Civic AI Education: Developing a Deliberative Framework

Weiyu Zhang, Department of Communications and New Media, National University of Singapore, Singapore. Email: weiyu.zhang@nus.edu.sg

Artificial Intelligence offers tremendous opportunities but raises concerns over harming civic values. AI Education has focused on advancing technologies but lagged in embedding civic competencies in the learning. This provocation aims to propose a framework of Civic AI Education that highlights deliberation, by centering on the informed and reciprocal participation of diverse civic actors from the early stage of AI design to the late stage of AI monitoring. This provocation tries to demonstrate that in order to make AI ethical, we have to first educate human actors to be civic.

CCS CONCEPTS • •Human-centered computing~Collaborative and social computing~Empirical studies in collaborative and social computing •Human-centered computing~Human computer interaction (HCI)~Interaction paradigms~Web-based interaction •*Human-centered computing~Collaborative and social computing~Collaborative and social computing systems and tools*

Additional Keywords and Phrases: Artificial Intelligence, citizens, civic engagement, civic tech, deliberation, education

ACM Reference format:

Weiyu Zhang. 2022. Civic AI education: Developing a deliberative framewrok. In 4th Annual Symposium on HCI Education (EduCHI'22), April 30–May 1, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 5 pages.

1 INTRODUCTION

Artificial Intelligence (AI) offers tremendous opportunities but raises concerns over harming civic values. There is an increasing level of public worry over the unethical consequences emerging out of AI development. From low-stake to high-stake situations, AI has become important decision makers who affect human life in all aspects. In the domain of education, AI was used to predict UK high school students' university entry exam grades[1], when Covid-19 prevented the exam from happening. The algorithm is a complex considering factors such as individual students and the schools they come from. The algorithm was picked over human (teacher) prediction to avoid problems such as grade inflation and variation in accuracy of predictions from individual teachers or schools. When the predicted grades were announced, students and parents took to the street to protest against them. The public outcry was so loud that the AI predictions were defunct only two days after

EduCHI'22, April 30–May 1, 2022, New Orleans, LA, USA.

[1] https://www.technologyreview.com/2020/08/20/1007502/uk-exam-algorithm-cant-fix-broken-system/

their announcement. If it is problematic that AI decides on student grades, in the domain of crime and justice, AI decides on who stays in jail. There are multiple projects trying to use AI to predict recidivism, and COMPAS is one of them. A study [4] has shown that "COMPAS is no more accurate or fair than predictions made by people with little or no criminal justice expertise." These incidents present a myriad of civic concerns over AI, potentially violating multiple moral values including fairness and care. The problem isn't just about the accuracy of AI but the process of generating the AI and motioning the usage of the AI. Social or human control over AI advancement has to be kept, in a continuous manner not at singular time points. This provocation tries to address the pressing concerns from the perspective of CHI education, focusing on the deliberative framework.

This provocation makes three bold proposals:

- Civic education has to be an essential part of AI education at all levels, including ordinary users, policy makers, corporate players, and technological developers.
- Civic AI education needs to go beyond ethical guidelines and toolkits to take up the deliberative framework, focusing on a dynamic procedure of informed and reciprocal participation from diverse civic actors.
- AI can be educated to be civic through educating the human actors related to AI development and implementation.

2 CIVIC EDUCATION AND AI

2.1 Civic Education and Its Challenges

A key goal of education is to prepare individuals for effective participation in societies. This area of learning is commonly known as citizenship education or civic education. Depending on the social and political contexts of different societies, other related terms used for this area of learning include values education, moral education or character education. Traditionally, citizenship education is largely confined within the physical boundaries of nation-states, and places strong emphasis on the formal understanding of their country's political, legal and economic systems, their citizenship rights and responsibilities, and how their government works. With globalization and technological advancement, such traditional conceptions of civic education is inadequate in preparing people for the complexities of societies. It is now necessary, and crucial, to broaden the traditional conception of civic education to involve the development of competencies for active participation in, including critical understandings and deliberations of the complexities inherent in exercising the rights and responsibilities of community life.

In our context, technological advancement, especially information and communication technologies such as AI, poses a major challenge to civic education. The AI incidents mentioned above fully demonstrate that not only designers but also users of AI are not educated enough to detect the civic hazards AI technologies may bring to the human society.

2.2 AI Ethics and Its Challenges

The challenges we see in AI ethics are illustrations of the challenges we find in civic education. Civic values are easy to say but hard to practice; easy to teach but hard to learn. Moreover, civic values can be understood in diverse ways among different social groups. Embedding civic values in an emerging technology like AI is like

shooting a moving target – when we are not even sure what exactly AI can do, how can we evaluate them against civic values? The current debate on AI ethics is one effort to help us define the target.

The current frameworks to build AI ethics can be categorized into three groups: The first group focuses on ethical AI guidelines, which are normative principles and recommendations that appear as checkbox lists for designers to consider. However, such universal rules, just like civic values, are difficult to be practiced in individual situations that are far from ideal. Reading ethical guidelines has no influence on designers, who have to grapple with personal disposition, organizational priority, and operational constraints. We need to "build tangible bridges between abstract values and technical implementations." [6] The second group refers to such tangible bridges, mostly in the form of toolkits. However, open source fairness toolkits were found to pose a steep learning curve even for technologists, lack mitigation strategies, and are hard to adapt to individual situations [9]. The third group are activity-based, using a workbook, a game, or a set of value cards to make designers reflect on their design decisions and play the role of other stakeholders [2]. Participatory design or co-design activities [11] were also proposed to educate the designers to be ethical. As we can see, all three groups of AI ethics frameworks still solely focus on AI technologists, without effective means to include the multiple civic actors who might bring diverse civic perspectives to AI ethics.

3 A DELIBERATIVE FRAMEWORK TO AI EDUCATION

The deliberative framework of civic education is based on the deliberative ideal of citizenship [7]. It values the opportunities for people to develop, and exercise their rights and responsibilities to participate actively, and in an informed and reciprocal manner, on effecting changes on a variety of social issues. This perspective takes a 'maximal' conception of citizenship that acknowledges the inevitable presence of 'politics', where people use their share in power to engage in activities to conciliate differing interests [3]. This requires people to have the knowledge, skills, dispositions and importantly, the institutional power to think critically about issues to consider how they can work within, change, or challenge the "given unit of rule" within which they can deliberate and act on societal issues [8; 12]. As Crick [3] elaborates, this perspective of deliberative citizenship is ideally premised on a fundamental commitment to the consideration of one's welfare in proportion to the survival of the whole community. In many democratic societies, deliberations and the consequent decision-making process are rooted in democratic values such as equity and social justice. Even in not so democratic societies, deliberations function as consensus making tools for citizens to learn about each other's views and to contribute to the often centralized policy-making.

A deliberative framework of civic education has at least two distinctive features. First, the deliberative framework is future-looking. The notion of citizenship is extended beyond just the political issue of rights and responsibilities, to include having the competencies to take on the "ethical challenge to narrow the gap between the promise and the reality of a global democracy" [5]. This means civic education has to perform the dual function of simultaneously socialising people to existing forms of participation, perspectives and understanding of the society, and develop competencies to help them exercise agency in a highly dynamic, controversial, and diverse future. Second, the deliberative framework emphasizes reflection and perspective-taking. In approaching deliberations, citizens need competencies to expand their self-knowledge and critical agency, both of which are developed in relation to the others. This means valuing the importance of a culture of questioning oneself and taking actions to transform oneself. Additionally, citizens need competencies to acknowledge and

understand the influence of human relations on shaping one's perspectives on social issues. Hence, they are made aware of the historically and socially built inequalities and injustices.

In practice, deliberation can be intrapersonal within one's own mind, using nudges and persuasions to elicit reflection on one's own views and taking other people's perspectives [13]. Most deliberation practices, however, emphasize the interpersonal communication among deliberators [10], including experts vs. lay people, people with diverse demographics, people with conflicting opinions, and people who occupy different social positions (e.g., policy makers vs. ordinary citizens). Through being exposed to diverse views, justifying one's own views in front of others, and offering and receiving arguments and empathy, such interpersonal deliberation is believed to be able to resolve disagreements and build consensus. There have been proposals to use the deliberative framework for responsible AI innovation, arguing that "deliberation serves an important function for both epistemic as well as moral justification in AI by highlighting particular tensions between common and ideal requirements for public reasoning on the one hand and particular challenges related to AI explainability and accountability on the other." [1]

3.1 Deliberation as a Civic AI Education Tool

Specifically to AI education, a deliberative framework requires that citizens need competencies to be multi-literate to access and understand information related to emerging AI technologies and the social systems they are embedded in. To be able to meaningfully deliberate on these issues, they must be able to engage, learn from, understand, be tolerant of and be responsible to matters of difference and otherness. This means that deliberation needs to be underpinned by values of mutual respect, communicative rationality, and ethical responsibility. Lastly, the institutional power needs to recognize the importance of deliberation outputs, granting them the binding influence to shape AI development and deployment.

Below I briefly describe an initiative conducted to develop university ethics around analytics and AI in education. University of Technology Sydney started to implement a set of data analytics and AI tools in its education system, such as simulation wards in the medical school and AI-assisted feedback. The university recruited a mini-public, representative of the educators and students population. Experts were called to provide information and answer questions. External professional facilitators moderated the deliberation sessions. Examples and dilemmas were given to the mini-public to collectively identify, discuss, and prioritize key ethical principles. After presenting the deliberation outputs to the university leaders, the principles were incorporated into UTS' Ed Tech Ethics policy to be launched in 2022.

3.2 The Limitations of Participatory Design and Co-Design

The deliberation initiative sounds familiar to the designers. In the HCI field, there are various design models that involve non-designers' participation. A rather simplistic summary is to differentiate them according to the role that participation from various civic actors plays in the design process. In user-centered and iterative design models, users may be extensively surveyed and interviewed about their needs and designers may iterate this user study for a few rounds, users are not the ones who decide on the design solutions. In participatory design and co-design, users are considered equally as experts in providing design insights and involved in the design process from the beginning to the end. Other design models such as design fiction or speculative design try to address the challenge of predicting the future. Value-sensitive or value-centered design (Menon and Zhang) prioritizes values over user needs.

However, a blind spot in participatory design or co-design is inherent to their emphasis on design, which pays attention only to the early stage of developing a technology. AI challenges do not stop after they are designed. An algorithm may decide on how data are generated, collected, sorted, and fed back into design. The same data can be re-purposed and re-used by another algorithm or other human decision makers. Taking Covid-19 contact tracing technology as one example. The Singapore government designed a privacy-preserving mobile app that relies on Bluetooth signals instead of location data to identify close contacts. The design intention here seems to be ethical enough. However, later it was found out that the police re-used such data in criminal cases, an intention not revealed to the public when contact tracing was made compulsory. We can imagine that even if co-design was implemented at the design stage, the co-designers would not be able to envision this re-use. Therefore, we need to build the deliberation approach into not only the design stage but also the later stage of monitoring.

4 SUMMARY

Developing a deliberative framework for civic AI education focuses on the informed and reciprocal participation of diverse civic actors from the early stage of AI design to the late stage of AI monitoring. Depending on the scope of one particular AI technology, relevant stakeholders should be included, and moreover, the stakeholders are equipped by not only technological information but also civic competencies to deliberate on the details of AI, from design to usage to re-purposing to monitoring. Civic AI education should not be limited to the AI developers but broadened to include ordinary users, policy makers, and corporate players.

ACKNOWLEDGMENTS

This work was supported by National University of Singapore (NUS) under Grant R-124-000-112-646. The author wants to thank the funder National University of Singapore for their generous support. Views expressed are those of the author alone and do not necessarily reflect opinions of the sponsoring agencies. The author wants to extend sincere appreciation to collaborators Yeow-Tong Chia and Jia Ying Neoh as well as participants at the 2022 NUS Symposium on AI Education for their inspirations.

REFERENCES

[1] Buhmann, Alexander, and Christian Fieseler. "Towards a deliberative framework for responsible innovation in artificial intelligence." Technology in Society 64 (2021): 101475.

[2] Cha, Inha, and Lim, Youn-Kyung. “Challenges in Devising Resources for Ethics: What ShouldWe Consider When Designing Toolkits to Tackle AI Ethical Issues for Practitioners?”. Workshop paper for “Co-designing Resources for Ethics Education in HCI”, the 2021 CHI conference on human factors in computing systems. 2021.

[3] Crick, Bernard. In defence of politics. A&C Black, 2005.

[4] Dressel, Julia, and Hany Farid. "The accuracy, fairness, and limits of predicting recidivism." Science advances 4.1 (2018): eaao5580.

[5] Giroux, Henry A., and Emiliano Bosio. "Critical pedagogy and global citizenship education." Conversations on Global Citizenship Education. Routledge, 2021. 3-12.

[6] Hagendorff, Thilo. "The ethics of AI ethics: An evaluation of guidelines." Minds and Machines 30.1 (2020): 99-120.

[7] He, Baogang. "Deliberative citizenship and deliberative governance: a case study of one deliberative experimental in China." Citizenship Studies 22.3 (2018): 294-311.

[8] Johnson, Laura, and Paul Morris. "Towards a framework for critical citizenship education." The curriculum journal 21.1 (2010): 77-96.

[9] Lee, Michelle Seng Ah, and Jat Singh. "The landscape and gaps in open source fairness toolkits." Proceedings of the 2021 CHI conference on human factors in computing systems. 2021.

[10] Perrault, Simon T., and Weiyu Zhang. "Effects of moderation and opinion heterogeneity on attitude towards the online deliberation experience." Proceedings of the 2019 CHI conference on human factors in computing systems. 2019.iyl

[11] Pillai, Ajit, et al. "Co-designing resources for ethics education in HCI." Extended Abstracts of the 2021 CHI Conference on Human Factors

in Computing Systems. 2021.

[12] Westheimer, Joel. What kind of citizen?: Educating our children for the common good. Teachers College Press, 2015.

[13] Zhang, Weiyu, Tian Yang, and Simon Tangi Perrault. "Nudge for Reflection: More Than Just a Channel to Political Knowledge." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021.