Teaching Data Ethics

"We're going to ethics the heck out of this"*

Tristan Henderson
School of Computer Science
University of St Andrews
tnhh@st-andrews.ac.uk

ABSTRACT

This paper outlines a new *Data Ethics & Privacy* module that was introduced to computer science students in 2018. The module aims to raise student awareness of current debates in computer science such as bias in artificial intelligence, algorithmic accountability, filter bubbles and data protection, and practical mechanisms for addressing these issues. To do this, the module includes interdisciplinary content from ethics, law and computer science, and also adopts some teaching methods from the law. I describe the format of the module, challenges with module design and approval, some initial comments on the first year's cohort, and plans for future improvements. I believe that the topic is currently important and this discussion might be of interest to other computer science departments considering the introduction of similar content.

CCS CONCEPTS

• **Social and professional topics**; • **Applied computing → Education**;

KEYWORDS

data ethics, education, algorithmic accountability

1 INTRODUCTION

1.1 What is it?

This paper describes a new postgraduate computer science module, *CS5055 Data Ethics & Privacy*, that was introduced at the University of St Andrews in January 2018.

---

*The subtitle is based on Matt Damon's character in *The Martian* (2015) as further paraphrased in *USA vs Jason Alexander Phifer*, case 17-10397. Whereas this character said that "we're just going to have to science the heck out of this," I believe that ethics is needed for to conduct science, and in particular that today's computer scientists need to taught about ethics.
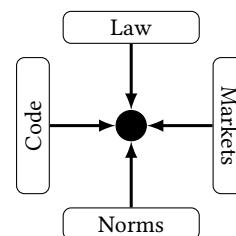

Figure 1: Lessig's "pathetic dot" has behaviour regulated by four possible mechanisms; code is just one of these [15]. I thus cover more than just code in this *Data Ethics* module.

The School of Computer Science at St Andrews is a medium-sized department by UK standards: 38 academic staff teach on a variety of four-year and five-year undergraduate degrees and nine different taught postgraduate degrees, to which approximately 90-100 undergraduate and a similar number of taught postgraduate students are admitted each year. As the School of Computer Science is not accredited by the British Computer Society (BCS), there exists no requirement to teach a professional ethics module. Until 2011 a final-year module *Professional & Social Aspects of Computing* was compulsory for all undergraduate students, but this was dropped due to low student satisfaction and a lack of qualified teaching staff.

2 BACKGROUND

2.1 Who else has done this?

My research has long involved around network measurement and using such measurements to better understand user behaviour. This led to an interest in the privacy and ethical aspects of such measurement, such as obtaining consent [13] or using simulated data as a proxy for real data. [17] I subsequently took a year's research leave to study for a Masters degree in Innovation and Technology Law, with a view to furthering my understanding of law as a mechanism for regulating and enabling ethical behaviour.

On my return to teaching, I wished to both adopt some of the teaching techniques that I had observed in my law degree, as well as some of the material. In particular I believed that many of the topics such as data protection and privacy, regulation of telecommunications networks and the ethics of dealing with new technologies would be of interest to computer science students and indeed useful for them in their subsequent careers. Lessig's "pathetic dot" theory [15] highlights how computer code is just one of the regulatory modalities that can alter behaviour (Figure 1), and so understanding other mechanisms such as the law, norms and ethics is vital for building technologies that do not result in unwanted behaviour.

At the same time there is increasing interest in the ethics of big data and artificial intelligence (AI). As data-driven analysis becomes used to make automated decisions affecting areas ranging from elections [2] and policing [19] to shopping [6] and university admissions [10], much interest has arisen in AI bias [1], and algorithmic accountability [5] and potential mechanisms for regulation [8, 14]. This has led to the formation of many new UK national initiatives such as the Ada Lovelace Institute,[1] the ethics emphasis of the Turing Institute,[2] and the Hall/Pesenti report on AI that proposed conversion courses for non-CS students as "a wider range of graduates will be needed in the AI workforce as it increasingly overlaps with ethics and social sciences."[11] Our existing computer science curriculum, however, did not cover any of these issues.

These personal and broader interests therefore led to the proposal of this new module. At the time of the proposal there were many professional ethics modules in BCS-accredited computer science departments, but these lacked the specific mix of big data, AI and law that I aimed to cover. Cambridge had an *Economics, Law and Ethics* module[3] and Oxford a *Computers in Society* module,[4] but these also had slightly different focuses. De Montfort University has long specialised in ethics and technology through the Centre for Computing and Social Responsibility,[5] and embeds ethics into a first-year undergraduate module[6] that again was slightly different to the Masters-level module that I proposed. I therefore chose to develop my own curriculum.

2.2 Where does it fit?

CS5055 is a "fifth-level" module; in the Scottish context, this means a module for Masters students. We currently offer nine different MSc degrees and CS5055 is available to all of these students. Fourth- and fifth-year (i.e., final-year) undergraduates at St Andrews are also permitted to "dip up" and take this module. Assessment is identical for both undergraduate and Masters students.

In the first year (2017/8), 27 students enrolled on the module. This was considered a healthy number and of the four new fifth-level modules introduced that year, it was only exceeded by the new machine learning module. More specifically, it was far beyond my expectations and I was worried about whether class discussion would scale to such a number. In practice this proved not to be an insurmountable problem.

As this is a relatively new area, there are few textbooks, and those that I did find are aimed at a social science rather than a computer science audience [12, 16]. In the absence of a book, the reading list[7] was instead made up of a number of research papers on each week's topics, with the occasional book chapter where relevant. The academic papers include both computer science research papers and law research papers, which could be challenging for some students, although I chose law papers with significant technical content.

2.2.1 Syllabus. The topics discussed in the module are listed and briefly described as follows:

- What is data ethics? A quick tour of some current controversies in the uses of data, and trying to frame the module.
- What is ethics? An introduction to some commonly used ethical frameworks and ways of approaching arguments.
- What is privacy? A history of privacy and a discussion of some of the many definitions of privacy (based mainly on Solove's taxonomy [20]).
- Relevant aspects of the law. This is mainly a discussion of data protection, placing the legal history into context and then discussing the new EU General Data Protection Regulation (GDPR).
- Consent. Since consent is what many people discuss when they think of technology ethics, we spend a session examining various types of consent and case studies (medical, data protection and so forth).
- Algorithmic governance and accountability. A discussion of various proposals and mechanisms for regulating algorithmic and automated decision-making.
- Ethical machine learning. A discussion, as well as some practical examples and exercises, of specific ways in which classification algorithms can be adapted or evaluated to mitigate potential bias or transparency problems.
- Ethics in practice. This session looks at various ways of enshrining ethical thinking into the software development process, e.g. codes of conduct and impact assessments.
- Research ethics. The final session critiques how university research ethics committees attempt to regulate ethical behaviour. The rationale for discussing this last is both to frame such committees in the wider module-long discussion of how to regulate behaviour, but also because it ties into the project ethics applications that the students are designing around the same time.

The official learning outcomes, as described in our course catalogue,[8] are for students to:

- be able to understand various conceptions of ethics and privacy
- be able to critically analyse research literature at the intersection of computer science, philosophy and the law
- be able to understand the effect of, and the source of, bias or discrimination in a data-intensive system
- understand the need for, and optionally be able to carry out, ethical, social or privacy assessment of data-intensive projects

In short, my unofficially desired outcomes were for students to be more critical about their software development activities during their careers, and to be aware of some of the many tools and levers that are available to them for assessing and designing the projects on which they might work.

2.2.2 Delivery and assessment. The module is delivered as ten weekly two-and-a-half hour seminar sessions. This was the most common model during my law degree and I found that it allowed

---

[1]https://www.adalovelaceinstitute.org/
[2]https://www.turing.ac.uk/research/interest-groups/data-ethics-group
[3]https://www.cl.cam.ac.uk/teaching/1516/EconLaw/materials.html
[4]http://www.cs.ox.ac.uk/teaching/courses/2015-2016/CIS/
[5]http://www.dmu.ac.uk/research/research-faculties-and-institutes/technology/centre-for-computing-and-social-responsibility/ccsr-home.aspx
[6]http://www.dmu.ac.uk/study/courses/undergraduate-courses/computer-science-bsc-degree/computer-science-bsc-degree.aspx
[7]http://resourcelists.st-andrews.ac.uk/modules/cs5055.html

[8]https://info.cs.st-andrews.ac.uk/student-handbook/modules/CS5055.html

students a chance to have deeper discussions than the hour-long lectures that I had previously used. Another observation from being a student is that it was sometimes hard to maintain concentration for over two hours. One professor brought tea, coffee and biscuits and held a break halfway through the session. I adopted this and provided a coffee break and biscuits. This allowed students to recharge and also to ask additional questions or generally chat about the material during the break.

Each week a number of readings were set for the next week's seminar. Students were expected to do these readings so that they could participate in the discussion. I started each seminar with a small number of slides that recapped some of the pertinent points in the reading, but interspersed these with questions that were discussed in groups. There were also other breakout group tasks such as investigating a website's privacy policy or sharing (or not sharing) a social media post with per group members.

I chose a 100% coursework (i.e., no examination) model for assessment. This was also common in my law degree. There are three pieces of coursework:

(1) Essay (worth 40%) - students were given a single essay title that was broad enough for them to pick their choice of case studies and ethical reasoning to argue their point of view.
(2) Peer instruction (worth 20%). For this I used the peer assessment tool PeerWise [4]. While developed in a computer science department and commonly used in the sciences, it was not used in our own School. I therefore first encountered it as a student during my law degree and was surprised by how well it worked in a non-science setting. Students were asked to develop multiple-choice essay-style questions on each week's topics. By having to research the topics and develop detailed questions and answers as well as explanations for why other answers were incorrect, this enabled deep learning and discussion amongst peers.
(3) A data protection impact assessment (DPIA, worth 40%). This instrument is part of the GDPR, and is so new that I used draft guidance from the UK Information Commissioner's Office as the final guidance was not complete.[9]

3 EVALUATION

3.1 Challenges

The first challenge in delivering this module was getting it approved in the first place. While the School of Computer Science was supportive, there were various issues in obtaining Faculty approval. The unconventional (for science) seminar and assessment formats caused particular problems. I originally planned to offer a choice of assessments, allowing students from technical and non-technical backgrounds to take the module (e.g. technical students could do a machine learning practical while less technical students could opt for a design practical), but in the end this had to be discarded.

A second challenge, common to many UK computer science departments in 2018, was the UCU strike over the USS pension scheme. As one of the "worst industrial action at universities in modern times" [21], it was inevitable that this would cause some

disruption to teaching. Given the focus of the module on current events, I chose to incorporate the strike into my teaching, using it as an example of power imbalances and data protection violations. Indeed, the latter led to the writing of a paper on how some universities violated rights by using lecture capture systems to break the strike [7], which may be added to future reading lists.

3.2 Does it work?

As this module has only run for one year, it is difficult to make any objective conclusions about whether it is a success. Instead, I present some anecdotal (quantitative and qualitative) evidence about this first year.

Students performed well, as might be expected for an optional module. Eight students (29.6%) received the equivalent of a first-class grade. There were some excellent coursework submissions, with many of the submitted DPIAs being more detailed than some of the official guidance documents.

Student feedback was good, with an overall module rating of 1.67 (on a 5-point Likert scale where 1 is the best and 5 is the worst). I was also pleased to see that the score for "I directed my own learning beyond the minimum module guidelines" was 1.44, since one of the aims of the module was to encourage additional reading and research. Free-text comments were also positive, with two particular themes being raised the most: students enjoyed the timeliness of the material (even praising the fact that they were working from draft guidance as it was so new), and they enjoyed the biscuits! Negative comments concerned the PeerWise assessment and students believed that the amount of work that was required to develop questions was not reflected in the coursework weighting. I have therefore increased the weighting for this year.

Engagement in the module was high. While I did not accurately record lecture attendance, the number of breakout discussion groups allowed me to estimate that almost all of the class routinely showed up.[10]

Misconduct was low, with one detected case of plagiarism. I was somewhat surprised by this as the topic is relatively new and the coursework required knowledge of current affairs, which would make plagiarism from older material less practical. But the case of plagiarism was a copied essay (as caught by the Turnitin plagiarism detection service) that had nothing to do with the coursework specification. As data ethics modules become more common, however, it may become more challenging to detect such misconduct.

4 FUTURE WORK

4.1 What will we do next?

The main challenge for the second year of the module is dealing with larger numbers. There are currently 44 students enrolled for 2018/9 (i.e., a 63% increase). On the one hand I am very pleased that so many of our students are interested in this topic; it justifies the effort expended in designing the module, and is heartening to see that students want to learn more about ethics, privacy and law.

---

[9]The ICO guidance has since been finalised and a template DPIA is available at https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/how-do-we-carry-out-a-dpia/

[10]I do not routinely record attendance, but from 2018/9 our institutional interpretation of the UKVI visa requirements means that we will have to do so; I may thus be able to monitor numbers over time. This of course raises an issue over whether one should repurpose visa monitoring data for lecture attendance monitoring, and again I may be able to use this as a teaching example for the module.

On the other hand, this creates potential problems in scheduling (finding a room in our institution that provides sufficient lecture space for this many students and at the same time enough room for breakout groups) and in maintaining the discussion format. I have requested resourcing for an assistant to help facilitate the discussions as it would be difficult for me to move around all of the groups on my own.

I plan to adapt the course content. The nature of the content, with an emphasis on current news stories, means that much of the news material needs to be updated each year. Fortunately (or perhaps unfortunately) there has been no shortage of new events, including various Facebook and Google data breaches, etc. The GDPR has also now come into effect, which introduces new possibilities for practical exercises such as asking students to exercise their data protection rights (this was possible under the GDPR's predecessor, but one practical advantage is that the GDPR allows data subjects to exercise rights without paying a fee). I also plan to refine the syllabus by looking at other offerings; there have also been a number of developments in the last year, including various AI ethics courses at other institutions,[11] and an attempt to crowdsource an "Ethical Tech 101" course.[12]

My long-term, and perhaps counter-intuitive, wish is for this module to disappear. It is oft-commented that ethics should be a core part of development culture,[13] and similarly I would like to explore how to consider ethics throughout a computer science curriculum, rather than tacking it on in a single module. Evan Peck provides examples of how ethical thinking can be embedded into a first progamming module [3, 18] and it would be interesting to see if this can also be applied to other modules at other levels. This will take time and effort, though. In the interim I am pleased that this new optional ethics module has shown such interest and engagement from our students, compared to our previous compulsory professional ethics module. The lessons learned here about the timeliness of the content and the various assessments can perhaps be used to explore how to add ethical thinking into our other modules.

5 CONCLUSION

5.1 Why are we telling you this?

I have described a new standalone Masters-level module, *Data Ethics & Privacy*, which is designed to introduce current debates in technological ethics, law and privacy to students. While it is too early to evaluate success just yet, I believe that the module design may be of interest to other educators. I would welcome further discussion of potential mechanisms for introducing ethical thinking into other parts of the computer science curriculum.

REFERENCES

[1] Reuben Binns, Michael Veale, Max Van Kleek, and Nigel Shadbolt. 2017. Like trainer, like bot? Inheritance of bias in algorithmic content moderation. In *Proceedings of the 9th International Conference on Social Informatics (SocInfo 2017)*. Springer, Cham, 405–415. https://doi.org/10.1007/978-3-319-67256-4_32

[2] Carole Cadwalladr. 2017. The great British Brexit robbery: how our democracy was hijacked. *The Observer* (7 May 2017), 12. https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy

[3] Janet Davis, Iris Howley, Darakhshan Mir, Evan M. Peck, and Deborah Tatar. 2019. Make and Take an Ethics Module: Ethics Across the CS Curriculum. In *To Appear: Workshop in ACM SIGCSE 2019*.

[4] Paul Denny, John Hamer, Andrew Luxton-Reilly, and Helen Purchase. 2008. PeerWise: Students Sharing Their Multiple Choice Questions. In *Proceedings of the Fourth International Workshop on Computing Education Research (ICER '08)*. ACM, New York, 51–58. https://doi.org/10.1145/1404520.1404526

[5] Nicholas Diakopoulos. 2015. Algorithmic Accountability: Journalistic investigation of computational power structures. *Digital Journalism* 3, 3 (2015), 398–415. https://doi.org/10.1080/21670811.2014.976411

[6] Charles Duhigg. 2012. Psst, You in Aisle 5. *The New York Times* (19 Feb. 2012), MM30. http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html

[7] Lilian Edwards, Laura Martin, and Tristan Henderson. 2018. Employee Surveillance: The Road to Surveillance is Paved with Good Intentions. In *Amsterdam Privacy Conference*. https://doi.org/10.2139/ssrn.3234382

[8] Lilian Edwards and Michael Veale. 2017. Slave to the Algorithm? Why a 'Right to Explanation' is Probably Not the Remedy You are Looking for. *Duke Law & Technology Review* 16 (Dec. 2017), 18–84. https://scholarship.law.duke.edu/dltr/vol16/iss1/2

[9] Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, and Effy Vayena. 2018. An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines* (Nov. 2018). https://doi.org/10.1007/s11023-018-9482-5

[10] Jay W. Goff and Christopher M. Shaffer. 2014. Big Data's Impact on College Admission Practices and Recruitment Strategies. In *Building a Smarter University: Big Data, Innovation, and Analytics*, Jason E. Lane (Ed.). SUNY Press, Albany, Chapter 4, 93–120.

[11] Wendy Hall and Jérôme Pesenti. 2017. *Growing the artificial intelligence industry in the UK*. Technical Report. Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy. https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk

[12] Gry Hasselbalch and Pernille Tranberg. 2016. *Data ethics: the new competitive advantage*. Publishare, Frederiksberg. http://dataethics.eu/en/book/

[13] Luke Hutton and Tristan Henderson. 2015. "I Didn't Sign Up for This!": Informed Consent in Social Network Research. In *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM)*. AAAI, Palo Alto, 178–187. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10493

[14] Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165, 3 (Feb. 2017), 633–705. https://www.pennlawreview.com/print/?id=553

[15] Lawrence Lessig. 2006. *Code: And Other Laws of Cyberspace, Version 2.0*. Basic Books. http://www.codev2.cc/

[16] Katherine O'Keefe and Daragh O Brien. 2018. *Ethical data and information management: concepts, tools and methods*. Kogan Page, London.

[17] Iain Parris, Fehmi Ben Abdesslem, and Tristan Henderson. 2014. Facebook or Fakebook? The Effects of Simulated Mobile Applications on Simulated Mobile Networks. *Ad Hoc Networks* 12 (Jan. 2014), 35–49. https://doi.org/10.1016/j.adhoc.2012.05.008

[18] Evan Peck. 2018. Ethical Reflection Modules for CS 1. Retrieved 2018-10-13 from https://ethicalcs.github.io/

[19] Jessica Saunders, Priscillia Hunt, and John S. Hollywood. 2016. Predictions put into practice: a quasi-experimental evaluation of Chicago's predictive policing pilot. *Journal of Experimental Criminology* 12, 3 (Sept. 2016), 347–371. https://doi.org/10.1007/s11292-016-9272-0

[20] Daniel J. Solove. 2006. A Taxonomy of Privacy. *University of Pennsylvania Law Review* 154, 3 (Jan. 2006), 477–560. https://doi.org/10.2307/40041279

[21] Nicola Woolcock. 2018. University lecturers to strike as students sit summer exams. *The Times* (9 March 2018). https://www.thetimes.co.uk/article/university-lecturers-to-strike-as-students-sit-summer-exams-2jmmxlsbx

---

[11]A good list of resources can be found at https://www.fast.ai/2018/09/24/ai-ethics-resources/

[12]https://ethical-tech.org/request-for-collaboration/

[13]see e.g. the AI4People project which calls for ethical principles to be "embedded in the default practices of AI" [9].