Engaging Science, Technology, and Society 4 (2018), 1-7 DOI:10.17351/ests2018.221

Engaging STEM Ethics Education

Kelly Ann Joyce[1]
Drexel University

Kendall Darfler[2]
Drexel University

Dalton George[3]
Drexel University

Jason Ludwig[4]
Drexel University

Kristene Unsworth[5]
Stockton University

Abstract

The automation of knowledge via algorithms, code and big data has brought new ethical concerns that computer scientists and engineers are not yet trained to identify or mediate. We present our experience of using original research to develop scenarios to explore how STS scholars can produce materials that facilitate ethics education in computer science, data science,

[1] Kelly Ann Joyce, Email: kaj68@drexel.edu

[2] Kendall Darfler, Email: kdarfler11@hotmail.com

[3] Dalton George, Email: daltongeorge033@gmail.com

[4] Jason Ludwig, Email: jdl84@drexel.edu

[5] Kristene Unsworth, Email: Kristene.Unsworth@stockton.edu

and software engineering. STS scholars are uniquely trained to investigate the societal context of science and technology as well as the meaning STEM researchers attach to their day-to-day work practices. In this project, we use a collaborative, co-constitutive method of doing ethics education that focuses on building an ethical framework based on empirical practices, highlighting two issues in particular: data validity and the relations between data and inequalities. Through data-grounded scenario writing, we demonstrate how STS scholars and other social scientists can apply their expertise to the production of educational materials to spark broad ranging discussions that explore the connections between values, ethics, STEM, politics, and social contexts.

Keywords
big data; ethics; algorithms

Introduction

Along with growing trends to bolster the workforce with the skills needed to code, mine "big data," and develop and run algorithms, there has been rising concern with understanding the values, norms, and decisions that go into these high-tech skills. In the early 2010s, for example, journalists and scholars began to question the decisions that animated Twitter's trending topics as well as the algorithms used in other corporate sites such as Facebook and Google. Concern about automated code popped up in media headlines that asked: "Can code be racist and sexist?" "What happens when machines discriminate?" and described "How white engineers built racist code -- and why it's dangerous for black people." Together, these and other reports called attention to how algorithms and big data may contribute to inequality and structural discrimination. Indeed, there is concern that these socio-technical systems may exacerbate existing inequalities even though such effects may be unintended.

Inspired in part by such reporting, some high profile breaches of ethical behavior, and a desire to train science, technology, engineering and mathematics (STEM) researchers to act professionally and responsibly, policy-makers and educators increasingly advocate ethics education in STEM fields as a way to prevent outcomes that negatively impact groups or society. In the United States, the National Institutes of Health (NIH) introduced its first policy on the responsible conduct of research (RCR) in 1989, requiring all grant awardees to provide evidence that they have been trained in RCR. In 1998 the Accreditation Board for Engineering and

Technology's (ABET) changed its criteria for the accreditation of university programs, Criterion 3f of ABET's Engineering Criteria 2000 (commonly referred to as EC 2000), calls for "an understanding of professional and ethical responsibility" as a student outcome for engineering programs to obtain accreditation

Concerned that such measures were not enough to prevent ethical breaches in STEM fields, the America COMPETES Act of 2007 charged federal agencies such as the National Science Foundation (NSF) and NIH to devise requirements for training undergraduates, graduates and postdocs in the responsible conduct of research. This act resulted in some positive initiatives. In response to this mandate, for example, NSF funded an Online Ethics Center to house training materials and required universities to come up with their own minimum standard of ethics training for students and postdocs. The COMPETES Act also helped create and legitimate funding streams that support the production of STEM ethics education material, including the one that funded our project. Taking a slightly different route, NIH required ethics training of all students and trainees funded through some NIH funding mechanisms, noting that this training cannot be solely completed online and must include a face-to-face component. In doing so, NIH sent a signal that stand-alone online training was unlikely to have meaningful impact on STEM researchers' behaviors, a stance supported by research by Smith-Doerr and others that shows that such training may actually turn off STEM researchers from discussing ethical issues. The COMPETES Act is limited, though, as it only applies to federally funded grants.

To provide minimum ethics training (and thus continue to be eligible for federally funded awards), many US universities subscribed to the Collaborative Institutional Training Initiative (CITI) Program, which was founded in 2000 to create and distribute online ethics training modules. Although online training provides a way for universities to check the ethics training box, universities also tend to offer domain specific courses in ethics (e.g., engineering ethics, ethics and information technology, information ethics). Such courses have historically been under the domain of philosophy. Philosophical approaches to ethics emphasized universal principles such as beneficence and strove to teach STEM students central categories of normative philosophy such as utilitarian, duty-based, and rights-based perspectives. As practitioners aim to find additional ways to bring ethics into the classroom, scholars such as Herkert, Han & Jeong, Lynch & Kline, and Smith-Doerr, have suggested incorporating lessons from Science, Technology and Society (STS) into engineering curricula. STS scholars are uniquely trained to investigate the societal context of science and technology as well as the meaning STEM researchers attach to their day-to-day work practices. Using STS research tools and approaches, scholars can embed in and investigate STEM worlds, providing rich material from which to build scenarios that can be integrated into courses. Building on this call to use STS approaches, we set out to understand the "Ethics of Algorithms" through a multidisciplinary study to understand how these practices are embedded with values, how these values shape decisions that shape the structures that get made, and the resulting societal impact of such decisions.

The Ethics of Algorithms and Big Data

To understand the ethics of algorithms, we observed three labs where teams worked with algorithms and code, and conducted open-ended, in-depth interviews with all members at each site. This research provided insight into computer scientists and engineers' articulated and unarticulated ethics and values as well as ethical dilemmas they face. We transformed these dilemmas into scenarios that challenged readers to think critically about the issue via an accompanying question set.

We want to highlight two example scenarios, composed for the project, that show why ethical engagement of algorithms and software matters. These scenarios are intended to be used in academic settings, yet the ethical dilemmas take place outside of the classroom or funded realm of research as well. The scenarios need to be relevant to "real world" or "on the job" situations where the students are required to take lessons they have learned to a variety of professional settings. It is important that the scenarios speak to the social worlds the students will inhabit as professionals, and do not merely suffice to check off the "ethics education" component of accreditation boards or funding bodies. Together, these scenarios show how the use of algorithms and big data can unintentionally exacerbate existing inequalities and may put individuals under increased surveillance and intervention.

The first issue revolves around how computer scientists and engineers typically approach data about humans. In big data work, computer scientists and engineers often treat data about humans as unbiased. The data are assumed to be transparent, meaningful, representative, and inclusive. STS scholars know, though, that data needs context to have meaning. Take, for example, the case of electronic health records [EHR]. Due to their training, computer scientists and engineers may take the validity of EHR content as a given whereas STS scholars know that the health data in EHRs is partial and may not reflect a person's illness or treatment. Doctors may, for example, give the patient a disease code that they know the insurance company will reimburse, instead of a code that reflects what they think the patient has. Or, they may make mistakes when entering data. Moreover, how should we interpret the health outcomes described in EHRs? Does this information tell us about the effectiveness of treatment (as may be assumed by data scientists)? Or, does it tell us more about the socio-economic class of clients? The scenario "Data Validity as an Ethical Issue" takes up the broader issue of context and health data. Assuming that data about humans is transparent, inclusive and representative ignores the contexts that impact data and make it meaningful.

Second, computer scientists and engineers are not trained to recognize how and when inequalities exist in their products, or to identify the inequalities that may already be embedded in policies, neighborhoods or technologies. This lack of knowledge means that new data systems can unintentionally exacerbate existing inequalities. Even if the inequalities are recognized, computer scientists and engineers are trained to give the paying customer or the boss the tool (algorithm) that retrieves the most information. Yet, there is an important distinction between retrieval of quantity versus quality. "An Algorithm Discriminates" explores how this might happen in a system intended to fine-tune hiring practices. Tasked with developing a system that will identify applicants most likely to stay in the job, the software developer realizes that zip code is a predictor of longevity at the company. Given this, the software developer labels applicants from zip codes that tend to be farther away as "non-recommended candidates." What the software developer may not be aware of is that zip codes are linked to income, and that the zip codes that are being eliminated tend to refer to neighborhoods populated by lower income African-Americans and Latinos. Unintentionally, the software developer excluded lower income applicants of color, and violated the disparate impact principle of the US Civil Rights Act of 1964. The software developer, though, is often not held accountable for these violations. Only after the discrimination has been recognized and litigation has begun does the information retrieved by the algorithm come into question. Is the software designer culpable or did she, in good faith, deliver the product requested by her boss or a client? These scenarios are intended to bring questions such as these to the fore. Ethics are not just important when we are dealing with human subjects review boards, or checking off teaching criteria. It is important to consider ethics and impact even if no one is monitoring the project.

As more and more federal and private monies fund algorithms and big data work, the work of computer science and engineering will increasingly affect a range of domains including city planning, child protective agencies, policing, education, and more. Given the guise of objectivity that surrounds algorithms and the belief that more data will yield better results, local governments and institutions are already having to decide whether to purchase a black box data set that promises to include more information than previously available or to pay someone's salary to manually interpret multiple smaller data sets in a longer amount of time. Guess which one is being chosen?

Connecting STS and STEM Ethics Education

As these examples show, computer scientists and engineers do not have the long history of working with human subjects that STS scholars do. As translators between social sciences and the disciplines we study, STS practitioners are uniquely situated to bring lessons about human subjects and societal impact into discussions of algorithms, code and big data. Our project highlights that when the data in big data is about people, it needs to be understood through a human subjects lens. Although computer scientists and engineers often consider issues related to privacy, the biases of data and the relation between data and inequalities need to be considered as well. Embedding scenarios in STEM courses allows students to explore more complex ways of viewing human subject data, and to be more aware of the challenges that will arise in their work. This is particularly important to do given the large investment in computer science, data science and engineering research, and the deep reach the resulting infrastructure will have in a variety of social arenas.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1338205. We would like to thank Laura Mamo, the ESTS editors Daniel Lee Kleinman and Katie Vann, and the anonymous reviewer for their thoughtful comments and feedback on previous drafts.

Author Biography

Kelly Ann Joyce is a Professor of Sociology and Director of the Center for Science, Technology and Society at Drexel University. As a sociologist of science, medicine and technology, she investigates the social, political and ethical dimensions of knowledge systems and expertise.

Author Biography

Kendall Darfler completed a master's degree in science, technology and society at Drexel University. She has research interests in health and the politics of design. She is currently project director for UCLA's evaluation of the California State Targeted Response to the Opioid Crisis.

Author Biography

Dalton George completed a master's degree in science, technology and society at Drexel University. He is currently pursuing a Ph.D. in Forestry and Environmental Resources at North Carolina State University.

Author Biography

Jason Ludwig completed a master's degree in science, technology and society at Drexel University. He is currently a fellow in the Congress-Bundestag Young Professional Exchange program in Germany.

Author Biography

Kristene Unsworth is a visiting assistant professor in the College of Business at Stockton University in Galloway, New Jersey. Her research examines the relationship between information policy and ethics as they relate to civil society and government.