Title Page:

Title: Co-designing AI with Youth Partners: Enabling Ideal Classroom Relationships through a Novel AI Privacy Framework

Author Name and Affiliations:

Michael Alan Chang\textsuperscript{a} (michaelc@berkeley.edu)
Mike Tissenbaum\textsuperscript{b} (miketiss@illinois.edu)
Thomas M. Philip\textsuperscript{a} (tmp@berkeley.edu)
Sidney D’Mello\textsuperscript{c} (sidney.dmello@colorado.edu)

\textsuperscript{a} University of California, Berkeley
2121 Berkeley Way, Berkeley, CA 94720
United States

\textsuperscript{b} University of Illinois
1310 S. Sixth St. Champaign, IL 61820-6925
United States

\textsuperscript{c} University of Colorado, Boulder
6661 CO-7, Boulder, CO 80303
United States

Corresponding Author: Michael Alan Chang

Declaration of Interest (for Title Page): This research was made possible by NSF grant DRL-2019806 and NSF grant #2021159.

(end Title Page)

(Cover Letter)
August 30, 2023

The Editors
Special Issue of Computers and Education

Dear Dr. Milrad, Dr. Cerratto-Pargman, Dr. McGrath –
Attached please find the manuscript entitled “Co-designing AI with Youth Partners: Enabling Ideal Classroom Relationships through a Novel AI Privacy Framework” for consideration in the Special Issue of *Computers & Education: Towards Responsible AI in Education: Challenges and Implications for Research and Practice*. This co-authored manuscript with Mike Tissenbaum, Thomas M. Philip, and Sidney D’Mello is original, not previously published, and not under concurrent review elsewhere.

Potential reviewers for this manuscript include:

1. Tamara Clegg, Associate Professor, University of Maryland.
   tclegg@umuc.edu
2. Betsy DiSalvo, Associate Professor, Georgia Tech.
   bdisalvo@cc.gatech.edu

We would appreciate it if you consider this manuscript for publication in the *Computers and Education* and would also appreciate any feedback you or your reviewers might provide.

Sincerely,

Michael Alan Chang, Ph.D.
Postdoctoral Researcher
Graduate School of Education
University of California
2121 Berkeley Way,
Berkeley, CA 94720
michaelc@berkeley.edu
408-568-9455

(end cover letter)

**Keywords:** Architectures for Educational Technology Systems, Cooperative/Collaborative Learning, Cultural and Social Implications, Interdisciplinary Studies

**Research Question (200 ch):** How does a workshop that supports students' explorations around re-imagining key relationships of schools and technologies could engage them as co-designers of collaborative AI partners?

**Advances the field. (200 ch):** This paper presents a participatory design approach for elucidating students’ dreams and concerns around AI in classrooms and a novel relational privacy framework for collaborative AI.
Highlights:
- Participatory Design Approach with Youth for Imagining Collaborative AI Partners
- Co-designed Novel Relational Privacy Framework for Collaborative AI Systems
- Workshop findings and framework operationalized in a Community Builder AI Partner

Abstract: In recent years, the design of AI-based tools for educational spaces have been largely driven by researchers who impart their past expertises and experiences on the development of research tools. While these approaches typically lead to technically feasible approaches and are often well-grounded in theories of learning, youth agency is typically limited in this process. In this paper, we argue that we have a significant ethical responsibility to incorporate youth voices – in particular, their dreams and concerns – into the design of AI tools starting from conception. Drawing from recent scholarship which advances ethics and relationality in participatory co-design with youth, we introduce a co-design methodology in which youth are supported in imagining expansive technical possibilities for K-12 public schools, grounded within affordances, limitations, and tradeoffs when using empirical machine learning techniques. This approach is demonstrated through our [Anonymized Workshop Title], which brought together 30 historically minoritized youth in conversation with experts in both education and technology. Through detailed case study on the enactment of the workshop, including a thematic analysis of the activities the youth engaged in and their outputs, we identified new, expansive relational possibilities for AI and developed a novel Relational Privacy framework that guides developers in the development of collaborative AI platforms. We conclude by connecting these findings and frameworks to the design of newly enacted AI-based applications and underlying data infrastructures.

Section 1: Introduction

While the use of intelligent software agents – software that can track student activity and products to help support student learning and orchestrate classroom activities – has been around for decades in educational settings (Andersen, Boyle & Reiser, 1985; Authors, 2019), the rise of Artificial Intelligence (AI) has rapidly advanced the computational affordances of these systems. These AI systems have also brought new challenges, requiring designers and researchers to consider the interplay between teachers, students, and the AI systems in order to effectively support teaching and learning (Lawrence et al., 2023).

Further complicating matters is the general lack of approaches for effectively bringing in the necessary stakeholders as part of the co-design process of AI systems (Xu et al., 2023). Recent work has aimed to address this by including teachers in the co-design process (Lawrence et al., 2022); however the work with students as co-designers has been relatively limited (Van Brummelen et al., 2023), despite its value across many other areas of computer-supported learning (Druin, 2002; Ahn, 2021). This is particularly critical for AI implementations that aim
to work alongside learners, as collaborators or conversational agents, as recent research has shown that students' trust of the AI systems, particularly around issues of privacy, are critical for their successful uptake in classroom settings (Ahn, et al., 2021; Ozmen Garibay et al., 2023; Rheu et al., 2021).

In response, this paper aims to outline an approach to participatory design that centers students in the design process, with a particular focus on their desires and concerns around the implementation of AI systems to support collaboration in their classrooms. In the sections below, we outline a participatory design approach that we entitled the [Anonymized Workshop Title] through which we sought to surface youth’s hopes for expansive collaborative possibilities with AI; students' concerns and opportunities for learning with AI that arose from the workshop; a novel AI Privacy framework for more responsible AI development to support collaborative learning; and an exemplar tool that built upon these findings.

Section 1.1: Participatory and Co-Design of Educational Technologies

The call for involving students in the design of technologies that will be used in their classrooms is not inherently new, and yet, far too often design educational technologies are designed as if children were "little adults", rather than existing with their own norms, needs, and concerns (Druin, 2002). In response, participatory design has been uplifted as an approach to the design of educational technologies that recognizes that students are best served when they are given a high degree of agency in the tools that are developed for them (Druin, 1999). Placing students at the center of the design process mirrors the principles of learner-centered design that sits at the heart of much of the field of computer-supported collaborative learning (Bonsignore, 2013).

More recently, there has been a growing recognition that in order to effectively engage students in participatory design approaches, we need to not just understand how they feel about the tool itself, but also apply human-centered design (HCD) approaches work students them to better understand their personal values, dreams, concerns, and socio-emotional needs (Iriarte et al., 2023; Shehab et al., 2021). As educational technologies become increasingly integrated into all facets of students' learning experience, there are increased issues around students' perceptions of privacy, autonomy, and sense of security (Hourcade et al., 2016; Kumar et al., 2018). By applying an HCD lens to the participatory design process, we can more effectively engage students in co-investigating the issues towards deeply understanding what matters to them, and design educational technologies that are more equitable and relational (Guha et al., 2013).

Such an approach is particularly important when researchers and designers are working with students around technologies that they may only be tangentially familiar with, despite their potential impact in their lives and classrooms (Kumar et al., 2018).

Section 1.2: Participatory Design of Artificial Intelligence Tools for Learning
The use of AI to support classroom learning and collaboration has seen a meteoric rise in interest in recent years (Touretzky et al., 2019). While there has been some work engaging teachers in the design of these systems and their related curricula (Chiu, 2021; Hrastinski et al., 2019; Lin & Van Brummelen, 2021; van Leeuwen et al., 2018), there has been relatively little work that engages students as partners in the design process (Holstein et al., 2019). This broad lack of engagement with students as participatory designers runs the risk of not attending to their concerns about the many valid ethical and privacy related concerns about the integration of AI into their classrooms (Akgun & Greenhow, 2022; Slade et al., 2019). A failure to attend to the concerns of students around issues such as privacy in the design of these tools, has resulted in students rejecting their inclusion in their classrooms and the tools needing to be pulled (Ahn et al., 2021).

Within the realm of AI supported learning, AI systems that support students during real-time collaborative learning activities raises a unique set of challenges for ethical design around the ways that students' data is collected, aggregated, processed, and ultimately used to impact their classroom behaviors (Adams et al., 2023). Within these systems, students' discourse, (among many other multimodal datum) is collected (Stewart et al., 2023). However, capturing all of their discourse raises new issues around students' privacy, agency, and sense of safety, as students' personal conversations can be recorded and processed. This kind of discourse is a common part of classroom activities (Dwyer et al., 2004) and a lack of care about students' perceptions about what is captured and what is done with this data, has the potential to threaten their sense of safety in the classroom (Slade et al., 2019). To this end, it is only through design approaches that include students from the outset that we can effectively design systems that leverage their discourse while still respecting their concerns (and dreams of what AI can be).

**Section 1.3: Methods for Enacting Participatory with Youth**

Engaging youth in participatory design requires different approaches than designing with adults (Druin, 2002). Traditional power dynamics, notions of authority, and adults "knowing more" can cause friction in the design process if not properly attended to (Guha et al., 2013; Kumar et al., 2018). At the same time, participatory design with youth can be particularly powerful as children are not as concerned about how things "should work" (Druin, 1999). However, many of these approaches often foreground the design process of the final product with less attention paid to youths' values, experiences, perspectives, and beliefs (Iriarte et al., 2023). For instance, product-oriented participatory design focuses on the design of an end product, which can limit the range of ideas and input that youth can provide (DiSalvo et al., 2017; Iriarte et al., 2023).

The previously referenced projects typically focus on educational and classroom contexts; youth occupy important positions in those classroom contexts and the focus rightfully is on shifting youth’s positionality from a “delegitimized stakeholder” to a meaningful partner in design. This re-mediation (Gutierrez, 2009) is a key first step, but much of this work tends to
assume human-human classroom relationships (e.g., the experienced relationship between a student and teacher) are static in this process. Recent work in the learning sciences around cooperative design and participatory design (Gutierrez & Jurow, 2016, Bang & Vossoughi, 2016) – while typically removed from co-designing technical tools – attend in particular to relational and ethical issues that are surfaced in co-design. This body of work suggests that if we are to move towards equitable and ethical futures, it is insufficient to just elevate the voices of youth in design; the co-design space must also support the youth in re-imagining relationships that happen within the institutional context being co-designed (Philip et al., 2023). In this case, our focus is schools. Concretely, the space of exploration shifts from: “how can I as a researcher support youth in feeling authoritative in designing AI” to “how can I as a researcher support ideal relationships in classrooms, and what possibilities for AI does that open up?” In our approach, we bring participatory design research into conversation with the literature on co-designing AI with youth; what happens when we simultaneously support youth in imagining possibilities spanning both schools and artificial intelligence?

Section 2: Material and Methods

We start by describing the author’s institutional context to explain some of the constraints that shaped the design of our workshop. Next, we describe the approach of the [Anonymized Workshop Title] and provide concrete details about the workshop implementation and participants. Finally, we describe our data and analysis processes.

Section 2.1: Author’s Institutional Context

The authors of this manuscript are members of an interdisciplinary AI Institute (hereon referred to as the Institute). The Institute was founded to develop AI-based tools that support small-group collaborations in United States K-12 classrooms. Institute researchers share a commitment to Responsible Unnovation (Stilgoe et al., 2013). Responsible Innovation orients us to inclusively anticipate concerns and dreams held by impacted actors, and also to be responsive to those dreams and concerns in our design efforts. Early on, Institute technical and educational experts stated a commitment to being responsive to the findings in our [Anonymized Workshop Title]; thus as designers of the workshop, we sought to walk a line between (a) youth’s expansive hopes and dreams for schools which may extend beyond our Institute’s capabilities and (b) the existing technical and educational expertises that were featured in our institute. Throughout our workshop, we explicitly communicated this tension to youth and made the Institutional context transparent.

Section 2.2: Workshop Approach

Our [Anonymized Workshop Title] sought to surface youth’s hopes for expansive collaborative possibilities inside schools by re-mentiating (Gutierrez et al., 2009) key relationships across both schools and AI. This focus on re-mediation (instead of remediation)
supports us in moving past deficit-oriented views of youth towards understanding and re-imagining the complex ecologies where youth occupy. Given the constraints described in Section 2.1, we surfaced and re-mediated the following key relationships:

Our first focus was the relationship between technology designers and youth, where youth expertise is commonly dismissed. We rectify this by positioning youth in our workshop space as experts in AI and schools, who have opportunities to freely imagine, propose, and create expansive AI possibilities. Facilitators specifically took a step back, primarily serving to facilitate discussion, and occasionally contributing technical expertise when appropriate to create a greater sense of feasibility around a technology-based proposal.

Second, we re-mediated the relationship between developers of Artificial Intelligence (AI) and users of the tools. At the time this workshop was held, our Institute primarily was interested in investigating empirical-based AI approaches, which relied on collecting large volumes of user data, and making inferences over that data using black box statistical approaches. Extensive data collection is essential to the effectiveness of these approaches, and has been shown to dovetail with broader systems of authority and surveillance, particularly around oppressed populations (Benjamin, 2019). Moreover, these largely black box models have been shown to take on harmful biases against minoritized groups. Thus, when positioning youth as developers of these technologies, we emphasized that youth had the ultimate say about (a) what data should and should not be collected about them and (b) what should and should not be inferred from that data.

Thirdly, we re-mediated the relationship between students and other actors (e.g., other students, teachers, and community) in American K-12 public school collaborative contexts, a context alike our institute’s existing district partnerships. Under the grammar of schooling (Tyack & Tobin, 1994) that has defined the American social imagination of public school education, collaboration is often oriented towards narratives of efficiency, driven and monitored by an authoritative teacher or high-status classroom actors. Compounding this challenge: even outside the explicit space of collaboration, school structures often delegitimize youth as stakeholders in public schools broadly as changemakers (Brion-Meisels & Alter, 2018). Thus, when positioning youth as agents of relational change in classrooms, we made clear that youth’s classroom experiences were of foremost importance in the dreaming of new worlds, and created opportunities to dream about their ideal classrooms alongside technological possibilities.

Section 2.3: The [Anonymized Workshop Title] Implementation

Our [Anonymized Workshop Title] occurred remotely over five days during the summer of 2021, for three hours each day. To prepare for the dreaming phases of our workshop, the first two days of our workshop sought to highlight the key affordances and limitations of supervised, empirical machine learning techniques, particularly applied towards natural language processing tasks. Here, we describe the early workshop activities but, for brevity, do not go into great detail about youth’s engagement. For more details about that, see [blinded] (Author, 2022).
In the opening activities of our workshop, we asked youth: “where do you experience AI in your everyday life?” Many surfaced their experiences using recommendation algorithms in social media sites (e.g., Instagram, TikTok). Building on youth’s personal experiences, we provided a brief lecture explaining the high-level mechanisms of supervised, empirical machine learning techniques. In particular, we broke down those techniques into training and inference. Training identifies patterns over (typically) large data sets, while inference uses previously trained models to categorize previously unseen data. We emphasized that AI models could infer unexpected patterns from the data, e.g., a retail store identifying a customer as pregnant from her purchasing activity.

After the lecture, facilitators led youth in an activity where they could design an AI algorithm that could support collaboration in social media sites. This activity offered us a chance to start making connections to small-group collaboration, our workshop’s focal classroom context. To scaffold this process, facilitators started by asking youth what AI actions and predictions might support them in collaborating. Then, facilitators asked youth to consider how raw data collected by social media sites today (e.g., likes, clicks, etc.) might inform predictive outcomes. In the final activity on day 2, facilitators introduced a sociopolitical perspective to AI and argued that technology takes on broader biases that exist in society. Youth then took part in a series of small group discussions, where they applied this newly learned lens towards a number of AI-driven applications: automated game moderation, a legal case involving AI and protected classes, Microsoft’s Twitterbot Tay, and the short-lived deployment of police robot “digidogs” in New York City.

This brief introduction to AI paved the way for the last three days of the workshop, which were spaces of dreaming, designing, and enacting. Details of these activities are sequentially embedded in the following Results section.

Section 2.4: Participants and Data Collection

Our [Anonymized Workshop Title] worked with thirty self-identified youth of color from California, Colorado, and Oklahoma. 14 youth identified as male, 14 identified as female, and 2 identified as non-binary. We had no particular selection criteria other than youth expressing an interest in learning more about artificial intelligence. Zoom video recordings were saved for large-group and small-group discussions. Additionally, artifacts (worksheets and drawings completed using Jamboard) created during the workshop were saved.

Section 2.5: Analysis

Between each day of the workshop, workshop facilitators met and recorded shared themes and wonderings. Following the workshop, the authors watched video recordings together and identified key themes. Across discussions in the workshop, a thematic analysis was used to pull out emergent ideas across all of these groups. Our interest in re-mediating relationships within collaborative contexts led us to start by observing what types of classroom relationships youth attempted to re-imagine. We transcribed the data line-by-line and started by coding each
turn where youth either expressed a hope that they hoped that AI could help with, or a concrete proposal involving AI in classrooms. Over initial dreaming activities in the workshop, youth re-imagined the following classroom relationships in their dreaming activities: student-student, student-teacher, student-administrator, and student-community. As a research group, we collectively observed that as the workshop moved past conceptualization into the implementation and evaluation of the AI tool, other classroom relationships became salient. For instance, youth made an initial proposal where the AI helps to connect groups based on topical interest, a student-student relationship. However, when considering the data required by the AI to make it work, youth made salient student-community connections since their topical interests were influenced by community events. Based on these discussions, we then conducted a second thematic coding for the implementation/evaluation parts of the workshop (day 4 and day 5), this time coding for relationships that were made salient after initial AI proposals were made. Because of this shift mid-workshop, we present our thematic analysis as a detailed case study, presenting each day of the workshop sequentially and highlighting identified themes along the way.

Section 3: Results

Prior to the activities described here in this Results, youth had learned about machine learning and a sociopolitical perspective towards technology. For each major activity in days 3-5 of the [Anonymized Workshop Title], we describe the design activity before diving into one representative vignette. We then show how these Results lead to the development of a novel Relational Privacy framework, and finish the section by showing how the Relational Privacy framework was operationalized in a novel AI tool built by our Institute.

Section 3.1: Ideal Schools and Imagining Possibilities (Workshop Day 3)

In small breakout groups, youth were given the following prompt: “what does ideal collaboration look like, and how might AI help to make that a reality in your schools today?” As the first dreaming activity, the focus was neither on ethics or technical feasibility, merely the ideation of exciting proposals. In this vein, we did not view the AI proposals made in this activity as final or above criticism. Instead, in our analysis, we found it meaningful to look beyond the immediate proposals and understand where youth were coming from as they constructed these tentative proposals. In this vignette, we illustrate one breakout group’s discussion around several AI possibilities. Participant quotes in this section are slightly modified to remove filler words.

In this small group, one participant Eric provided some framing to begin the discussion: “Okay, so my thought about this is not really technology, it’s theory. Collaboration should be about equity not equality. In collaboration everyone should aim to give as much as they can and get as much as they can.” Eric made central equity as a key goal of AI-supported collaboration, and more specifically, and emphasized that each student, depending on their circumstances,
could “give” and “get” from each other in different amounts. This framing immediately implicates AI’s role in mediating relationships between students, rather than just between the AI and an individual student.

Following Eric’s remark, Ricky declared that teachers “can only help so many students… if you had AI, it could tailor itself to every student’s needs.” Ricky did not counter Eric’s equity framing, but offered a different imagination about how AI supports equitable goals: the AI acts as a proxy for a teacher and individually supports students. While Ricky advocated for a student-AI tool, his framing of the tool (“teachers can only help so many students”) suggests that the proposal addressed a key frustration, namely that teachers could support the youth at a personal level, but they lack the resources to give sufficient attention to students.

Another student, Akil invoked a number of different possibilities and built on Eric. Akil first considered how AI tools could keep student peers from getting in “trouble for talking to your friend or asking a question.” Next he discussed how “Google or Siri” could help answer a “question you weren’t sure how to work or were embarrassed to ask.” Finally, he imagined a proposal that brought the community into the classroom and capitalize on his community passions: “an AI that could connect you to different resources. Instead of googling jobs in my neighborhood, a career test! Where you excel, from there you do really well in these areas, we think these organizations would go well with you.”

Esperanza later chimed in: “I feel like collaboration shouldn’t be dreadful. In my classroom, it’s dreadful many times.” Later, she added: “I feel like people take [classes] to get credit, some people don’t take it really seriously. I think it’s a problem… I just feel like some people don’t take it seriously in general. The whole class, the teacher, the topic!” After some additional prompting by the facilitator, Esperanza replied hesitantly with an AI that re-imagined student-student interactions: “Maybe the ones who are more experienced in it or more interested in it could somehow motivate the others to be interested in it as well.”

Over the course of this discussion, youth surfaced a variety of frustrations and hopes for what collaboration in school could be like, spanning relationships with each other, with teachers, and with the broader community. Our analysis in this section reveals that ultimately, youth had deep, substantive hopes for re-mediating key relationships they experienced inside classrooms. Our co-design space created the opportunity for youth to imagine how they might tackle some of these frustrations and hopes using AI. While these proposals may raise many concerns about technical feasibility and learning theory, the discussions underlying the proposal makes clear that if we are to realize youth expansive hopes using artificial intelligence, we must better understand how AI can help support, accompany, and uplift existing classroom relationships.

In other groups, youth described wanting to feel cared for, respected, and heard in classrooms. Common AI proposals made in response to these experiences included a friendship finder (“Tinder but for friends”) and a collaboration matcher based on shared interests. Another common frustration was youth’s frustration with other youth for perceived disruptions to their collaborative efforts; they imagined an AI that would keep those disruptive youth accountable. These proposals formed a starting point for development and critique in the later days of the workshop.

Section 3.2: Designing AI, Considering Data (Workshop Day 4)
During day 4, we planned to take previously conceptualized AI proposals and better understand their ethical implications, particularly around an essential part of empirical AI systems: training models on large amounts of user data. The discussion was supported by a worksheet (Figure 1), which was custom created for the purpose of this exercise. The worksheet asked youth to collectively consider their proposed AI agent’s embodiment, the actions taken by their imagined agent, the corresponding inferences the AI would need to make, as well as the raw data necessary to make those inferences. Additionally, for each of those dimensions, the worksheet explicitly asked youth to describe what they hoped the AI agent would not do.

In coding these excerpts, we observed that the worksheet raised new considerations for previously proposed AI features. These considerations were driven by youth’s experiences around classroom relationships that differed from the specific relationship being re-imagined. We coded for these different relationships, and sought to understand how these newly salient relationships shaped how the young folks viewed their original relational possibilities for the AI.

In the following vignette, we share the discussion around one oft-proposed metaphor, which we have termed the Collaboration Matcher, an AI that supports youth in identifying collaborative partners. Initially, the Collaboration Matcher was designed to help mitigate the feeling of dread when it came to collaborating with each other: the focus was thus on the relationships between students and other students.

**Figure 1**

A completed data privacy worksheet used by youth to imagine the privacy tradeoffs involved in proposed AI tools
In this breakout group, youth explored how an AI could identify how individuals learn by examining their individual attributes (“hav[ing] similar personalities”) or learning styles (“some learn by reading, doing some activity”). These initial ideas, building on notions of “learning styles” that remain common in popular discourse even though they have been debunked in educational research (Newton, 2015), were later problematized by the workshop participants during this discussion; we omit those discussions for brevity.

Youth then collectively filled out the data privacy worksheet for the Collaboration Matcher. While filling out the worksheet, the participants stated explicitly that they did not feel comfortable having an educational AI agent collect personal data outside of the physical school boundaries. Picking up on this initial boundary, the facilitator seeded a discussion by asking, “How is the AI supposed to learn your learning style without that [out of school] information though?” Melinda quickly responded, “I don’t think it needs to know our personal relationships with people.” Another participant, Sam, further refined the boundary: “I don’t want them to know who I’m going out with.” From these excerpts, we don’t know the specific relational experience motivating Sam and Melinda’s discomfort with sharing personal relationships, but we infer that they are speaking to a relational dynamic where that data may be used to harm or embarrass them. The facilitator (a former teacher), pointed out that such a boundary around using personal data might constrain the capabilities of the Collaboration Matcher, arguing that teachers often observe romantic crushes between youth, and assign them into the same group in order to motivate them. Melinda affirmed the facilitator: “Teachers assign this because the boo know that they might be trying to impress them.” The Collaboration Matcher was initially designed to support student-student relationships; because we engaged the youth in considering the data implications of the AI tool, the relationship between student and teacher also became salient and led to certain relational possibilities for the Collaboration Matcher being closed off.

Sam then made an argument that the Collaboration Matcher could work even without personal information: “Compatibility has nothing to do with the level of friendship! I can never be in a group with [my friends]. I would never get anything done unless it’s an art project. You put me in a history group with my homies, F. Terrible! You put me with a random dude who I said wassup to a couple times in the corner of the room, we ace-ing that.” While the initial part of Sommer’s argument seemed to de-emphasize the importance of personal relationships, Sommer’s later argument in fact illustrated how personal relationships do indeed contribute to the effectiveness of the Collaboration Matcher. Faced with a relatively new set of considerations around a relationship-oriented AI metaphor, Sam and Melinda were actively making sense of tradeoffs and tensions.

This sense-making around boundaries continued. During a lull in the conversation, Sam wondered: “What if you are tardy [in class] for health reasons? Do you want the AI to know that? A lot of people would see that as an invasion of privacy. I would be okay with that. A recurring issue.” Sam acknowledged that others might not share his comfort, but for this specific application, he would give away private health data. Melinda, reflecting on Sam’s comment, later stated: “Schools have a lot of our information already. Would it just have all the information that
the school already has? I have therapy, and the school already knows that. Still kind of sketch—but still!” Sam and Melinda continued to break away from the hard line between personal and academic data for AI use in classrooms. They both acknowledged the sensitivity of health information, but nevertheless deemed the tradeoff to be worthwhile. Now considering the student-administrator relationship, the lack of perceived empathy towards youth’s health conditions was sufficient to outweigh the potential harm from sharing that data. Sam and Melinda further nuanced the seemingly hard boundaries they drew at the outset.

Similar discussions occurred in other Day 4 breakouts in the workshop, where youth grappled with the tradeoffs of designing an AI that supports them in building relationships with each other. In other groups, Nick and Eno discussed sharing their individual “weaknesses,” Juan and Kory contemplated biological signals like their “brain waves,” and Larry wondered about out-of-school interests (“relate certain subjects to things out of schools”). Across these discussions, several findings become clear: designing relationships with AI lead to an unfamiliar and slightly uneasy set of discussions around what data is collected and what is inferred from that data.

Taken together, youth’s relational hopes for an AI agent (e.g., Collaboration Matcher) may not cleanly align with the relational data necessary to make those AI agents a reality. Relationships outside of the relationship under design become salient and close off or open up possibilities. Moreover, these expectations are constantly under negotiation; even over the course of our 45 minute discussion, youth increasingly nuanced an initially hard boundary between personal data and “academic” data.

Section 3.3: Theatro Activity (Workshop Day 5)

Our last day of the workshop involved a modified Theatro Activity (Boal & McBride, 2013). In Theatro, actors first perform a scenario where a perceived injustice occurs. After an initial performance, the actors repeat the scene but give way to audience members who “tap into” the scene and experiment with different interventions. To better understand the role of AI in collaboration, we designed the activity such that scenes must feature a virtual AI assistant (e.g., Amazon Alexa) and audience members could only step up to replace this virtual assistant. The Theatro script was co-created with workshop participants, who incorporated immediate needs they experienced while collaborating in schools. Through this activity, youth were able to test the boundaries of their comfort with an AI agent’s capabilities (Holstein et al., 2017) and determine how an AI feature could effectively carry out its mission.

We highlight one small-group discussion, where youth experimented with an AI that kept “disruptive” youth from interfering with classroom collaborative efforts. This AI proposal was controversial, as one of the most commonly proposed AI tools that was also frequently complicated by youth. In the breakout, youth iterated through the script four times, trying out increasingly punitive measures to augment the AI’s authority. These proposals ranged from threatening “disruptive” students by deducting grades, summoning the teacher, and finally, in the
most extreme case, summoning parents. The proposal – which started off as a student-student relational tool – evolved to bring in teachers and parents. After this activity, participants debriefed their experience.

Eric initiated the conversation: “For it to be more effective, it would need more power, the ability to mute someone. I don’t know what other power it would need before it becomes illegal.” Larry and Jake followed up by suggesting other suggestions that might position the AI as a point of authority: “fail a class” or to give an “individual a bad grade,” giving the AI authorities traditionally held by a teacher. Alana went further and suggested that all dialogue transcripts should be recorded and then directly sent to teachers for review.

Much like the data privacy based activities previously conducted, the format of the activity – AI grounded within familiar collaborative settings – helped to make salient other, overlapping classroom relationships. In particular, youth could not envision certain AI possibilities effectively functioning without sharing information with authoritative others: teachers, school administrators, and parents. In other words, some AI possibilities will go awry unless data is shared outside of the domain and relational context where it was collected. Youth felt (in the most extreme case) that entire collaborative transcripts – which may contain dialogue that youth do not wish to share with teachers – need to be shared with teachers if they are to work effectively. When we start to view AI as a tool to help design ideal relationships, we start to see how the design of the AI must carefully consider existing authoritative relationships in the context where they are to be eventually deployed.

Section 3.4: A Relational Privacy Framework derived from Findings

Our findings presented in sections 3.1-3.3 offers a mixture of tangible design considerations as well as a number of provocative design tensions. In this section, we share a novel framework (based directly from our findings) that systematically supports educational and technical designers in gradually realizing expansive relational possibilities using AI. We refer to this framework as a Relational Privacy Framework for Collaborative AI Systems.

In figure 2, we systematically illustrate the design tensions that emerge when building an AI agent that supports ideal classroom relationships. Each step in the machine learning pipeline raises particular constraining factors which shift the agent towards existing ways of doing things in schools, instead of its intended goals of fostering ideal classroom relationships. Our first finding shows that underlying youth’s AI proposals was a desire to meaningfully change key relationships between each other that they felt dissatisfied about in modern K-12 classrooms. In the training of machine learning models explored in Vignette 2, we saw how youth preferences for data privacy estranged certain relational possibilities for the AI agent. A particular ethical quandary emerges for designers; youth’s preferences for data collection inadvertently made it harder to build AI that youth themselves wanted. Finally, through an improvisational activity grounded in youth’s own classroom experiences in Vignette 3, our third finding draws an explicit connection between the intended uses of an AI tool and the institutional contexts where the tool is deployed: AI proposals which reproduce authoritative classroom relationships (e.g., an AI that
keeps youth on-task) requires support from authoritative figures (e.g., snitching to a teacher). In totality, if we do not carefully attend to these tensions, designer’s aspirations to build out youth’s ideal AI tools may unintentionally reproduce existing (unideal) classroom relationships.

Figure 2

A summary of workshop findings. If we break down a machine learning pipeline in a co-design process, ideal AI-supported classroom relationships are narrowed and ultimately come closer to reproducing existing classroom relationships.

Towards operationalizing the findings in Figure 2, we brought to bear Nissenbaum’s theory of *contextual integrity* (2004). Critical to Nissenbaum’s framework is the notion that all information flow occurs within some social context (e.g., education, healthcare, religion, etc.), marked by specific practices, goals, and ends; participants within these contexts are aware they are participating in this practice and can explicate the norms of this context. For the purpose of modeling information transmission within these contexts, the context is parameterized by four categories: sender, recipient, information “types,” and transmission principles. Every agent operating inside a context has a role (e.g., student, teacher, administrator). Transmission principles represent terms and conditions that dictate how information is shared between participants (consent, coerced, in confidence, via a warrant, etc.). Contextual integrity is met when 1.) information types are appropriate for a given context; and 2.) the flow of information follows the transmission principles. If contextual integrity is not met, user’s expectations of privacy are violated, provoking public outcry, leading to lack of trust and adoption towards the intervention.
Contextual integrity offers a developer-friendly formalization of our privacy within complex relational spaces, while at the same time making salient the relational possibilities and concerns, which featured so prominently in our Findings. While the static notion of “roles” can be a stable and useful construct, the educational classroom context in particular demands a more flexible understanding of these norms; as we saw in our second vignette, personal and academic boundaries are constantly being breached and re-assessed. As Criado and Such (2015) argue, context, roles and associated informational norms are “implicit, ever changing and not a-priori known.”

Figure 3

Transfer of data recorded by an AI partner in a classroom, through the lens of contextual integrity

Bringing this discussion specifically into our goals of supporting classroom collaboration, Figure 3 shows how contextual integrity frames one model of AI-supported collaboration. At the most immediate level, a sender is conveying some information $x$ to the original recipient(s), and like we saw in Vignette 1, an AI Partner might be supporting the Sender and Original Recipient in developing some desired collaborative relationship with each other. The shared information is subject to particular expectations of privacy, based on the following contextual integrity tuple: *(sender, original receiver(s), the information type of $x$, the (sender-original recipient) transmission principle)*. In addition to sharing with the original recipients, the AI partner also collects that information, renders some computational model over the data and, depending on the application, shares the result $f(x)$ back to the sender/original receiver and/or possibly shares some
result $g(f(x))$ with secondary recipients. The expectations for privacy on this exchange depend on the following tuple: \text{\textbf{(sender, secondary receiver(s), the information type of x, the (sender-secondary recipient) transmission principle)}}. At the highest level, the central tension that we identified are the conditions for passing of information $x$ between the immediate relational context and the secondary relational context.

Below in Table 1, we translate our findings into an iterative design framework for AI in educational classrooms. This framework aims to clearly explicate the tensions between the data required to build AI systems dreamed up by youth, and the students' own sense of safety, agency, and autonomy. Navigating these tensions requires compromise rather than consensus from design teams and stakeholders (Druin et al., 2013), as articulated by students' own acknowledgement of these issues Vignette 2. Each of the driving questions below are starting points for designers and researchers to unpack these tensions and understand where compromise is needed and how it will impact the desired design of the AI systems.

Table 1

Relational Privacy Framework for Collaborative AI Systems

<table>
<thead>
<tr>
<th>Steps in Relational Framework</th>
<th>Connection to Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question 1: What are the ideal relationships being supported by the AI?</strong> As we showed in Vignette 1, the core of what youth were designing for in our workshop was ideal collaborative relationships. There might be multiple relationships being supported by the AI in the immediate relational context, and designers should explicitly name all of these relationships in ways that can be understood by the students.</td>
<td>Vignette 1: Designing for Ideal Classroom Relationships</td>
</tr>
<tr>
<td><strong>Question 2: If we are to implement the designed ideal relationships from Question 1, what do researchers, designers (i.e., Secondary recipients) consider to be key information types/transmission principles that are relevant to sustain the actor relationship?</strong> In Figure D2, secondary recipients are defined as individuals who are not in the immediate relational context where information is exchanged. For instance, in a student-student interaction, teachers or researchers might be secondary recipients. In this question, secondary recipients make explicit the kind of information they would prefer to have access to in order to fully</td>
<td>Not directly inspired by vignette (added to support design process)</td>
</tr>
</tbody>
</table>
realize the ideal relationship imagined by youth. These considerations are likely influenced by learning theories or past experiences in classrooms.

**Question 3:** If we are to implement the relationship being designed for, what do youth (i.e., the sender) consider to be their preferred information types/transmission principles that are relevant to sustain the actor relationship? To sustain the ideal relationships in a classroom space, designers need to comprehend the salient information types that might emerge, as well as the transmission principles that govern the flow of that information type. Fundamental to this question is the existing relationship between students and teachers. Drawing from Gutierrez et al. (1995), teachers’ willingness to engage with the third space (i.e., the space that lies in between traditionally “formal” academic talk and the classroom underlife”) deeply shapes student’s comfort level with sharing particular kinds of information with teachers. As we illustrated in vignette two, answering this question is impossible without engaging co-design partners and likely varies classroom-to-classroom. More concretely, in Figure D2, this requires explicitly naming the CI-tuples with youth: (sender, receiver, information type of x, and its (sender-original recipient) transmission principle). Worksheets like the one used in Vignette 2 can be helpful towards surfacing these transmission principles.

**Question 4:** Does either of the ideal actor-relationships require data sharing beyond the local immediate context? Does that data sharing violate the initial transmission principles articulated in question 2? As we showed in vignette three, some AI tools require the sharing of data $g(f(x))$ with authoritative sources in order to work effectively. Sharing that data may violate the transmission principles between the sender and the original recipient, and thus compromise the ideal relationships that are being designed for in Question 1. For instance, an AI that helps youth form classroom friendships might be compromised if the information types salient only to a friendship are shared with an
untrusted teacher. Appropriate remedies should be considered, some of which we detail in our working example of this framework.

**Question 5:** How are information types/transmission principles contested and re-negotiated? How does that expand or narrow our design in response to question 3? As we showed in vignette two, youth were constantly redefining their ontological understandings of information types, and laying down new boundaries based on their gradual sensemaking. These experiences have the possibility to both relax and tighten the transmission principles originally constructed in question three of this framework. Upon going through this process, designers should revisit design Questions 3 and 4.

**Vignette 2:** Youth are constantly making sense of their data preferences

---

**Section 3.5: Relational Privacy Framework in Practice: [Blinded Tool]**

In this section, we show how our workshop findings and design framework has shaped a new AI tool called the [blinded tool] (AnonTool). AnonTool’s role in small group collaborations is to support student learners in building customized, ideal collaborative relationships with each other – a framing that derives directly from the first finding presented in this paper. The AnonTool display is shown in Figure 4; AnonTool uses AI-based computational models to automatically provide feedback to youth on how well they were holding accountable to collaborative agreements. There are three major agreements: Committed to Community, Moving Thinking Forward, and Being Respectful. In order to build the aforementioned computational models, annotators used small group collaboration data to code student utterances that demonstrated the agreements in practice (Authors, 2023).

**Figure 4**

*The Institute’s Display, and the agreements that AnonTool helps students stay accountable to. The fourth agreement, “Being equitable”, is still under development, and we do not discuss in this paper.*
Not only are these agreements rooted in past work on modeling collaborative problem solving (Sun et al., 2022), they are directly connected to the ideal relational hopes described by youth in our workshop. In Table 2, we sequentially go through each of the relational experiences described by youth in Vignette 1, and show how we connected them to relational agreements supported by AnonTool through machine learning. Other connections were made between AnonTool agreements and other small group breakouts, but we omit those for the sake of brevity.

Table 2

Connecting AnonTool-supported agreements with youth’s classroom collaborative experiences in the order described in Vignette 1

<table>
<thead>
<tr>
<th>Youth Relational Experience</th>
<th>Youth Proposed AI Solution</th>
<th>AnonTool Community Agreements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prioritization of Equity over Equality: Eno: “In collaboration everyone should aim to give as much as they can and get as much as they can.”</td>
<td>No explicit proposal made by Eno during Vignette 1</td>
<td>Being Respectful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Responds to others’ questions/ideas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Asks others for suggestions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Compliments or Encourages Others</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Apologies for one’s mistakes</td>
</tr>
<tr>
<td>Sense of being forgotten by teachers: Ricky: “Teacher can only help so many students. They can’t tailor it to every student’s needs.”</td>
<td>Ricky: “If you had AI, it could tailor itself to every student’s needs.”</td>
<td>See Discussion on AnonTool Student-Teacher Design Principles</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Embarrassment about asking a question Akil: “you had a question you weren’t sure how to word or were embarrassed to ask”</td>
<td>Akil: “It could be another search engine, tailored to this course or this subject.”</td>
<td>Committed To Community - Provides instructional help to each other - Asks others for suggestions</td>
</tr>
<tr>
<td>Frustration from looking for community organizations: Akil: “Instead of googling, googling jobs in my neighborhood, like a career test, where you excel.”</td>
<td>Akil: “An AI that could connect you to different resources.. we think these organizations would go well with you. I think you would partner well with etc.”</td>
<td>Future Work: not currently handled by AnonTool</td>
</tr>
<tr>
<td>Dread towards working with people who don’t take collaboration seriously: Esperanza: “I feel like people take [classes] to get credit, some people don’t take it really seriously”</td>
<td>Esperanza: “Maybe the ones who are more experienced in it or more interested in it could somehow motivate the others to be interested in it as well.”</td>
<td>Being Respectful - Responds to others' questions/ideas - Asks others for suggestions - Compliments or encourages others</td>
</tr>
</tbody>
</table>

In many learning analytics settings, computational tools collect information about youth, and by default share that information with teachers. Specifically, these tools monitor student behaviors while individually or collaboratively interacting with learning technologies and provide analytics such as student progress, their use of meta-cognitive skills, and even their engagement and affect (Aslan et al., 2019; Holstein et al., 2018; VanLehn et al., 2021) with teachers for possible interventions. While initial models of AnonTool engaged similar models, ultimately we constructed a different approach to privacy based on our usage of the data privacy framework. In the following subsection, we detail how we leveraged the data privacy framework in a simplified case where AnonTool is only supporting two relational agreements that were both proposed in
the workshop and are reflected in AnonTool’s set of agreements: “being respectful” and “moving thinking forward.”

**Relational Agreement Example 1: “Being Respectful”**

First we envision one scenario supported by AnonTool where youth identify *being respectful* as a central goal of their collaborative activity. From Table 2, we can see that AnonTool supports this through four agreements: responds to others’ questions, asks others for suggestions, Compliments and Encourages Others, and Apologies for one’s mistakes. While these might be seen as formulaic individual moves made by collaborators, we also seek to recognize the creative and non-dominant ways that youth call each other into a conversation. We have come to recognize that this requires engaging with youth’s everyday experiences, and the “underlife” of the classroom (Gutierrez, 1995). One might imagine, for instance, that a skilled youth collaborator may invite other youth into the discussion by sharing frustration about a teacher’s harsh grading tendencies; such moves have been observed in past work on equitable collaborations (Langer-Osuna et al., 2020). Sharing this information outside of the student-student relational context may invite unpleasant discussions with a teacher.

We work through this agreement in the left column of Table 3. As shown in the table, our original design sent analytics and examples of exemplar utterances from individual small groups. for their analysis; this design violates the transmission principles of the intended relationship; comments made in a confidential student-student context (e.g., building community through complaining about a teacher) would have been shared with a teacher. Based on this contradiction, our revised planned design instead shares only aggregated data across the whole class about how well the class as a whole accomplishes its collaborative agreements, and explicitly gives permission to the AI before specific quotes are shared outside of their immediate collaborative context.

**Table 3:**

*The two exemplar agreements viewed through our Relational Framework*

<table>
<thead>
<tr>
<th>Question 1</th>
<th>Agreement: “Being Respectful”</th>
<th>Agreement: “Moving Thinking Forward”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Supporting caring, collaborative relationships between students</td>
<td>Academic Collaborators Trying to Complete Task</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 2: Designer-Ideal Transmission Principles</th>
<th>Information Type: Transmission Principle</th>
<th>Information Type: Transmission Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditionally Academic Talk: Shareable outside of immediate</td>
<td>Task-Relevant: Shareable outside of immediate relational context (e.g., with teachers)</td>
</tr>
<tr>
<td><strong>Question 3:</strong> Youth-Preferred Transmission Principles</td>
<td><strong>Information Type: Transmission Principle</strong></td>
<td><strong>Information Type: Transmission Principle</strong></td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>(based on youth responses in Vignette 2)</td>
<td>Traditionally Academic Talk: Shareable outside of immediate relational context (e.g., with teachers)</td>
<td>Task-Relevant: Shareable outside of immediate relational context (e.g., with teachers)</td>
</tr>
<tr>
<td></td>
<td>Personal Information: Confidential</td>
<td>Personal Information: Confidential</td>
</tr>
<tr>
<td></td>
<td>Romantic Information: Confidential</td>
<td>Off-topic: Shareable outside of immediate relational context with conditions (e.g., personal details anonymized)</td>
</tr>
<tr>
<td></td>
<td>Critiques of Teacher: Confidential</td>
<td></td>
</tr>
</tbody>
</table>

| **Question 4:** Institutional Alignment               | **Original Design** (deemed inconsistent with principle): Analytics and examples of exemplar utterances from individual small groups. for teacher’s analysis | **Original Proposal:** Selected diarized transcripts and on-task/off-task classification shared with teacher |
|                                                      | Revised Design (with framework analysis): Aggregated collaborative scores are made available to teachers, specific quotes that exemplify teaching are first shared with students in group, who give approval before sharing with class | **Revised Designs:** None necessary, original proposal is consistent |

<table>
<thead>
<tr>
<th><strong>Question 5:</strong> Re-visiting Norms</th>
<th>TBD (will solicit from students once AnonTool is experimented with in classrooms)</th>
<th>TBD (will solicit from students once AnonTool is experimented with in classrooms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>relacional context (e.g., with teachers)</td>
<td>Off-topic: sharable outside of immediate relational context</td>
</tr>
<tr>
<td></td>
<td>Personal Relationships: shareable outside of immediate relational context</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Relational Agreement Example 2: “Moving Thinking Forward”

Next, we describe a second scenario where youth are focused on the agreement of “Moving Thinking Forward.” Similar to our third vignette, AnonTool records collaborative discussions, “scores” the collaboration based on the dimensions of “moving thinking forward”, and relays per-group analytics to teachers. The teacher can leverage these transcripts to better support young people in staying on-task during collaboration. Under this design, as noted in Vignette 3, an AI which reports off-tasks behaviors that mimic the institutional authority of schooling requires teacher support to work effectively. Within AnonTool monitored activities, collaborators trying to complete a task should primarily be engaged in academic talk that can be shared with a teacher - thus the data sharing arrangement is consistent with the data transmission principles.

Simultaneously supporting “Respectful Collaborations” and “Moving Task Forward”

AnonTool is designed to support multiple categories of agreements simultaneously; in the ideal case, a classroom environment would support both respectful collaborations and move the conversations forward. We look across the results of our framework detailed in Table D2, and observe a key tension. In the ideal case for both agreements, the classroom culture is such that youth are comfortable sharing academic and personal information with teachers and other secondary recipients. From our [Anonymized Workshop Title], however, we see that youth have very different expectations of the transmission principles that vary based on the agreement. For instance, youth expressed an explicit desire for the AI not to record personal information. For supporting equitable collaborations, sharing the AI’s findings with teachers may compromise the youth’s comfortability with building rapport amongst each other and creating the conditions for welcoming collaborators into the conversation. Most critically, we observe the key relational contexts across the two agreements and how that leads to important differences in the transmission principles. In order for the “Moving Task Forward” to work effectively, we learned from the LFW that youth believed that an AI would need to report off-task remarks to an authoritative figure like a teacher. On the other hand, for the “respectful collaborations” to work, personal details would need to remain confidential; reporting those remarks as off-task would violate the transmission principle and create a sense of discomfort and possible outrage. If AnonTool were to report personal remarks to teachers, “Moving Task Forward” would function effectively but “respectful collaborations” would be compromised. On the other hand, if AnonTool were to withhold remarks to teachers, “respectful collaborations” would function effectively while “Moving Task Forward” would be compromised. In the short-term, AnonTool has addressed this tension by aggregating all predictions at the class-level. Individual (and even group) identities are protected because of this level of aggregation; while this seems safe privacy-wise, pedagogically, it leaves something to be desired.

Section 4: Discussion

As such, this paper has three distinct outcomes for other researchers who which to enact more student-engaged AI designs: 1) The [Anonymized Workshop Title] as a participatory design approach for elucidating students' dreams, needs, and concerns around the enactment of AI systems in their classrooms broadly; 2) The "Relational Privacy Framework for Collaborative AI
systems" which highlights critical issues to attend to when designing collaborative AI systems specifically; and 3) An instantiation of a tool (AnonTool) built on the principles of the framework that attends to the principles of the framework while also outlining the tensions that arise when designers and researcher attempt to thread the needle between compromise and consensus (Druin et al., 2013). In this section, we briefly discuss some limitations of our approach and our future work outside of technology design that we believe are crucial towards being responsive to youth’s voices.

**Limitations:** It should be noted that the work outcomes described in this paper were focused specifically on extracting factors for collaborative AI systems. These particular principles may be useful for other AI designs (e.g., personal agents or orchestration agents), but such designs may require different framings during the participatory workshops.

**Future Work: Institutional Changes:** Our paper findings reveal a key ethical quandary: the promise of youth-imagined expansive AI-supported classrooms relationships vs. youth’s preferences for confidentiality around personal data inside a classroom. Our intention in this paper has not been to position youth’s preferences as oppositional or unreasonable, merely to elucidate how it can narrow expansive relational possibilities. As an Institute being responsive to this quandary, we take it upon ourselves to create the educational conditions and infrastructure that support youth in changing their transmission principles. Through our work building Research Practice Partnerships (Coburn & Penuel, 2016) with local districts, we have begun to align our technical development of relational possibilities with the (a) professional development of teachers who use AnonTool and (b) the creation of a curriculum unit that explicitly highlights how AnonTool collects, processes, and secures data (Mawasi et al., 2022). At the professional development level, our approach has been shaped to accommodate care and equity as central to teaching practice, where teachers embrace the third space, and come to recognize and respond to everyday forms of generating knowledge. At the curriculum level, we hope to better support youth in better understanding the inner workings of AnonTool. Taken together, our hope is these factors being worked on by our Institute will create the conditions for youth to more broadly embrace transmission principles – to help realize the expansive possibilities for AI that they themselves came up with.

**Section 5: Conclusions**

This work described a novel approach for supporting students as participatory designers of collaborative AI systems, through the implementations of our [Anonymized Workshop Title]. While participatory approaches have been around for decades, their use for revealing the wants, desires, concerns, and feelings of the students who will be directly impacted by them has been particularly scant (Van Brummelen et al., 2022). Given youth's concerns around the use of their data, their privacy, and the impacts of this on the uptake of AI systems in classrooms (Ahn, et al., 2021), this work comes at a particularly timely juncture as a means of outlining how these important design conversations can take place. As more AI-based systems find their way into classroom environments, we hope that this and other studies show the importance of design with
students, rather than simply designing at students, to ensure that these innovations attend to their what matters to them, while simultaneously respecting their personal values, dreams, concerns, and socio-emotional needs. If we as researchers fail to respect the needs, wants, and concerns of students in our design, we will engender distrust in them, rather than build them up as willing partners engaged in an increasingly AI-mediated future.

References


Authors. (2019). [Title omitted for blind review]

Authors. (2022). [Title omitted for blind review]

Authors. (2023). [Title omitted for blind review]


https://urresearch.rochester.edu/institutionalPublicationPublicView.action?institutionalItemId=23125

https://doi.org/10.1080/10508406.2016.1204548


