New Perspectives in Science Education
International Conference
NEW PERSPECTIVES in SCIENCE EDUCATION

Balancing Science Education Toward Engaged Citizenship: AI Ethics and Humanities

Chiaki Sekiguchi Bems

Riga Business School / Riga Technical University, Latvia

Abstract

This paper asks what kind of education is needed for science students when the emergent technology of artificial intelligence (AI) is integrated into nearly all human activity and is rapidly affecting the social fabric of our society. With accelerated interest and market demands in fields such as science and mathematics, the value of a university education seems to heavily lie on scientific knowledge and practical training that lead to employability of students. However, this paper argues that traditional educational approaches by isolated disciplines are insufficient for science students to succeed in their professional and personal life. This is because the AI-powered science and technology they develop touch wider fields than their specialization in our increasingly complex, uncertain, digital world. In such a world as this, ethical, responsible AI development, collaboration with people in other fields of expertise, and the willingness to think through uncertainty toward collective societal well-being are of utmost importance. Drawing on my experience of teaching computer science and business students an interdisciplinary AI ethics course called "Humans and Machines," which incorporates the perspectives of the humanities and social sciences, the paper emphasizes the need for a broader approach to education. As Audrey Shafer, MD wrote about the significance of discussing works of art such as Frankenstein with artists, humanists, and social scientists at a medical school, to help us understand what it means to be human, we can broaden science education so our students will have not only scientific but also humanistic understanding of our AI-driven age. The Executive Director of the Modern Language Association of America, Paula M. Krebs also discussed successful pedagogical interdisciplinary collaborations between humanities and computer science and technology. The challenges that the emerging technology has brought—data privacy, algorithmic bias, and ethical and cultural implications, for example—go beyond the technological sphere; therefore, cross-disciplinary re-imagining of a university science education is urgently needed. This collaborative university learning process will, in turn, nurture thoughtful, engaged citizens with a broad, open mindset necessary to face our societal and global challenges.

Keywords: *computer science, humanities, university education, AI ethics, interdisciplinary, engaged citizenship*

1. Introduction: Why Humanities Matter in AI Ethics

Higher education is experiencing turbulent times. In the United States, the decrease in the perceived value of college degrees, the problem of affordability and the "degree gap" between those with college degrees and those without them, the decline of humanities, and the increased demands for artificial intelligence (AI) skills all pressure universities and colleges to reimagine themselves to be relevant to today's society [1, 2, 3, 4]. With accelerated interest and market demands in fields such as computer science, engineering, and mathematics, the value of a university education seems to heavily lie on scientific knowledge and practical training that lead to employability of students. Science education, especially those fields dealing with emergent technologies, then may seem promising for further development with public interest and sometimes with funding support. However, when AI technology is integrated into nearly all human activity and is rapidly affecting the social fabric of our society, what kind of education is needed for ever expanding science students? The answer does not seem entirely clear.

This paper argues that traditional educational approaches by isolated disciplines are insufficient for science students to succeed in their professional and personal life. We live in an increasingly complex, uncertain, digital world, where science students must exhibit ethical, responsible AI development and use, collaborate with people in other expertise, and willingly think through uncertainty toward collective societal well-being. The AI-powered science and technology they develop touch much wider fields than their specialization of particular science, as recent research in social science and philosophy has shown: the concerns for algorithmic bias, data privacy and surveillance, ethical dilemmas, and social and cultural implications for example. Where, how, and with whom could science students learn about

these issues? Some may claim that this could be done in computer science classes when they discuss policies and regulations. AI ethics, however, extends much beyond these discussions and requires a wider understanding of societal structures and cultural values. It forces students to think beyond numbers, theories, and mathematical abstraction to face complex real life consequences technology incurs where no one answer is rarely right. It is therefore best to do this learning, before students go into professional practice, in an interdisciplinary college setting, so they can collaboratively learn with others and test their ideas in a safe classroom environment.

Drawing on my experience of teaching computer science and business students an interdisciplinary AI ethics course called "Humans and Machines," which incorporates the perspectives of the humanities and social science, the paper emphasizes the need for a broad, dynamic approach to education. Kevin C. O'Rourke in discussing today's higher education, argues for the "necessity for a broader education" and humanities' unique attributes to allow us to question our values through critical analysis and think in non-binary, "gray" terms traditional science education may not favor [5]. We are entering an increasingly uncertain, uncharted territory in education as well as in society, so broadness across disciplines and open mindedness are an imperative for success for both students and universities. Elsewhere and earlier, Audrey Shafer, MD wrote about the importance of cross-disciplinary collaboration at a medical school: "Studying and discussing works of art and imagination such as *Frankenstein*, and exchanging ideas and perspectives with those whose expertise lies outside the clinic and laboratory, such as artists, humanists and social scientists, can contribute not just to an awareness of our histories and cultures, but also can help us probe, examine and discover our understanding of what it means to be human" [6]. We should broaden science education so our students will have not only scientific but also humanistic understanding of our AI-driven age and its particular challenges.

Partnership between science and humanities has proven not only successful but also indispensable to discuss AI in a responsible, holistic way. It has been said that humanities help students in STEM majors become well-rounded, and that reading books of literature helps teach empathy and moral compass [7]. It is yet humanists' specific tools that are also relevant in teaching AI ethics. The Executive Director of the Modern Language Association of America, Paula M. Krebs has discussed successful pedagogical interdisciplinary collaborations between humanities and computer science or biology classes on the college campus where she used to teach. "Close reading, cultural competency, attention to language and the ability to build an argument—these are the tools of humanists," Krebs writes, pointing to their usefulness in such fields as environmental humanities, health humanities, and critical technology studies [8]. These tools provide powerful means with which students can discuss ethical dilemmas, ethical, societal, and cultural implications, power dynamics, and dominant narratives all included in AI ethics.

Presumably, it is because of these soft, social skills that humanities also help students' future career progression. According to a report in The Times, the wage gap between humanities and STEM graduates gets smaller slowly in about ten years. Humanities graduates are good at rising to the top. A study from the HIgher Education Statistics Agency found that only 3 percent of science graduates become "managers, directors and senior officials" while the rate is twice as high among those from non-science degrees. Data suggests, according to the Times's Data Editor, humanities graduates are more adaptable and flexible than their STEM counterparts, and they are more likely to move between jobs voluntarily [9]. Providing science students with some of the useful humanities tools, like those mentioned above, only helps their prospects in the competitive professional world.

In what follows, this paper shows an example of an experimental course on AI ethics in higher education. By demonstrating the interdisciplinary, dynamic nature of the course and its progression over the years, the paper aims to call for the re-imagining of a university science education as a cross-disciplinary, collaborative, and adaptive learning. This learning model, with its multifaceted attention to and reflection on society, technology, and our values, in turn, nurtures thoughtful, engaged citizens with a broad, open mindset necessary to face our societal and global challenges.

2. AI Ethics Course: "Humans and Machines"

2.1. Summary

The course "Humans and Machines" is an interdisciplinary, undergraduate AI ethics course that examines broadly the ethical challenges and societal impact of AI technology. It has been offered twice to computer science and business students at Riga Business School/Riga Technical University and continues to evolve into a new form incorporating necessary changes to meet the needs of

students and society. It's a small-size (up to 20 students), discussion-based course with hands-on activities and comprehensive discussions to address major ethical challenges such as AI bias, privacy and surveillance, robots and agency, and human/AI relationships. Students learn from material in AI ethics, social science and literature, and develop their individual projects to pursue a topic of their interest. As part of the semester-long student-centered learning, the course has a community learning component in which students organize and lead a film event/discussion with the school community at a local venue.

2.2. Origin

The course "Humans and Machines" began several years ago to complement what seemed to me to be limited learning experience my computer science and business students were receiving at the time at the university where I worked in Latvia. In a relatively new computer science program housed in a business school, they were avidly learning such subjects as programming, mathematics, data science, data visualization, and AI in the technical sense, but there was no AI or technology ethics course or a course where students learn about societal impact of technology. From my earlier experience of working with them over critical thinking essays on social issues, I knew that there was a strong interest in ethical concerns with AI among the students because many of them had chosen to write about this issue to make an argument. At the time, in Washington D.C., the Facebook whistleblower, a data scientist, Francis Haugen, was claiming that the company prioritized profit over people by delivering harmful products for youth mental health. My students responded strongly to this news. Elsewhere in their research, some computer scientists and social scientists had shown societal harms of algorithms used in AI technology. Thus, despite my background in humanities, more specifically American literature, I proposed and designed an AI ethics course where students could learn about social, cultural, and ethical implications of AI in the hope that my students would engage in responsible AI development and business practices in their professional life.

2.3. Course Objectives

Over the years, the course objectives have shifted from broad to focused on the topic of AI ethics.

In the first year, they were 1) identify major critical questions and concerns regarding AI, 2) have a historical, philosophical understanding of human's relationship to intelligent machines, 3) recognize the non-neutrality of data science and its impact on social power structures, 4) explore the socially responsible ways in which humans can live with advanced technology, 5) think beyond the human-machine framework toward sustainable future, 6) gain insight into humanity and human nature, 7) have collaborative skills to lead and engage in discussions, and 8) develop an active thinking habit to ask critical questions and pursue them.

In the second year, the course's goals were two-fold: to equip students with, first, critical AI literacy and a thorough understanding of the ethical considerations, then, with the competency for ethical use and development of AI technology. To achieve these goals, the course helped students 1) identify important concepts and debates in AI ethics, 2) recognize the societal impact and implications of AI technology, 3) understand different AI types and what technical parts can drive social problems, 4) analyze and/or create ethical AI practice and/or product, 5) know where significant AI ethics research and practices are happening in the world, and 6) develop critical thinking to ask important ethical and technical questions.

In the third year, the course's goals are envisioned to equip students with, first, critical AI literacy; second, competency for ethical use and development of AI technology; third, readiness to work or pursue studies in the field of AI ethics; fourth, crossover learning that students in tech and humanities backgrounds learn from both disciplines. The targeted actions to achieve these goals stay the same as the second year. The shifts in the objectives reflect a growing sense of the need to offer both theoretical and practical knowledge and to integrate science and non-science students in crossover learning.

2.3. Course Structure, Format, Modality

The course structure consists of synchronous 90-min classroom discussion or presentation with mini lectures and asynchronous class preparation with reading, hands-on AI-related activities or online forum discussions on course materials. Separately, students have individual projects. In the second year, these individual projects were divided into two tracks: the humanities track and tech track.

The class structure significantly changed from the first year to the second year because the course objectives became more focused on targeted topics in AI ethics, and the nature of the course shifted from an initial exploratory survey to observations and examinations of the problems. Accordingly, in the first year, the class was conducted in a seminar format where students led the discussion with questions on the assigned material. In addition to the primary textbook we read, Mark Coeckelbergh's *AI Ethics,* students explored various themes they found in such literature as Mary Shelley's *Frankenstein* and Karl Čapek's robot play *R. U. R*. Discussion took the form of Socratic discussion, debate, and what I call "disbate," a debate-inspired student moderated discussion. These various types of discussions as well as the reflections we posted on our online class forum served the purpose of introducing students to the topics of AI ethics while delving deeper into the questions that matter to the students.

In the second year, with clearer focus on specific topics in AI ethics—algorithmic bias, privacy and surveillance, robot and agency, environment etc.---the course took a different format. Instead of student-led seminars, student-led presentations of AI ethics topics helped our class understand each topic more precisely. Not only did students present their gained knowledge, they also showed their response to it in their unique ways. For example, in the discussion of robots, moral agency, and robot rights, a student connected his learning to the AI film *Ex Machina*. In both the first and second years, the class began with a very short presentation of recent AI news or the "industry research" which introduced a noteworthy AI ethics initiative or organization to help them stay informed.

Online forum discussions between classes form an integral part of our class discussions. Prior to each class meeting, students are expected to post thoughts, responses, analyses, or AI experiments. Examples of forum discussions include authentic responses to reading material, bias analysis of AI generated images, and critical analysis of the poetry of Joy Buolamwini, "poet of code". Unlike class discussions where outspoken, extravert students may express their views more often than other students, asynchronous online discussions give students time and space to reflect in the manner that suits them. Writing their thoughts and reading others' thoughts in a shared space reinforce a truly collective class learning experience that classroom discussions alone perhaps cannot achieve.

The modality of instruction is hybrid. Online classes are mixed with occasional in-person classes. Class recording is available for later viewing and reviewing. As Ann Kirshner and Jeffrey Selingo in their recent opinion article "How higher ed can reinvent itself" recommends as one of the reimaginings higher education needs to do to stay relevant, online learning serves various needs and schedules of students and instructors [4]. This works particularly well with an international teaching team like ours.

The course consists of three people in the teaching team. While I am the primary instructor of the course, two exceptional teaching assistants—Eduards Lapiņš and Dāvis Benefelds—in the computer science field help run this interdisciplinary course. This collaboration and co-teaching bring tremendous value to the course that makes a bridge between humanities and science.

2.4. Course Materials

Placing Mark Coecklebergh's *AI Ethics* as a primary textbook, the course has provided various complementary materials from literature and social science. In addition to Shelley's *Frankenstein* and Čapek's *R. U. R.*, students read excerpts from books such as Mark O'Connell's *To be a machine*, The Royal Society's booklet on AI narratives, "Portrayals and perceptions of AI and why they matter," Safiya Umoja Noble's *Algorithms of Oppression*, Cathy O'Neil's *Weapons of math destruction,* Sasha Costanza-Chock's *Design Justice*, Caroline Criado-Perez's *Invisible Women,* and Catherine D'Ignazio and Lauren F. Klein's *Data Feminism*. In the second year, we added excerpts from materials such as Emily Bender et al.'s "On the Dangers of Stochastic Parrots," Joy Buolamwini's *Unmasking AI*, Chris Wiggins and Matthew L. Jones' *How Data Happened*, Nita A. Farahany's *The Battle for Your Brain*. In addition, numerous audio visual materials were also provided.

2.5. Community Learning

At the end of a semester, we in the teaching team, together with our students, organize a community learning evening with AI documentary screening. In the first year, we showed Shalini Kantayya's *Coded Bias* to bring awareness to our school community. Students, faculty, staff, prospective students, and alumni gathered at a local theater to view the film and discuss it with the students in the course. Students led the discussion with the audience, making the occasion a public forum on the important subject of algorithmic bias. In the second year, we showed Ann Shin's documentary about digital cloning, *A.rtificial I.mmortality* to encourage our community to think about our fundamental values on

the matter of life and death in the digital age.

2.6. Final Projects

The final projects are individual research projects students develop throughout the semester. Students choose one AI ethics topic and pursue a research question academically in a medium of their choice. In the second year a more practical, hands-on assignment, ethical algorithm creation, was added as a tech track, making the original assignment a humanities track. Students had an option to choose from the two tracks. In the technical project, students not only made an AI model but also conducted ethical analyses: local and global. The local ethical analysis included areas on privacy and data security, bias, transparency and explainability, and accountability, while global ethical analysis included areas on impact on society, global inequality, scalability, and robustness. The list of the student final projects is found in Fig. 1 below.

Track	Title/Content	Medium
Tech	Algorithm on fetus abnormality detection	Algorithm, ethical analysis
Tech	Algorithm on maze robots	Algorithm, ethical analysis
Tech	Algorithm on real estate pricing in New York	Algorithm, ethical analysis
Tech	Titanic survivability challenge	Algorithm, ethical analysis
Humanities	Deepfakes and Their Impact on Society	Academic essay
Humanities	Silicon Shephard (AI in agriculture)	Booklet
Humanities	Mind over Data (AI in mental health)	Magazine
Humanities	AI Ethics: How Our Future Looks Like (AI robots that dance)	Magazine
Humanities	AI in Education	Interview
Humanities	Possibility of AI Taking over Healthcare Professions	Academic essay
Humanities	AI and Cyber Warfare: The Inevitable Evolution of Military Conflicts	Academic essay
Humanities	Global AI Perspectives: The Influence of Governance and Economic Status on Societal Views of AI and Its Development	Academic essay
Humanities	Effective Accelerationism	Academic essay
Humanities	Sentient AI and Robot Rights	Academic essay
Humanities	AI in Art Industry (copyright issues)	Portfolio

Fig. 1. Student final projects from the “Humans and Machines” course, spring 2024.

3. Assessment and Path Forward

Students’ responses to this experimental course have been positive and encouraging. In course evaluations, the strong points were course organization, instructor feedback, usefulness of resources, critical thinking, and creative thinking, scoring higher than 4 out of 5. The lowest scores were computer literacy, teamwork, and problem solving, scoring about 3.5 out of 5. Course evaluations help design a better course and it was students’ suggestion for including practical, tangible examples that led us to install the two-track system for individual projects in the second year. And this brought forth a great mix of theoretical, practical, critical, and creative outcomes in student projects. Responding to the lower-scoring points on computer literacy, teamwork, and problem solving, the third offering of this course is designed to improve on these areas. In order to increase computer literacy, humanities or

business majors could learn fundamental skills and knowledge of AI from the teaching assistants in workshops, and final projects could be done in small groups that are ideally made up of both humanities and science students. To improve on problem solving, student collaboration with the local AI industry on AI ethics issues in their respective companies might be a possible option, if such cooperation can be arranged. This would help connect classroom learning with real life experiences and could better prepare students for life after university.

Written comments in the evaluations indicate that students enjoyed this not-completely science or math course with abundant discussions on societal issues that have no clear one, right answer. Their comments also show how much they enjoy expressing their opinions critically as well as creatively and learning collectively from classmates as much as from the course materials. In my experience as a humanities instructor at a business school for business and computer science students, I have learned that this type of open, cross-disciplinary discussion course is rare in science education but students greatly appreciate it if included in their university curriculum.

This openness or thirst for learning beyond their science specialization brings us back to the need for broad, dynamic education in our AI-driven world. Many students are open to learning from other disciplines including humanities if there are opportunities. However, throwing them into a mere physical, interdisciplinary juxtaposition may not help students and instructors achieve their goals. We need to teach students how to learn in cross-disciplinary ways by providing science students with humanities's tools and humanities students with science's tools and show them how they could be helpful in their current and future challenges. Guidance is key. To do this, universities need to move toward flexible collaborations across disciplines beyond static, rigid organizational structure. At the same time, the instructors need to adapt to this open structure and be willing to learn beyond their expertise.

The impact of this teaching is increasingly, tangibly felt. There were numerous repeaters at the second AI film community event. More than a few students in the inaugural class of this course (13 students) are pursuing AI ethics by taking another course at a different university, writing a bachelor thesis on it, or teaching it. Of course, the course cannot be the only trigger, given the wider recognition of the subject and a shared sense of urgency in the world.

4. Conclusion

What if Victor Frankenstein lived in our current era and attempted his scientific experiment? Scholars who read the novel as a cautionary tale point out the similarities between Victor with his unchecked ambition and AI developers and researchers who seek knowledge to push the boundaries [10]. Audrey Shafer, MD in her article suggests that safeguards, protocols and institutional approvals are used to prevent "a lone wolf" like Victor from undertaking his "garret experiments", but she acknowledges their limitations and goes on to stress the importance of interdisciplinary talk [6]. For the prevention to be effective, we must build the culture of collaboration and conversation across the disciplines, as well as within the discipline and the classroom. And this collaboration should begin at university's broad science education where students can develop the habit of engaging with others on important societal issues brought by technology—-to fulfill a sense of engaged citizenship. Toward the end of the book *AI Ethics*, Coeckelbergh alarms us against science's tendency toward abstraction and alienation by referring to Hannah Arendt which may blind our attention to "our messy, earthly, embodied, and political life" [11]. This is the place where citizens live with their full embodiment, impacted by technology positively and negatively. Students in science should be willing to recognize this citizens' life with full complexities behind algorithms and theories—-to enact their own citizenship through such engagement.

Acknowledgement
I would like to thank Eduards Lapiņš and Dāvis Benefelds for their collaboration in the AI ethics course "Humans and Machines."

REFERENCES

[1] Fry,R., Braga, D., and Parker, K. "Is college worth it?", Pew Research Center, 23 May, 2024, https://www.pewresearch.org/social-trends/2024/05/23/is-college-worth-it-2/. Accessed 17, March 2025.
[2] Barshay, J. "The number of college students in the humanities drops for the eighth consecutive year", The American Academy of Arts & Sciences, 22 Nov, 2021,

https://www.amacad.org/news/college-graduates-humanities-drops-eighth-consecutive-year. Accessed 17, March 2025.
[3] Ellis, L. “Business schools are going all in on AI”, Wall Street Journal, 3 April, 2024, https://www.wsj.com/tech/ai/generative-ai-mba-business-school-13199631. Accessed 17, March 2025.
[4] Kirshner, A. and Selingo, J. “How higher ed can reinvent itself”, Boston Globe, 17 February, 2025. https://www.bostonglobe.com/2025/02/17/opinion/higher-education-reform/. Accessed 17, March 2025.
[5]O’Rourke, K. C. “What is intelligence anyway? AI and higher education in the 21st century,” Using GenAI in teaching, learning and assessment in Irish universities. Ch. II. Eds. Ana Elena Schalk Quintanar and Pauline Rooney. Trinity College Dublin and University of College, Cork. 2025, para. 13. https://ucclibrary.pressbooks.pub/genai/chapter/__unknown__-9/. Accessed 17, March 2025.
[6] Shafer, A. “Why Frankenstein matters,” *Stanford medicine magazine.* 26 February, 2018, para. 16. Para. 8. https://stanmed.stanford.edu/why-issues-raised-in-frankenstein-still-matter-200-years-later/. Accessed 17, March 2025.
[7] Streitfeld, D. “Sam Bankman-Fried’s wild rise and abrupt crash”, New York Times, 2 November 2023.
[8] Krebs, P.M. “Acknowledging Humanities Expertise”, MLA Newsletter, Winter 2022, 54:4, para. 3.
[9] Calver, T. “Get rich slow with a humanities degree—it’s not all about Stem”, The Times, 18 August, 2024, https://www.thetimes.com/comment/columnists/article/get-rich-slow-the-data-shows-how-humanities-catch-up-with-stem-6n86c7q3t. Accessed 17, March 2025.
[10] Patowary, U. “Artificial Intelligence and Mary Shelley’s Frankenstein: A Comparative Analysis of Creation, Morality and Responsibility”, Integrated Journal for Research in Arts and Humanities, 3:4, July 2023, pp. 121-127.
[11] Coeckelbergh, M. *AI Ethics*. MIT Press, 2020, p. 195.