a similar purpose, they differ greatly in the process and deliverables. For instance, some define and rely on a domain-specific modelling language to support the reasoning or the presentation of the results; it is the case of SoSA Architecture Decision Maps [La19] (in short, SoSA), the Square of Values [Ra19] (SoV), and Strategic Mapping for ICT [Wa21]. Other methods propose tabular structures to guide or structure the reasoning process and its outcomes; the Ethics Canvas [Re18], the Ethical Matrix in Digital Innovation [SS20], the Design Solution Matrix for Value-Sensitive Design [Ja20] (VSD), and MEESTAR [Ma15] belong to this category. Finally, other methods have a more textual format and either result in a structured argumentation, such as the Ethical Framework in Information System Decision Making [Bo12], or scenario-like narratives, such as Techno-Ethical Scenarios [BSS10] and Ethical Dilemma Scenarios [Wr14].

Within our course, we have included course components to train the students in the methods summarised in Table 1. For a more detailed description of the methods, please refer to [Es23c]. Fig. 1, 2 and 3 are models submitted by student teams that serve as exemplars of the typical results of the three ethical reasoning methods under discussion. Fig. 1 displays a model created by a team of students applying the SoSA Architecture Decision Map method to a case related to Tesla electric cars, where several sustainability concerns are interrelated to show cause-effect links. Fig. 2 shows the result of applying the Design Solution Matrix
to the case of Airbnb, discussing in each row the current feature or business aspect (a.k.a. initial solution) that affects negatively a given ethical value (a.k.a. norm), as well as a potential feasible solution (a.k.a. solution after workshop), and an ideal solution used as a reference (a.k.a. regulative idea). Fig. 3 shows the result of applying the Square of Values method again to the case of Airbnb, depicting a value whose attainment or maximisation is desirable (i.e. having a lively neighbourhood), along with an also positive sister value (i.e. quiet neighbourhood), their exaggerations (i.e. the ones at the bottom) and several design scenarios should describe ICT design decisions or features that contribute to the achievement of some of the values. In this deliverable, the students made the mistake of explaining or discussing the values rather than reasoning about the implications of the ICT in the ethical dilemma.

Tab. 1: The three ethical reasoning methods introduced in our course

<table>
<thead>
<tr>
<th>Method</th>
<th>Main tenets</th>
<th>Purpose</th>
<th>Process outline</th>
</tr>
</thead>
<tbody>
<tr>
<td>SoSA Architecture Decision Maps</td>
<td>Software has impact at several levels and sustainability is an essential quality</td>
<td>Make sustainability-driven ICT design decisions, despite trade-offs.</td>
<td>1. Determine sustainability concerns.</td>
</tr>
<tr>
<td>[La19] (SoSA)</td>
<td></td>
<td></td>
<td>2. Determine impact levels of concerns.</td>
</tr>
<tr>
<td>Design Solution Matrix for Value-Sensitive Design</td>
<td>Software should be developed following value- or virtue-sensitive design paradigms</td>
<td>Structuredly comparing different ICT designs decisions, from an ethical point of view.</td>
<td>3. Define relationships among concerns.</td>
</tr>
<tr>
<td>[Ja20] (VSD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square of Values for Business Informatics</td>
<td>Values are guiding principles or Aristotelian virtues that can be pursued or avoided.</td>
<td>Modelling ethical dilemmas and alternative designs using a quadrant that delimits the possible choices.</td>
<td>1. Model initial and sister values.</td>
</tr>
<tr>
<td>[Ra19] (SoV)</td>
<td></td>
<td></td>
<td>2. Model exaggeration values.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Describe current, alternative and negative designs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Discuss options and agree on design.</td>
</tr>
</tbody>
</table>

Abb. 1: Example of a SoSA Architecture Decision Map
Abb. 2: Example of a Matrix of Design Solutions
Abb. 3: Example of a Square of Values model
3 Preliminary teaching experiences

Course context and structure. The Responsible ICT course debuted in September 2020, and had four editions so far. This course is allocated 7.5 ECTS credits, and we expect students to devote 210 hours of work. It is an elective of the Master in Business Informatics of Utrecht University (UU, the Netherlands). Each year, around 70 students will participate in this course. The course focuses on the social and environmental, positive and negative impacts of ICT, and introduces ethical reflections on all the stages of the ICT lifecycle. It covers theories and skills that will allow students to deepen into the interrelation between ICT, society and the natural environment to critically assess the roles ICT plays both at the organisational and systemic levels, its capability to be an ingredient in recipes proposed as solutions to humanity’s outstanding challenges, and also the trade-offs it entails. It has a blended learning approach. The first session is devoted to the motivation for this course, the outline, course rules and organisation. From then onwards, the students are expected to read the course material that we offer through a learning management system before the live sessions. Live sessions are contact time where one or several of the following learning formats are included: core lectures (focused on the core content of the course and delivered by the official lecturers of the course, guest lectures (covering related topics and delivered by colleagues from UU or other Dutch universities), tutorials (intended to train the students on software tools or integrated development environments), or workshops (time used by teams of students to work on a joint assignment). The editions of the course had a minimum of two lecturers involved and a maximum of three. Additionally, a student assistant supports with some tasks, including grading.

Course content. The course content covers an introduction to ethics as a philosophical background (e.g. we discuss deontological, consequentialist and virtue ethics) [Cr98], then we present the Theory of Basic Human Values as a psychological background [Sc17]. We devote a session to codes of conduct and organisational ethics, as reifications of values in either professions (e.g. [As18]) or organisations (e.g.[Mi23]). We then discuss the concepts and taxonomies of values at the economic, societal and environmental levels (e.g. [DS07; Ho11; Sc07], respectively). Throughout the lectures and workshops, we relate these topics with the ICT domain, for instance exemplifying ethical dilemmas or sustainability concerns aggravated or alleviated by ICT interventions.

Training on ethical reasoning methods for ICT. At this point of the course, students are prepared to reason about the impacts and ethical consequences of ICT by themselves. We devote a four-hour session to this purpose (structured as four 45-minute blocks with 15-minute brakes in-between). We first present the concept of ethical reasoning method in a generic way. For each of the three methods mentioned in Table 1, we introduce it, explain its process and deliverables, and present an illustrative example. Then we provide them with a case description that they can use to trigger an individual reflection and apply the ethical reasoning method. When they are done (ca. 20-25 minutes) we show and discuss a reference solution to the exercise. Students can also discuss their solutions and find out whether they have applied the method correctly.
Workshop on ethical reasoning methods for ICT. The next four-hour session is devoted to the Ethical Reasoning Methods Workshop. In this workshop, students are presented with cases such as Tesla, Airbnb, and originally, Corona, which was later replaced by ChatGPT in the most recent year due to its growing relevance. Students are grouped into teams of four to six members. Each team of students is asked to apply the three methods to reason about three distinct cases. The method-case combinations are as balanced as possible. This diminishes the chances that some student teams share information with others. But more importantly, it allows us to create a dataset of method applications that lends itself to be analysed. We have assessed the quality of the deliverables submitted by the teams. During the course, this served the purpose of grading the student performance. As part of this research, this serves the purpose of understanding whether the students are learning properly. Each deliverable has been assessed by either a lecturer of the course or a teaching assistant; random and difficult cases have been discussed to reach consensus and homogenise the assessment process. For each deliverable, we have assessed two aspects. Methodical accuracy refers to whether the method guidelines are correctly applied, including tabular or graphical notations, and whether the results make sense as an analysis of the ethical consequences of the ICT under study. Textual explanation and justification refers to whether the reasoning process and, if applicable, the rationale for the design decisions is clear, and whether the team has documented the deliverable elements with sufficient depth.

Preliminary results of the workshop assignments. We have analysed the quality of the workshop assignments delivered by student teams over a four-year period. Descriptive statistics reveal a consistent trend of high achievement. The Method grade, averaging 7.65 with a standard deviation of 1.13, indicates a strong grasp of technical skills. The Explanation grade, although slightly lower with an average of 7.49 and a broader standard deviation of 1.29, reflects a varied proficiency in articulating methodologies. The Composite grade, with an average of 7.56, encapsulates the overall performance, balancing methodological competence with communicative clarity. These findings suggest a commendable level of skill and understanding among the students, albeit with room for enhancement in the explanation of their methods.

In Fig. 4, we observe that students’ grades using SoV and SoSA have shown a decline over the four-year period, while VSD shows an initial decline followed by a slight increase in the last year. Fig. 5 indicates that the grades for the Tesla case have slightly decreased, while Airbnb shows more variability but generally trends downward. In the year that ChatGPT was introduced as a case study, students achieved a higher average grade in their assignments compared to the initial grades of previous cases. This might suggest that the ChatGPT case was particularly well-suited to the workshop’s educational goals or that it sparked greater interest among that year’s cohort. Across the other cases, there are variations in grades that could reflect the particular challenges each case presents or the varying effectiveness with which different student cohorts apply the ethical reasoning methods. These variations are not indicative of a learning curve but rather the diversity in the ability and perhaps the
Preliminary experiences teaching ethical reasoning methods for ICT

interests of each new group of students.

4 Discussion and conclusions

4.1 Main findings

In the context of pedagogical efficacy, the results shown as Fig. 6 delineates a comparative analysis of average composite grades, segregated by method and case study across four sequential academic years. Notably, the VSD method applied to the AirBnB case in the 2020-2021 academic year exhibited superior performance, positing a potential methodological congruence with the case material. Subsequent academic years elucidate a consistent pattern wherein the SoV method appears to enhance student outcomes, particularly in the Corona case. These findings suggest an intrinsic compatibility between certain methodological approaches and specific case content, which may significantly impact educational outcomes. The data underscores the necessity for an adaptive pedagogical strategy, considering the dynamic interplay between case content and analytical methods, to optimise educational efficacy.

Furthermore, at the conclusion of the course, a survey was conducted to gather student feedback on the ethical reasoning methods taught. Opinions on the SoSA method were split, with 40% of students deeming it the least useful, while another 40% found it beneficial. The VSD method was deemed most useful by 60% of students, indicating high acceptance. Only 20% of students considered SoV to be the least useful, suggesting that while it wasn’t the majority’s first choice, negative views were less common. The divided opinions on SoSA might indicate a need for more real-life examples or alternative explanations to make it understandable to everyone. The high acceptance of the VSD method suggests its
practical nature and ability to reflect real-world ethical dimensions resonated with students, pointing towards the importance of continuing to find and design course content that is both theoretically sound and closely aligned with real-world immediacy. The minimal unpopularity of SoV hints at the necessity of presenting SoV in a manner more directly related to students’ aspirations and anticipated career paths. Bridging the gap might be achieved by integrating SoV into scenario-based learning or linking it with case studies that students are keen on.

![Average Composite grade by Case, Method, and Year](image)

Abb. 6: Bar chart of Average Composite grade by Case, Method, and Year.

### 4.2 Limitations

Teaching ethical reasoning methods for ICT over a longer time span will surely produce deeper insights. So far, we have four years experience and yet we keep polishing the way we deliver the training or structure the workshop assignments. While the recommendations above are likely to remain stable, they might be nuanced or become more precise in the upcoming years.
Ethical reasoning methods still leave space for constructing argumentations that contain fallacies or inaccuracies. Introducing a critical reasoning component in our courses can help students recognise logical fallacies and build strong arguments that properly support their position [Ta17]. Inaccuracies related to the lack of correspondence between premises and the real world are difficult to identify and correct without rigorous literature reviews or empirical research.

Raising ethical awareness within ICT students is still a challenge. Ethical reasoning methods alone might not suffice. Other authors have proposed approaches that could be complementary. For instance, a case-based tool can offer students guidance when considering social, legal or professional implications of the actions and decisions of participants in ethically challenging situations [SRF05]. Taxonomies of principles (e.g. [FC22]) and codes of conduct (e.g. [Go18]) can also play an important role.

4.3 Future work

In light of the findings and observations from the current study, future work regarding teaching ethical reasoning methods for ICT will focus on several key areas to further enhance pedagogical effectiveness. Firstly, reflecting on the student performances in the Ethical Reasoning Method Workshop, it is evident that such workshops are invaluable for students, many of whom may not have previously engaged with this kind of ethical analysis. The experiences of the past few years have highlighted a key insight: students tend to grasp the models more effectively when the case studies are familiar and relevant to their current context. To maximise this understanding, future workshops should incorporate a wider array of contemporary and technologically pertinent case studies. The use of the ChatGPT case study this year serves as a prime example of how integrating current and relatable content can enhance students’ comprehension and application of ethical models.

Additionally, the future direction of these workshops could benefit significantly from incorporating more interactive elements. One approach could be to facilitate group discussions among students who have analysed the same case using the same method. This would allow for deeper, collaborative exploration of the ethical dimensions of the case. Another innovative approach might involve having groups analyse the same case, but each using a different ethical reasoning method. This would enable students to compare and contrast conclusions drawn from varied perspectives, fostering a richer understanding of how diverse ethical frameworks can lead to different insights and conclusions. Such interactive and comparative methods are likely to engage students more actively and deepen their understanding of ethical reasoning in complex, real-world scenarios.

We expect that, eventually, alumni will nudge the organisations they work for to adopt and apply ethical reasoning methods in the ICT projects they participate in. It will be interesting to conduct case studies to understand how these methods are being applied, whether stakeholders perceive their value, whether the resulting insights influence organisational
goals and processes [Ru14] and also to measure the actual impact on the projects and organisations [Es23a]. We hope that ICT will gradually be developed more responsibly and become an ingredient in solutions to humanities most pressing problems.

Acknowledgements

We are thankful to the students of our Responsible ICT course editions, for their active participation and constructive feedback. Sergio España is supported by a María Zambrano grant of the Spanish Ministry of Universities, co-funded by the Next Generation EU European Recovery Plan.

Literatur


Preliminary experiences teaching ethical reasoning methods for ICT


