

Curricular Framework for **Undergraduate and Graduate Education in Ethical & Responsible Computing**

Report by the Ethical & Responsible Computing Working Group, Faculty of Computing & Data Sciences

August 27, 2020
(Updated September 9, 2020)

Executive Summary

This curricular framework for undergraduate and graduate education in ethical & responsible computing (ERC) presents the recommendations of a working group of the Faculty of Computing and Data Sciences (CDS). Based on an extensive literature review and a survey of faculty, this framework defines key terms, presents a roadmap for ERC educational initiatives in the form of six priorities, and identifies steps for implementation of its recommendations. Principles guiding the framework include the following.

- ERC educational initiatives should be responsive to the unpredictability of emerging technologies and new uses of data, helping students learn how to discern morally charged issues in novel situations.
- ERC educational initiatives should support a dual focus on information acquisition and skill cultivation.
- ERC educational initiatives should introduce students to historical and cultural awareness, and support the cultivation of the interpersonal skills needed to navigate the complex social realities they may face in their professional careers.
- ERC educational initiatives should be taught in many different university units and departments, and ERC coursework should be accessible to students from sociology, political science, public health, and other data-centric social sciences.
- ERC educational initiatives should seek to uphold the principle of porosity by engaging experts in a variety of university units.

CDS leadership should use this document to stimulate discussions with deans, chairs, and directors of cognate units and programs; to solicit input on ERC educational initiatives within CDS and across the university; to help build the network of ERC faculty, departments, programs, and units; and to identify faculty representatives for appointment to an implementation committee.

Members of the ERC Working Group:

- Wesley Wildman (Chair), Professor of Philosophy, Theology & Ethics (STH + CDS)
- Alisa Bokulich, Professor of Philosophy (CAS/PH)
- Stacey Dogan, Professor of Law (LAW + CDS)
- Elaine Nsoesie, Assistant Professor of Global Health (SPH + CDS)
- Adam Smith, Professor of Computer Science (CAS/CS + CDS)

1. Preamble

1.1. Ethical and Responsible Computing (ERC) Defined

Ethics is the domain of inquiry concerned with identifying the morally good and bad in all domains of reality, including human behavior, institutions, and technologies. Ethics can involve identifying moral principles that characterize what's good and bad, arguing for maxims to guide moral decisions, and critiquing individual behavior and social practices in light of well-articulated, defensible moral norms. By extension, ethics can connote the application of ethical inquiry to specific domains of professional activity (e.g., medical ethics, bioethics, ecological ethics, business ethics, legal ethics), and articulating responsibilities and practices of ethical analysis suitable for use by professionals working in those domains. Thus, **ethical and responsible computing** refers both to (1) codes of ethics that ought to govern behavior and decision making in professional domains related to computing and data sciences, with associated institutional mandates and requirements; and (2) modes of ethical analysis that can be learned and adaptively applied by professionals at all levels working in those domains.

1.2. CDS and ERC in Undergraduate and Graduate Education

The Faculty of Computing and Data Sciences (CDS) is not a traditional school or department but a unit that arcs across other university units. In regard to Boston University's educational mission, CDS has been organized as a flexible and porous structure capable of exercising both content-related and catalytic roles. As such, in relation to ERC, CDS should aim both (1) to help CDS graduate and undergraduate students acquire ERC expertise and also (2) to serve as a resource for other university units having related aims. The first of these aims calls for CDS to create and sustain ERC educational initiatives in diverse forms, potentially including dedicated ERC classes, ERC modules in other classes, ERC lecture events, ERC certificates, and perhaps ERC degree specializations. All such initiatives should seek to uphold the principle of porosity by engaging experts in a variety of university units. The second aim calls for ease of cross-unit access to ERC initiatives within CDS, and partnerships around ERC initiatives elsewhere in the university.

1.3. Learning Outcomes and Assessment

CDS educational initiatives should foster the personal and professional development of Boston University students. This includes preparing students for future careers in computing and data sciences, fulfilling their professional duties ethically. Given the unpredictability of emerging technologies and new uses of data, students need to learn how to discern morally charged issues in novel situations.

Additionally, ERC education connotes preparing students to deal with interpersonally and socially charged issues, such as gender and racial inequality, or cultural parochialism and paternalism, and to engage diverse stakeholders and fellow professionals respectfully and productively. To this end, ERC educational initiatives should introduce students to historical and cultural ethical awareness, and support the cultivation of the interpersonal skills needed, to navigate the complex social realities they may face in their professional careers.

CDS educational initiatives should assess these envisaged learning outcomes in ways appropriate to each type of educational activity.

2. Survey Findings: Summary

The CDS ERC Planning Committee distributed an information-gathering survey to collect faculty and administrator thoughts about ERC educational initiatives and existing university resources. Nine faculty from a variety of university units responded with detailed narrative answers to a variety of questions about ERC education at Boston University. A detailed summary of survey findings is in Appendix 3. Here we present a brief narrative summary.

The survey demonstrates that a variety of university units offer resources relevant to ERC education. Most take the form of units or lectures in classes primarily about other topics, or lecture events organized by university schools and centers. The (mostly graduate) “Law for Algorithms” class is well regarded and includes significant emphasis on ERC issues. A rich variety of faculty in a dozen different departments or schools were identified as having active interests in teaching and research related to ERC issues, and several BU Centers and Initiatives were mentioned as possible collaborative partners in ERC initiatives.

Our informants told us that the ERC should support a dual focus on information acquisition and skill cultivation. On the information acquisition side, the scope of ERC should be conceived broadly, embracing issues such as data privacy, proper analysis and use of data and algorithms, diligence in testing and validation, bias evaluation, anti-racism literacy, the impact of the internet on democracy and public discourse, the ethical use of social data, and the impact of research on marginalized communities. On the skill-cultivation side, ERC should help produce the capacity to act against self-interest when the common good is at stake, the desire to model ethical behavior so as to create ethical individuals rather than learned unethical individuals, good behavior in relation to women (underrepresented in the industry), anti-racist behavior in relation to racial minorities (some are underrepresented in the industry), the ability to detect ethical implications of computing even when they are nearly hidden, a practical ethics toolkit applicable to industry realities, the capacity to speak up and intervene when colleagues aren’t addressing issues, the insight to anticipate the social impact of the technical capacity to manage and analyze data, and awareness of implicit bias including implicit ethical and political choices.

Respondents told us that ERC should be taught in a lot of different university units and departments, and that ERC coursework should be accessible to students from sociology, political science, public health, and other data-centric social sciences. Some thought ERC education should be incorporated into all courses in computing and data analytics, and all courses where the material could straightforwardly lead to significant damage to society. There was divergence around whether each department should develop their own material independently of others or whether a small number of units should take the lead. There was strong consensus that interdisciplinary faculty should lead such educational efforts, perhaps pairing faculty who understand the technology with humanities and social science faculty who have thought about ethical questions for a long time.

Favored methods for teaching and learning ERC are case studies, projects, and interactive discussion. Some stressed the importance of industry placements to make learning more concrete. One person suggested an annual symposium on ERC trends focusing on new cases studies and areas of activism.

Some respondents, whether in CDS or not, seemed to think that CDS should develop ERC courses and make them available for students from across the university, being careful to remain porous, to avoid

stifling innovation and initiative in other units and departments, and to invite and inspire faculty to engage ERC issues in their teaching. Others thought that CDS is not better positioned than Philosophy or the Law School to take the lead, or that it might be too early for CDS to tackle this subject.

3. Road Map of Priorities

This section lists priorities for CDS educational initiatives in ERC. In part, these priorities reflect timing realities: the still-forming CDS unit is naturally in a position to advance on some issues before others. Also, some priorities require extensive discussion and coordination with university departments, schools, and centers; it takes time to create such discussions, to internalize insights, and to adapt university policies to make collaboration optimally smooth. These priorities should be read with the principles of catalyzing wider university involvement and bi-directional porousness between CDS and other university units firmly in mind. The six priorities, in order of decreasing urgency, are:

1. CDS doctoral requirement in ERC
2. CDS courses in ERC
3. CDS graduate certificate program in ERC
4. CDS undergraduate graduate certificate program in ERC
5. CDS undergraduate minor in ERC
6. CDS masters and PhD subject area specializations in ERC

3.1. CDS doctoral requirement in ERC

The newly approved CDS PhD program includes a requirement in ERC. Appendix D presents the wording in its entirety. The main requirement is that all PhD candidates in CDS must complete a two-semester cohort-based training course (4 credits total) to be offered in CDS covering various aspects of the responsible and ethical conduct of computational and data-driven research. It is anticipated that this class would be offered each year and that CDS PhD students would take it in their first year; that it will be available to students from professional, social science, and humanities disciplines; that it will include modules that cover relevant topics from the Responsible Conduct in Research (RCR) training of students and mentors engaged in NSF/NIH-sponsored research; and that it will feature rounds of presentations by various CDS faculty and guest experts, exposing students to socio-technical tussles that develop at the interface between CDS technologies, social, humanistic, legal, and public policy considerations. The training supplied by this requirement should provide a multifaceted understanding of the promise, consequences, and dilemmas brought about by specific CDS capabilities such as data mining, machine learning, automated decision-making, artificial intelligence, and unbreakable cryptographically secure data storage, processing, and communication.

Recommendations:

- CDS immediately needs to plan a two-semester class offered every year beginning in Fall 2021 to meet this curricular requirement. This sequence could be made available to cognate PhD programs (CS, ECE, Stats, Bioinformatics, etc.). Indeed, these other programs might consider weaving a similar requirement into their doctoral training, which this sequence could help students meet.
- This class should involve a lot of different perspectives from diverse faculty and varied disciplines, using a case study and discussion model.
- Evaluation should focus on projects and participation in discussions.

- It would also be worth considering ways of coordinating this two-semester sequence with the other two-semester CDS PhD first-year cohort-wide seminar. For example, while the CDS introductory seminar is covering data privacy technologies, the ERC seminar could address the ethics of data privacy.
- The class sequence needs to have strong core leadership, present every week, tying everything together.
- The first edition of this class should be mounted by two volunteers from within the CDS faculty, representing (1) technical fields and (2) social sciences, humanities, or law fields. They should design the course around the foregoing principles and the CDS faculty should comment on syllabus drafts.
- Successive editions of the class should accumulate an archive of material that can be handed off to the next team (case studies, assessments, lecture materials, presentations, etc.).

3.2. CDS courses in ERC

CDS will offer courses before long. The shape of an undergraduate curriculum is yet to be finalized but it may still be possible to discuss how ERC education might fit within it. The CDS priority is to support the undergraduate program, but some more advanced classes may also be open to graduate students.

The most critical and demanding element is the creation of ERC-specific classes—a general introductory class as well as classes on a variety of specialized topics—targeted both to CDS undergraduates and to undergraduates elsewhere in Boston University.

Recommendations:

- CDS leadership should endeavor to work with other university units and departments to secure faculty coverage of undergraduate classes dedicated to ERC themes is a question.
- CDS should explore cross-listing existing courses with CDS, as and when they arise within the university (currently there are few dedicated ERC-related classes for undergraduates).
- CDS ERC classes should be considered for the fulfillment of requirements in other degrees or BU Hub.
- CDS should incorporate an ERC unit within relevant CDS-generated undergraduate classes. For example, a CDS machine-learning class might incorporate one or more ERC units dealing with the ethics of machine-learning algorithms.
- CDS should explore the possibility that large multi-section classes, such as the CAS-Philosophy class in ethics, could have a section dedicated to ERC issues and perspectives.

3.3. CDS graduate certificate in ERC

A useful model for a CDS graduate certificate in ERC is the Women's and Gender Studies Graduate Certificate (<http://www.bu.edu/wgs/graduate-certificate/>). Requirements are a single core class, two graduate classes from across the university, a pedagogical workshop, and participation in a symposium series. This certificate depends on the rich network of collaborating units, departments, and faculty across Boston University.

Recommendations:

- CDS should create a graduate certificate in ERC with the following requirements: (1) the core class is the ERC sequence required for the CDS PhD, (2) two ERC-relevant graduate classes drawn

from anywhere in the university, including the “Law for Algorithms” graduate class, (3) participation in a symposium series constructed by requiring attendance at ethics-focused events within the existing Cyber Alliance series and Hariri Workshop series.

- The CDS ERC graduate certificate should be conceived as built atop a rich collaboration of university units, departments, and faculty; this collaborative network should be identified and maintained within CDS and have a web presence to match.
- A CDS graduate certificate in Data Science is planned for a variety of key topics (e.g. Machine Learning and Statistical Analysis). In coming up with such a certificate, ERC should be integrated and possibly also covered as a key topic.

3.4. CDS undergraduate certificate in ERC

An undergraduate certificate in ERC is also relatively low-hanging fruit.

Recommendations:

- CDS should create an undergraduate certificate in ERC with the following requirements: (1) a core class, which would be an undergraduate class focused on ERC, (2) two ERC-relevant undergraduate classes drawn from anywhere in the university, (3) participation in a symposium series constructed by requiring attendance at ethics-focused events within the existing Cyber Alliance series and Hariri Workshop series.
- The CDS ERC undergraduate certificate should be conceived as built atop a rich collaboration of university units, departments, and faculty; this collaborative network should be identified and maintained within CDS and have a web presence to match.
- A CDS undergraduate certificate in Data Science is planned for a variety of key topics (e.g. Machine Learning and Statistical Analysis). In coming up with such a certificate, ERC should be integrated and possibly also covered as a key topic.

3.5. CDS undergraduate minor in ERC

If ERC classes within CDS and related university units become sufficiently coordinated and numerous, and if the undergraduate certificate (3.4 above) is sufficiently popular, it may become possible to offer an undergraduate minor in ERC. Here again, a serviceable model for a CDS undergraduate minor in ERC may be the Women’s and Gender Studies undergraduate minor (<http://www.bu.edu/wgs/minor/>), which draws on a large pool of classes related to Women’s and Gender Studies from ten units across the university, requiring students to complete six classes, including two core classes.

3.6. CDS masters and doctoral specializations in ERC

Masters and doctoral degree specializations in ERC become possibilities as the wealth and variety of classes across the university increases, as there is persuasive evidence of demand through the graduate certificate, and most importantly as there arises a critical mass of faculty wanting to define ERC as an area of their scholarship and research. While ERC-related graduate specializations won’t be feasible quickly, it is an important possibility because of the strong likelihood that this will become an increasingly important domain of research and employment.

4. Next Steps

First, CDS leadership should use this document to stimulate discussions with deans, chairs, and directors of cognate units and programs; to solicit input on ERC educational initiatives within CDS and across the university; to help build the network of ERC faculty, departments, programs, and units; and to identify faculty representatives for appointment to an implementation committee.

Second, CDS leadership and any future ERC Implementation Committee should make use of industry contacts as consultants on industry needs. ERC education should not be in thrall to the interests of industry partners but it can be informed by those interests.

Third, CDS leadership and any future ERC Implementation Committee should reach out to university institutions and networks beyond Boston University for best practices in ERC educational design. Examples are universities already showcasing similar programs (such as UC Berkeley and CMU) and university networks (such as the Academic Data Science Alliance).

Appendix A: Potential Stakeholders

ERC curricular design should consider a variety of stakeholder perspectives to maximize the relevance of Boston University ERC educational initiatives to the workplace realities that students will face after graduation. The following stakeholder perspectives are relevant.

- Research Institutions
- Private Companies
- Legal Bodies and Courts
- Public Policy Makers and Analysts
- Social Ethicists
- Industry Professionals
- Governments and International Organizations
- Media Companies and Creators
- General Populace as Customers, Users, and Research Subjects

Appendix B: Areas of ERC Concern for CDS Educational Initiatives

Based on the list of topics mentioned in the survey (above) and the topics that emerged in a literature review, the following areas of concern should be kept in mind when considering the scope of coverage in any ERC educational initiative.

- Disinformation, Propaganda
- Addiction & Dopamine Economy
- Economic & Asset Inequities
- Machine Ethics
- Implicit Trust & Understanding
- Hateful & Criminal Actors
- Political Disenfranchisement, Underrepresentation
- Declining Transparency/Rising Opacity
- Algorithmic Bias and Injustice

- Diversity, Equity, and Inclusion in Tech
- Diminished Data Privacy/Security
- Technological 'Psyops'/Manipulation (see Surveillance Capitalism)
- Increasing Concentration of Power
- The Surveillance Society - Surveillance State, Surveillance Capitalism (Data Monetization & Behavioral Nudging)
- The Attention Economy
- Growing Machine Autonomy/Declining Human Control and Accountability
- New 'Digital Taylorism'
- Declining Social Trust/Civic Virtue
- Environmental Sustainability of Digital Culture
- Tech Monoculture
- Democracy vs. Technocracy/Tech 'Solutionism'
- Ethical Research and Data Collection
- Labor in the Creation and Development of Technology
- Civic Responsibility
- Lag of Regulations and Policy

Appendix C: Survey Findings

The CDS ERC Planning Committee distributed an information-gathering survey to collect faculty and administrator thoughts about ERC educational initiatives and existing university resources. Nine faculty from a variety of university units responded with detailed narrative answers to a variety of questions about ERC education at Boston University. A detailed summary of survey findings in in Appendix 3. Here we present a brief narrative summary.

Note: The ERC committee is aware that there are experts, university units, and curricular elements related to ERC other than those mentioned by survey participants and listed below.

C.1. Existing BU classes, units of classes, or educational events related to ERC

- ERC-related units in CAS-Philosophy classes in ethics and philosophy of science
- ERC-related module for the BU Spark! Fellowship program
- Ethics and AI lecture in CAS-CS class on machine learning
- CAS-CS and LAW collaboration on ERC-related course "Law for Algorithms" (the CS students in this class are mostly PhD level) and on courses related to cybersecurity, which include significant emphasis on ERC
- CAS-CS PhD topics courses on privacy and fairness
- CAS-CS masters classes on algorithmic interpretability and applied cryptography include ERC components
- CAS-CS undergraduate class on computer security includes an ERC-related module
- ERC-related colloquia through the Center for the Philosophy and History of Science (CAS-Philosophy)
- ERC-related conferences and lectures through CAS-Philosophy and COM-Emerging Media Studies

- SPH course intended for CS students addressing research computing in epidemiology, including research ethics.
- Joint lecture event sponsored by Hariri and “science for the People” on “Technology, Power, and Resistance in the New Gilded Age”

C.2. BU Faculty with established expertise related to ERC

- CAS-Computer Science: Ran Canetti (CDS), Sharon Goldberg, Bryan Plummer, Leonid Reyzin, Kate Saenko (CDS), Adam Smith (CDS), Mayank Varia
- CAS-Philosophy: Alisa Bokulich, Juliet Floyd, Victor Kumar, Russell Powell
- CAS-Physics: Pankaj Mehta (CDS)
- COM-Emerging Media Studies: James Katz
- ENG-Electrical and Computer Engineering: Manuel Egele, Ioannis Pachalidis (CDS), David Starobinski, Ari Trachtenberg
- ENG: Wynter Duncanson
- LAW: Danielle Citron, Stacey Dogan (CDS), Andy Sellars
- MED-Medicine and Biostatistics: Evan Johnson (CDS)
- QST: Nina Mazar
- SPH: Nina Cesare, Elaine Nsoesie (CDS), Prasad Patil
- STH: Wesley Wildman (CDS)
- CDS: Ziba Cranmer

C.3. Possible or Likely Collaborating BU Centers and Institutes

- Biostatistics and Epidemiology Data Analytics Center (BEDAC)
- Diversity & Inclusion Initiative through the Provost’s Office
- Institute for Philosophy and History of Science
- Susilo Institute for Ethics in the Global Economy
- University units such as LAW, CAS-CS, ENG-ECE, and SPH.

C.4. Scope of ERC issues

- “Protection of sensitive data, proper use of data, proper analysis of data and interpretation of results, proper sharing of analysis code and reproducibility of the science.” This could be extended to include diligence in testing and validation, not just for reproducibility but for evaluating bias, impact, etc.
- “I feel that it is impossible to teach ethics without shared axioms (e.g. an unprovable but commonly accepted religious, ethnic, or professional statement of principles). With a diverse community at BU, I don’t see such a common basis ... so the only recourse is to look at history and teach about the intended and unintended consequences of bad behavior. Ultimately, we need some sense of shared purpose and character. We need to build in students the ability to act against their own individual self-interest when there is a greater common interest at stake. This happens from role modeling, not from classroom instruction, and it has to start with the administration and faculty before it can transfer to the students.” Note: “Employers are looking for ethical individuals, but ERC training just provides learned unethical individuals (which, in many ways is worse).”
- “The dangerous role of the internet and social media in undermining democracy and public discourse.”

- “My general perspective is that ‘grad initiation’ or ‘TF training’ classes are likely the only chance we have to train not only our teaching assistants but also future professors who will go on to represent BU in other institutions. I personally wish someone taught me in grad school about implicit bias and why Black people are so underrepresented in STEM, or, as an immigrant to the US, even the basic history of racial injustice in the US. Also about not making misinformed statements and microaggressions. I think that we should be teaching our grad students at least the basics about how to behave around underrepresented minorities, e.g. women, racialized minorities, LGBTQ, etc.”
- “While there are courses on ethics, the issue with ERC is that it is harder to understand the ethical implications of computing. It may often enable unethical practices and actions and part of our responsibility is to be aware and do everything we can to block and mitigate these effects. Educating students about these effects is needed. Idea for class 1: the effect of social media to the public discourse and democracy.”
- “Undergraduates need a set of principles, a checklist, they can take with them into their careers in industry. Graduate students/ PhDs need a more in-depth with toolkit to ensure they are not glossing over these issues in their research. Ideally, their ability to critically identify ERC is a requirement of graduation. Both groups also need training on how to speak up and intervene when their colleagues or affiliated faculty aren’t addressing issues.”
- “For the CS department: classes, and modules for classes, that get students thinking about ethics/responsibility challenges related to computing technology. It would be great to expose them to a variety of disciplinary perspectives (though we have to be careful not to keep it compelling and accessible to them).”
- “I think it depends on the student’s background. If they are primarily trained as computer scientists, they may require foundational courses that address the ethical use of social data and the impact of research on marginalized communities. If they come from an applied/social science background, course work would need to address the capacity of research computing and the way in which it might create or exacerbate inequity.” Note: most employers “are interested in the technical capacity to manage and analyze data, but there is increasing interest in being able to anticipate and unpack the social impact of this work.”
- “That all science involves ethical and political choices, much of them implicit.” Note: employers are looking for “something that makes them look good. I think they have no deep attachment to content—just form.”
- “Diversity, equity, and inclusion as well as anti-racism should be explicitly mentioned and integrated into the theoretical framework [for ERC]. There is an emerging field of antiracism literacy in tech.”

C.5. Which BU units should be teaching ERC?

- CAS-CS, undergraduate and graduate, including graduate initiation classes
- CAS-Mathematics and Statistics
- CAS-Philosophy
- CDS
- COM-Emerging Media Studies
- ENG
- LAW

- ALSO: ERC needs to be integrated in all courses in computing and data analytics, and all courses where the material could straightforwardly lead to significant damage to society.
- ALSO: ERC classes should pair faculty who understand the technology with humanities and social science faculty who have thought about ethical questions for a long time.
- ALSO: ERC coursework should be accessible to students from sociology, political science, epidemiology, and other data-centric social sciences.
- ALSO: If there is no common moral sense, “we are left with only professional ethics, so each department should develop their own material” versus “I would be wary of allocating this to specific departments.”
- ALSO: “I think it’s critical that this course work is developed and curated by interdisciplinary faculty. Not only will this ensure curricula are characterized by flexibility and breadth, it will make coursework seem more accessible to the students who work with affiliated faculty – those who may not think of themselves as ‘data scientists’ but who work on the cutting edge of research computing.”

C.6. Other suggestions for educational initiatives

- “It would be great to have an annual symposium on trends, spotlighting new case studies, areas of activism (ACLU could have a standing slot to evaluate our approach or present new issues), and perhaps a bi-annual review/ update of the curriculum modules.”

C.7. Preferred methods for ERC educational initiatives

- #1 case studies
- #2 projects
- #3 discussion
- #4 legal analysis
- ALSO: industry placements; abstract analysis is not enough
- Note: Some respondents think philosophical ethics is useless and others that it is critical.

C.8. On the CDS role in catalyzing ERC across BU

- Most common comment: CDS should develop the courses and requiring them will set the stage so that others will follow.
- Alternative: “I don’t think the Faculty of Computing and Data Sciences is better positioned than Philosophy or the Law School to take the lead.”
- NOTE: “FCDS should play a role *but* it should absolutely not stifle innovation at the department level. For example, offering resources to instructors developing new course materials under the condition that they share them and get involved in the university-wide process so lessons are shared.”
- ALSO: “I think initially this might be too hard [for CDS]. I would wait 2-3 years to see how CDS develops and revisit the question.”
- ALSO: “This has to be a ‘pull’ model and not a ‘push’. Faculty should be incentivized to include this content in their own course, if they see fit to do so.”

Appendix D: ERC Requirement for the CDS PhD

The newly approved CDS PhD program includes a requirement in ERC. The wording in its entirety is as follows.

The proliferation of computational and data-driven processes into every aspect of our society necessitates that CDS curricula incorporate the ethical and public policy dimensions of computing in general and of data science and AI in particular. This will ensure that CDS graduates understand the importance of responsible use (and dangers from irresponsible abuse) of tools, techniques, and systems they develop – not to mention the very definition of responsible computing in the context [of] ethical, cultural, legal, and social norms[,] all of which are also impacted by computing. This is a two-way street. ... [W]hile it is important that future computational and data-driven technologies be developed by people who understand these norms, it is equally important that future social scientists and humanities thinkers understand what these technologies are capable of doing.

All PhD candidates in CDS must complete a two-semester cohort-based training course (4 credits) to be offered in CDS covering various aspects of the responsible and ethical conduct of computational and data-driven research. Featuring rounds of presentations by various CDS faculty and guest experts, this course will expose students to a number of socio-technical tussles that develop at the interface between CDS technologies, social, humanistic, legal, and public policy considerations. The training is meant to provide a multifaceted understanding of the promise, consequences, and dilemmas brought about by specific CDS capabilities such as data mining, machine learning, automated decision-making, artificial intelligence, and unbreakable cryptographically secure data storage, processing, and communication.

To broaden students' perspectives and their sense of social and professional responsibility for decisions related to technology development and deployment, for each socio-technical tussle students will be organized in groups that engage in debates that leverage literature [on both] the technological and ethical dimensions. While required of all graduate students in CDS, this training will be available to students from other professional, social science, and humanities disciplines, and will include modules that cover relevant topics from the Responsible Conduct in Research (RCR) training of students and mentors engaged in NSF/NIH-sponsored research.

The approved CDS PhD proposal document further notes the ethical analysis furnished in the Stanford Encyclopedia of Philosophy (<https://plato.stanford.edu/entries/ethics-computer/>; 2001, revised 2015), the formation of communities around various aspects of responsible computing (e.g. <https://responsiblecomputing.org/>), curricular initiatives (e.g. <https://ethics.csc.ncsu.edu/>), professional codes of conduct (e.g. from the Association for Computing Machinery (<https://www.acm.org/code-of-ethics>) and the American Statistical Association (<http://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx>)), and Toyama's nuanced Geek Heresy arguments (<https://geekheresy.org>) about social development that leverages, rather than merely relies on, technology to create change.

Appendix E: ERC Educational Initiatives in Other Universities

Peer institutions present two broad approaches to ERC education: departmental models and center models. There are also hybrids of the two approaches. A list of examples of each model follows the in-depth descriptions. We also list examples of concern with ERC awareness, education, and professional training in organizations outside of universities.

E.1. Departmental Models

In a departmental model, faculty affiliated with a single department or a perhaps two departments, are responsible for ERC content. Most courses arise in this way, which is both more *ad hoc* and more sustainable because it is based on faculty interests and perceptions of student training needs.

The most comprehensive review of courses in the ethics of technology is by Casey Fiesler, an associate professor in information science at the University of Colorado, Boulder. Most of the listed courses are generated as departmentally supported initiatives of individual faculty, though some are created by interdisciplinary faculty teams, usually as a dialogue between two disciplines, such as computer science and law. Boston University's own Law and Algorithms course, which usually contains around half computer science students and half law school students, is a successful example of the two-department type. Fiesler invited (and continues to invite) submissions on ethics and technology courses from readers of her Medium blog and from her followers on Twitter. She maintains a public database of such courses, which was most recently updated in spring semester, 2020, and numbers over 260 courses.

(<https://docs.google.com/spreadsheets/d/1jWlrA8jHz5fYAW4h9CkUD8gKS5V98PDJDymRf8d9vKI>). This survey was presented in *Proceedings of the 2020 ACM Technical Symposium on Computer Science Education* (<https://doi.org/10.1145/3328778.3366825>). Fiesler's own analysis reveals a great variety of emphases in ERC courses. Our own analysis (using the archive function of Google docs) shows that many of these courses have been iterated and rebranded with new titles, which is typical in departmental models. The most popular type of course involved policy and legal aspects of technology.

E.2. Center Models

In the center model, faculty may be associated directly with the center based on a fit between faculty interests and the center's mission. Sometimes faculty affiliated with a center are not employed by the university that runs the center. For example, the Berkman Klein Center for Internet and Society at Harvard University (<https://cyber.harvard.edu/>) and the University of Southern California's Center for Artificial Intelligence in Society (<https://cais.usc.edu/>) have affiliated faculty from other universities. These centers usually do not offer degrees but facilitate teaching and research of affiliated faculty and PhD students.

These centers typically involve faculty from a rich variety of disciplines. The Berkman Klein Center was originally founded more narrowly as a "Center for Law and Technology," before broadening to include more disciplines: cultural studies, philosophy, computer science, engineering, and the social sciences. A narrower approach is exemplified by the Markkula Center for Applied Ethics at Santa Clara University, which has a robust focus on the internet and technology and derives its resources primarily from philosophical methods and from faculty related to philosophical ethics in particular. Most centers covering ERC education are aggressively multi-disciplinary.

E.3. Hybrid Models

Hybrid models combine departmental and center approaches to ERC education by means of a course-offering university unit having faculty appointments, like departments, but with a university-wide mission, like centers. ERC education within Boston University's Faculty of Computing and Data Sciences arguably falls into this category.

A parallel example is the University of California at Berkeley's Computing, Data Science, and Society program (<https://data.berkeley.edu/>), which offers undergraduate degrees, with pronounced sensitivity to ERC issues. This "division," as it is called at Berkeley, allows students to specialize in several different topic areas, called "domain emphases." These domain emphases are quite varied, from neuroscience to linguistics to urban science, and often touch on ethics concerns. Selecting a domain emphasis usually means taking courses in another department relevant to data science. Notably, this program does not yet offer graduate degrees, which are instead handled by the School of Information. The division offers data-science modules to faculty in any other school for integration into existing classes. The division also has a coordinated program in Human Contexts and Ethics (HCE), promising (undergraduate) HCE training for all, regardless of disciplinary home (<https://data.berkeley.edu/academics/undergraduate-programs/data-science-offerings/human-contexts-and-ethics-program/hce>).

E.4. University Examples

Departmental Models

- Georgetown, Kennedy Institute of Ethics
- Purdue University, Weldon School of Biomedical Engineering – PRIME Ethics
- Tufts University, Human Robot Interaction Laboratory
- UC Berkeley, Hariri Institute for Computing and Computational Science and Engineering – AI Research Initiative
- University of Colorado, Boulder, College of Media, Communication and Information – Information Science
- UMass Amherst, Electrical and Computer Engineering - Sustainable Computing Lab

Center Models

- De Montfort University Leicester, Centre for Computing and Social Responsibility
- Harvard University, Berkman Klein Center for Internet and Society
- MIT, Internet Policy Research Initiative
- New York University, AI Now Institute
- New York University, The Institute for Human Development and Social Change – Center for Critical Race and Digital Studies
- Princeton University, Center for Information Technology Policy joint with Center for Human Values – Dialogue on AI and Ethics
- Santa Clara University, Markkula Center for Applied Ethics
- Stanford University, Digital Impact
- Stanford University, Institute for Human-Centered AI joint with McCoy Family Center for Ethics in Society
- University of Birmingham, Centre for the Study of Global Ethics
- UC Berkeley, Center for Human-Compatible Artificial Intelligence

- UC Berkely, Simons Institute for the Theory of Computing
- University of Ottawa, Centre for Law, Technology, and Society
- USC, Center for Artificial Intelligence in Society
- University of Toronto, Centre for Ethics – Ethics of AI Lab
- University of Tübingen, International Center for Ethics in the Sciences and Humanities
- University of Warwick, Warwick Business School, Artificial Intelligence Innovation Network
- Utrecht University, Human-Centered Artificial Intelligence

Hybrid Models

- UC Berkeley, Computing, Data Science, and Society
- University of Cambridge, Cambridge Centre for Data-Driven Discovery
- University of Edinburgh, Data, Culture, and Society
- University of Illinois at Urbana-Champaign, National Center for Supercomputing Applications
- University of Padua, AI Society
- University of Sheffield, Digital Society Network

Other Organizations

Non-Profit/NGO

- AI Commons
- AI Ethics Lab
- AI Robotics Ethics Society
- AI4ALL
- AI4People
- Alexander von Humboldt Institute for Internet and Society
- Algorithmic Justice League
- Algorithm Watch
- Allen Institute for AI
- Athena SWAN Charter
- Brookings Institution Artificial Intelligence and Emerging Technology Initiative
- Center for Democracy and Technology
- Center for Human Technology
- Center for Internet and Society
- Center for Mind and Culture
- Center for Technology Ethics
- Code Cooperative
- Computer Ethics Institute
- Confidential Computing Foundation
- Data and Society
- Data Justice Lab
- DataEthics
- DataKind
- Design Justice Network
- Digital Standard

- Electronic Privacy Information Center
- Ethical OS
- Digital Wellness Collective
- Fair AI Certified
- Institute for Ethical AI and Machine Learning
- Linking Artificial Intelligence Principles
- Lifeboat Foundation
- Machine Intelligence Research Institute
- Montreal AI Ethics Institute
- Open Data Institute
- Partnership on AI
- Partnership to Advance Responsible Technology
- People of Color in Tech
- Rathenau Instituut
- Responsible Computer Science Challenge
- PERVADE
- Tech She Can Charter
- Wadhvani AI

Professional

- Association for Computing Machinery
- Association for Information Systems
- Computer Professionals for Social Responsibility
- Data Science Association
- Developers Alliance
- FAT ML
- Institute for Electrical and Electronic Engineers
- National Society for Professional Engineers
- Physicians for Safe Technology