Responsible Thinking Educating Future Technologists

Christopher Frauenberger, Peter Purgathofer
first.last@tuwien.ac.at
HCI Group, TU Wien
Vienna, Austria

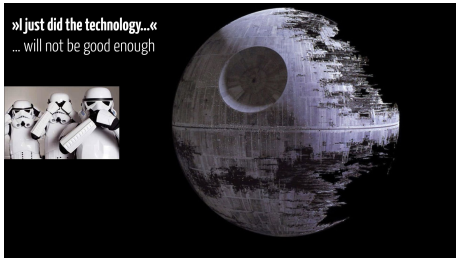
ACM Reference Format:
Christopher Frauenberger, Peter Purgathofer. 2019. Responsible Thinking Educating Future Technologists. In *Proceedings of CHI Conference on Human Factors in Computing Systems (CHI 2019).* ACM, New York, NY, USA, 5 pages.

INTRODUCTION

As a contribution to this workshop on Responisble Innovation, we want to briefly report on a new entry-level course for students of Informatics at TU Wien, Austria[1][2]. It was in part inspired by a recent workshop on "Values in Computing" that we co-organised at CHI'17, where a group of leading experts in the field of human-computer interaction discussed how the field can contribute to establishing a culture of responsible thinking in its research and practice. The outcome was distilled into the *Denver Manifesto*[3] which centres on a call for a shift in the education of future technologists to ensure they can not only write software, but are also critically reflecting on their moral positions and their contribution to society.

In response, we designed a course with the goal to enable students who start out to study computer science, to think about given problems in different ways and from different perspectives. Thus, we aim to plant a seed of critical reflection and equip students with the intellectual tools to see everything they hear later in their studies as part of something bigger. Consequently, we have called this new course "Ways of Thinking in Informatics". We have developed the syllabus[4] over the course of one year and have first taught it to over 960 bachelor students in the fall semester 2017 and again in the fall semester 2018 to a similar number. In what follows, we describe the context in which this course took place, the rationale for its structure and fundamental concept and zoom into the chapter on Responsible Thinking.

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK
2019.



[1] http://informatik.tuwien.ac.at/english

[2] Some descriptions here are also part of a more detailed article on the course to appear in the *Communications of ACM* July 2019

[3] http://www.valuesincomputing.org/the-denver-manifesto/

[4] An english version of the syllabus is available at https://wot.pubpub.org/

With the intimate entanglement of digital technology with humans and their social way of being, computer science has changed. While some of the problems we deal with are (still) well defined and mostly computationally solve-able, many problems are now found to be wicked and ill-defined. People and technologies are now part of an interwoven socio-material web in which humans are not the only actors anymore. This pervasive complexity rises challenges for computer scientists and technologists that go well beyond of what could be addressed by a traditional understanding of engineering the most efficient computational tools. It requires us to rethink what the core competencies of computer scientists of the future need to be. New skills stemming from the social sciences or philosophy need to complement engineering skills to create digital technologies within lived experiences. With it comes a major shift in responsibility. In an article for the New Your Times, Farhad Manjoo argued why 2017 could be seen as a turning point for big technology companies as they "began to grudgingly accept that they have some responsibility to the offline world"[5]. Technologists, whether working in dominating corporations, small start-ups or within academia, can no longer pretend they only solve tech problems, but are required to engage in a moral discourse as most of their products or results are essentially social interventions.

[5]https://www.nytimes.com/2017/12/13/technology/tech-companies-social-responsibility.html

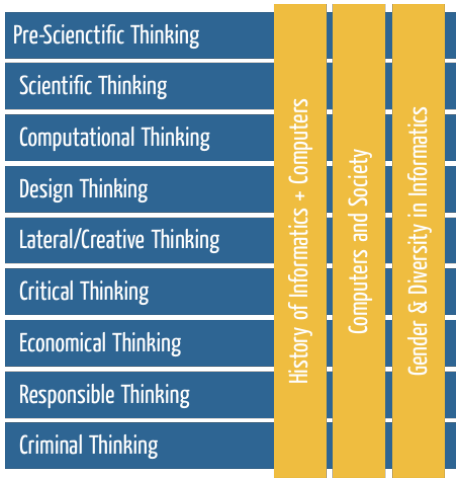
CONTEXT & INSPIRATION

TU Wien is the largest technical university in Austria with close to 30k active students. With over 5k students, the Faculty of Informatics is the second largest in the university with an annual intake of 580 students. We offer 5 Bachelor programmes with foci on visual and human-centred computing, medical informatics, software & information engineering, computer engineering and business informatics. Particularly within the first semesters there is substantial overlap between these in foundational courses covering programming, mathematics, logic, algorithms and software engineering. As the background of our students varies, depending on the type of school they have attended, these foundational courses establish a common ground and basic knowledge to be built on. "Ways of Thinking in Informatics" too became one of the compulsory courses for first-year students of all programmes. This means an annual cohort of 700-1000 students taking this course every fall as additional students from other studies register for our course and some are repeating the course.

The faculty decided to introduce this course as an orientation and to highlight the diversity and potential of computer science in a rapidly changing world. We were inspired by UC Berkley's introductory course "The Beauty and Joy of Computing"[6] which was one of a series of pilot courses that aimed to teach Computer Science's big ideas and the most important computational thinking practices [6]. However, we also felt that the focus on computational thinking maybe is only part of a bigger picture, which is how we set out to build a curriculum that emphasised the diversity of thinking styles. Importantly, we also wanted to follow our university's leitmotiv "Technik für Menschen" (Technology for people) and situate these thinking styles within the context of society.

[6]http://guide.berkeley.edu/search/?P=COMPSCI%2010

[7]See, for example, the collection of ethics related courses started by Casey Fiesler here: https://bit.ly/2IZ2L3O



The need for including aspects of societal impacts and ethics into computer science curricula has been highlighted by the ACM/IEEE CS Curriculum 27 years ago [8]. However, as Goldweger et al. note, the uptake and implementation of such themes in the CS curriculum has been slow. In their survey, they note that while most institutions teach relevant topics, most narrowly focus on ethics or computing history and teach them relatively late and detached from other topics [2]. Considering the recent surge in relevant courses being taught at many univeristies [7], this may be slowly changing in response to the public discourse around societal impacts of digital technology.

However, Ways of Thinking does not aim to be a course about societal aspects of computer science. Rather, we argue computer science is inherently social and no social aspects can be meaningfully separated from computer science. Consequently, we teach very fundamental concepts such as abstraction, cryptography or complexity theory alongside and interwoven with algorithmic bias, the historical role of women in programming or questions of privacy. This resonates with Skirpan et al., who recently argued that ethics education in CS needs to be continuous, in-situ and perspectival [5]. In other words, talking about ethics needs to be part of talking about computer science which in turn requires the ability to take multiple perspectives on computing problems.

WAYS OF THINKING

The key premise of the course is to enable students to apply different lenses on problems of computer science, situated in the world. Each such lens, or way of thinking, frames these problems differently, provides different theoretical foundations and brings with it a unique orientation for questions and methods through which they can be explored. The sequence of "Ways of Thinking" is designed to be semi-historical and somewhat follows a logic increasing complexity. However, we emphasise that these ways of thinking are neither qualitatively ranked, nor mutually exclusive. In fact, the final assignment to students asks them to analyse one given context through multiple lenses with the aim to reveal the benefits of applying different ways of thinking at the same time.

We also acknowledge that the range of thinkings we cover in this course is neither a complete nor the only possible way of categorising the perspectives one could take. In fact, we would encourage educators in the field to develop other ways of organising such a course that teaches perspective-taking as a prerequisite for becoming a reflective practitioner or researcher. Based on initial feedback and on our experience teaching the course, we believe that our choices have been effective in this respect.

In total, we walk through nine different ways of thinking and have included three cross-sectional topics which are interwoven with them. The main chapters are: Pre-scientific Thinking, Scientific Thinking, Computational Thinking, Design Thinking, Creative Thinking, Critical Thinking, Economical Thinking, Responsible Thinking and Criminal Thinking. The cross-sectional topcis are History of Informatics and Computing, Computers and Society, and Gender & Diversity in Informatics.

[8]http://collectivedebate.mit.edu

[9]http://www.nhmrc.gov.au/guidelines/publications/e72

[10]http://moralmachine.mit.edu

[11]https://autonomousweapons.org/slaughterbots/

For the purpose of this submission, we briefly zoom into the chapter on Responsible Thinking.

RESPONSIBLE THINKING

The field increasingly recognises its responsibility within society, as technology becomes a constituent part of today's world. To deflect this responsibility by stating "I was just the scientist" will not be good enough [4]. The chapter starts with common excuses of technologists for avoiding criticism. A prominent example is the notion of "Everyone else is doing it anyway" or in its variation "If I am not doing it, somebody else will". Other common statements include "I couldn't know what it will be used for" (variation: "Guns don't kill people, people kill people"), "It's just always been this way" or "I just did my job" (and "I needed the money"). None of these excuses are necessarily entirely invalid and the moral judgement depends on the precise circumstances. Discussing these common excuses in class serves both, to sensitise students and to illustrate that responsibility is typically complex, situated and distributed.

We then introduce the three main positions from moral philosophy: virtue ethics, deontology and consequentialism, and we demonstrate how ethical judgements shift when these are applied to a controversial social media study in which researchers from Cornell University and Facebook unknowingly involved 689003 users to investigate emotional contagiency [3]. The main learning point here is that ethical judgements are dependent on a moral position, that we often tacitly assume. We then discuss the notion of values and their sources and how they relate to believes. We demonstrate how values are used in data-driven psychological profiling and how this is applied in political campaigning for the purposes of manipulation. We discuss online examples, such as the MIT project *Collective Debate*[8] to demonstrate the implications for public discourse.

We then explore two main areas of moral import for future computer scientists: ethical conduct in science, and responsible research and innovation (RRI). Going through classic examples such as the Milgram experiment and the Tuskagee study, and linking back to the Nuremberg trials, we derive some fundamental principles for research ethics which are still the backbone of most formal ethics frameworks: informed consent, respect, fairness and judging the balance between knowledge gain and risk to participants. We review some existing ethics frameworks (e.g. the Australian National Statement on Ethical Conduct in Human Research[9] and step through two hypothetical research studies to identify common ethics issues. We stress that anticipatory ethics has its limits and that an approved ethics application does not replace an ethical mindset while conducting a study [1].

We then turn to RRI framework implemented by the European Union [7] to discuss central pillars for a reflective and responsible practice: anticipation, reflexifity, inclusion and responsiveness. We apply these principles to the case of predictive policing and point out the many moral dilemmata involved. We then discuss the MIT project "The Moral Machine"[10] and point out the limitations of such a data-driven engineering approach to moral judgements. A video "Slaughterbots"[11] by the

[12]https://www.fastcompany.com/90244860/silicon-valleys-new-playbook-for-tech-worker-led-resistance

Campaign against autonomous weapons serves as chilling example of why responsible innovation is all the more important at the intersection of AI and military applications. Connecting to this point, we discuss the case of Google workers forcing their company out of a related AI project[12] and that the value of programmers as highly skilled and sought-after workforce, makes them increasingly powerful as activists for societal responsibility.

We end with the 6th law of Technology by Kranzberg "Technology is neither good nor bad; nor is it neutral" and discuss a quote by Ryan Calo (Stanford Law School)

> The most interesting aspect of cyberspace flows from its status as an engine of realization: cyberspace widens the range of what we think of as possible. It's often said that where there's a will, there's a way. I don't agree. [...] I'd say the converse is more plausible. Where there's a way, there's a will. If one day a new road for thought yawns into the distance, some adventuring mind will take it.

CONCLUSION

With this short overview of our course and in particular the Responsible Thinking chapter, we hope to be able to contribute to the workshop by making the case for weaving RRI into the education of future technologists.

REFERENCES

[1] Christopher Frauenberger, Marjo Rauhala, and Geraldine Fitzpatrick. 2017. In-Action Ethics. *Interacting with Computers* 29, 2 (March 2017), 220–236. https://doi.org/10.1093/iwc/iww024

[2] Mikey Goldweber, Joyce Currie Little, Gerry Cross, Renzo Davoli, Charles Riedesel, Brian R. von Konsky, and Henry Walker. 2010. Enhancing the Social Issues Components in Our Computing Curriculum: Computing for the Social Good. In *Proceedings of the 2010 ITiCSE Working Group Reports (ITiCSE-WGR '10)*. ACM, New York, NY, USA, 117–133. https://doi.org/10.1145/1971681.1988996

[3] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. [n. d.]. Experimental evidence of massive-scale emotional contagion through social networks. 111, 24 ([n. d.]), 8788–8790. https://doi.org/10.1073/pnas.1320040111

[4] Ann Light, Chris Frauenberger, Jennifer Preece, Paul Strohmeier, and Maria Angela Ferrario. [n. d.]. Special Topic: Taking Action in a Changing World. 25, 1 ([n. d.]), 34–45. https://doi.org/10.1145/3169128

[5] Michael Skirpan, Nathan Beard, Srinjita Bhaduri, Casey Fiesler, and Tom Yeh. 2018. Ethics Education in Context: A Case Study of Novel Ethics Activities for the CS Classroom. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18)*. ACM, New York, NY, USA, 940–945. https://doi.org/10.1145/3159450.3159573

[6] Lawrence Snyder, Tiffany Barnes, Dan Garcia, Jody Paul, and Beth Simon. 2012. The First Five Computer Science Principles Pilots: Summary and Comparisons. *ACM Inroads* 3, 2 (2012). https://homes.cs.washington.edu/~snyder/Snyder-Pilots.pdf

[7] Jack Stilgoe, Richard Owen, and Phil Macnaghten. [n. d.]. Developing a framework for responsible innovation. 42, 9 ([n. d.]), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

[8] Allen B. Tucker (Ed.). 1991. Computing Curricula 1991. *Commun. ACM* 34, 6 (June 1991), 68–84. https://doi.org/10.1145/103701.103710