A Stakeholder-Based Framework to Highlight Tensions when Implementing Privacy Features

Julia Netter, Tim Nelson, Skyler Austen, Eva Lau, Colton Rusch, Malte Schwarzkopf, Kathi Fisler
Brown University

Abstract

Preparing university students to build privacy-preserving systems requires preparing them to design around societal contexts and stakeholders. While legislation such as GDPR and CCPA provide regulatory frameworks for such design, discussions of privacy and stakeholder values can be fairly abstract for students. From an educational perspective, teaching abstract concepts such as the "right to be forgotten" in the concrete context of technical implementation can help students grapple with what these concepts mean in practice.

This paper proposes a framework for designing technical assignments that ask students to resolve tensions between conflicting stakeholders while implementing a specific technical feature. We describe a privacy-facing assignment for a second-year introductory computer systems course, and explore its efficacy. We find that students make different design choices and implement for different values based on the specific stakeholder conflict with which they work. We also find that the assignment design engages students in thinking about how abstract values affect technical design decisions in the context of privacy.

1 Introduction

Designers of software systems must navigate design and implementation decisions that consider privacy and related values. These concerns may be inherent, or come about as a result of regulations such as GDPR [17] and CCPA [7]. Making decisions requires negotiating among competing values, such as personal privacy versus public good or the boundaries of free speech. From an educational perspective, we want students to grapple with the tensions that underlie the design of privacy-facing features. We could ask students to read the GDPR and answer questions about its requirements, but that is unlikely to help students make design tradeoffs related to privacy and data protection in practice. We believe that it is more important for students to develop a general understanding of what is at stake in making decisions about user privacy than to teach them the specifics of current privacy-facing regulations, which are almost guaranteed to evolve over time.

In recent years, we have given assignments that asked our second-year university students to read and reflect on the implications of privacy regulations like the GDPR. The submitted work was unsatisfactory: student answers tended to be shallow, perhaps dashed off quickly as students turned their attention to what they felt was more relevant technical work. In 2023, we experimented with a novel assignment design in which students had to design and implement a privacy-facing feature motivated by "right to be forgotten" clauses common in privacy regulations. To highlight the design tensions, students had to justify their approach in the context of a concrete pair of *stakeholders* with conflicting goals for how the feature would behave. The stakeholder pairs grounded the seemingly technical implementation task in the kinds of real-world tensions between user needs and preferences that system developers must navigate. We designed the pairs such that no implementation would satisfy the privacy-related concerns of both parties, as we wanted students to appreciate that privacy-facing decisions are complex and messy.

This paper describes our stakeholder assignment framework and analyzes its use in a specific assignment on managing the "right to be forgotten" in a social media system. Our analysis explores the design decisions that students made and how those varied with specific stakeholders. We describe the values that students invoked when justifying their decisions (such as freedom of speech, data ownership, or technical feasibility) and which values students saw as being in tension. We report on the aspects of privacy that our students referenced, partly as a way to understand additional topics that might need coverage in our course or curriculum.

The contributions of the paper are:

1. the stakeholder assignment framework itself;
2. our research-based analysis of how it worked for teaching about privacy; and
3. our observations of the nuances that (at least our) students bring to thinking about the "right to be forgotten" in a social media context.

Our work highlights the pedagogic value of intertwining technical tasks with learning about societal or regulatory concerns.

2 Related Work

Multiple perspectives exist on how to define privacy [13]. In this work, we don't promote any specific definition. Instead, we draw on a variety of approaches for understanding and justifying privacy-related claims which help to explain common controversies around data protection, making them relatable to computer science students as part of a technical assignment. One key controversy relates to the foundations of what gives privacy value in the first place: whether it is intrinsically valuable and requires no grounding in other values [26], or whether it is merely instrumental [18]. Instrumental justifications for privacy invoke a variety of other values and public goods which privacy protects or promotes. Those range from autonomy [19] and the capacity for personal development [22] to dignity [1] and bodily security [25]. Privacy is also considered a material precondition of democracy through providing protected spaces for developing a range of opinions and positions [10]. For our project, we designed opposing stakeholder pairs, where one stakeholder's concerns reflect one or several of these values. The other stakeholder's concerns were rooted in values considered in tension with privacy [15], such as public accountability or freedom of expression.

Our work is not about teaching people to protect their privacy [8,20], but rather on teaching future technologists how to think through conflicting values when designing for privacy. Previous work shows that computer science students often fail to think to build privacy into projects (as do professional developers!). In one recent study [23], Tahaei et al. asked 20 CS students to describe how they would design a simple mobile app; despite many students having prior coursework in security, privacy features were absent from their designs. Such concerns apply more generally to value-centered topics, including accessibility, even with professional developers [27]. Tang et al. propose activities based on ideation cards to make privacy decisions more actionable for software developers [24]. Our stakeholder framework generalizes beyond teaching privacy and is intended for use across multiple assignments to help students develop long-term skills in reasoning about values and applying them to technical choices. From this educational perspective, frameworks like Privacy by Design [6] also focus on goals that differ from ours.

Stakeholder-based activities have often been proposed as a way to engage students in values-based design. The ImpactCS framework [14] was one of the earliest such approaches: it had students explore technical issues through the lens of values but did not have students actually implement a technical feature while considering those values. The more recent Embedded EthiCS project [11] develops case studies that deploy within technical courses, but these also do not weave deeply into implementation-based assignments. Our approach is unique in putting specific stakeholders in specific conflicts that result in unresolvable design decisions for a feature that students have to implement. We believe this combination of features is essential for getting technically-oriented students to critically apply values-based reasoning to system building.

Kohno et al. presented a framework to guide ethical reasoning in the context of security-related decisions [12]. Their work supports having conversations about ethics through different philosophical lenses. Their focus is on arriving at a decision, whereas ours is to help students connect values-based tensions with implementation choices during the earlier years of a university education. A guide such as theirs might be more useful after a student has learned to connect value tensions and technical decisions through work such as ours.

Studies have explored how university students perceive social media privacy from a user perspective [21]. We expect that such experiences might impact how students make design choices, but wrestling with personal privacy behavior is fundamentally different from deciding how to implement privacy-facing features for others.

3 Our Learning Goals for Teaching Privacy

Conceptually, privacy is a value that is commonly cited as the rationale and yardstick for determining which data protection measures are adequate or desirable. Our first goal is to help students **recognize and understand** the different facets of privacy and appreciate the fact that those interpretations of the value of privacy are contentious (cf. Section 2).

In the context of courses for computer scientists, privacy is often a broad term that can include topics such as cryptography, regulation, and related security mechanisms. Our work takes a narrower focus, contextualized to the design and implementation of specific systems with privacy-facing components. In practice, the function privacy performs—the goods it protects, the harms it prevents, the capabilities it enables—is context-specific. In order to make implementation choices appropriate for a specific context, students must be able to reason and make nuanced judgments about the concrete value of shielding and revealing personal data in that context. For example, privacy claims may shield individuals from scrutiny where transparency and accountability are called for as well; they may conflict with notions of justice and threaten the integrity of public records for researchers and historians. If we are to train students to design with privacy in mind, they need to identify these tensions and justify their decisions on how to navigate them.

Beyond teaching students to reason competently about the nuances of privacy as a concept, our second goal is to enable them to **translate their reasoning to concrete technical decisions** they make as engineers. To that end, they must recognize how different technical implementation choices available to them relate to different aspects of privacy and to what extent they satisfy their interpretation and their judgment on the trade-offs with other related values.

These two learning goals go hand in hand: in order to make competent technical choices, students must make competent judgments on how to interpret privacy in the broader context

of a technical application and its real-world context. We came up with a new approach because we made multiple unsuccessful attempts to design assignments that would meet these goals while adequately engaging students in the non-technical side of such work. The result is a framework for designing assignments that deeply intertwine discussing value-rich concepts with the technical implementation work that students (for better or worse) associate with computer science courses.

4 A Stakeholder Framework for Designing Technical Assignments

How can we engage computer science undergraduate students in understanding privacy and its tradeoffs? Computer science students are, unsurprisingly, focused on learning computer science. An assignment or abstract discussion about the values underlying privacy runs the risk of appearing unrelated to the technical skills that students expect to practice in computer science courses. This is less an issue of students being closed-minded and more a reflection of well-understood principles of education and learning: learners need to see ideas in the context in which they are expected to use them [9].

We propose a high-level framework for designing assignments that engage computing students in connecting technical and value-based concepts. The key objective of the framework is to make abstract value-laden concepts tangible and relevant to concrete technical decisions.

Under this framework, a value-enhanced technical assignment has the following components:

1. A **technical problem** that students perceive as relevant to their studies, independent of any consideration of a value-based component.
2. A **specific value** (such as privacy) that would be exercised or impacted in a real-world version of the technical assignment.
3. **Background information** that might be needed to understand the value in a technical context.
4. Two or more concrete **scenarios in which pairs of stakeholders hold conflicting views** of how the technical artifact under design should express the value.

The stakeholder conflicts represent different interpretations of the value and related concepts in a concrete and tangible way, making it easier for students to recognize their relevance and relate them to the technical choices they face.

The assignment itself asks students to design and/or implement a solution to the technical problem while weighing the conflicting concerns of the stakeholder pairs. Students must describe the design, discuss how it does—or does not—satisfy each stakeholder's goals, justify their decision based on their understanding of the value, and implement their design.

4.1 Skills Within a Stakeholder Assignment

A stakeholder assignment requires students to engage five separate skills:

1. detecting the nature of the concrete conflict (e.g., protecting a home address versus enabling research);
2. abstracting the conflict to the higher-level *values or principles* that are in play (e.g., personal privacy versus public interest in preserving the historical record for research);
3. justifying a technical decision in terms of the values or principles (e.g., redacting the username in posts is a way to protect personal information);
4. choosing how to implement that technical decision in a concrete system, when the data structures, APIs, or resource constraints may not fully support the chosen strategy (e.g., there's no easy way to scan all posts that might contain leaked sensitive information); and
5. recognizing and articulating the limitations of the chosen strategy (e.g., redacting usernames in posts doesn't protect privacy all that much if the user's identity can be reconstructed through contextual information).

The first skill is a form of reading comprehension; students who take the assignment seriously should demonstrate this with ease. The second skill should be largely straightforward for students who have been exposed to the scenario topic as part of their daily lives. The remaining skills are more interesting. The third skill gets to justifying technical design choices, which is a key professional skill in and of itself. Our approach is rooted in the assumption that computing students do need to learn to think and communicate beyond the level of code, but often are not required to do so.

In the fourth and fifth skills, students must confront the nuts and bolts of how societal constraints interact with building computational systems. Ideally, this is where students would confront that the data structures, APIs, and algorithms in a system influence what sorts of policies are feasible to implement at all. We hope that with repeated exposure to activities like this, students would develop strong skills for building robust, fair, and privacy-preserving socio-technical systems.

5 Assignment: Right to Be Forgotten

Our study explores a specific instance of the stakeholder framework. The technical component of this assignment was to implement a privacy-compliant key-value store designed to back a social network (similar to TAO [4] and memcached [16] at Facebook). Students had to implement a function to delete information from a social media system inspired by the GDPR's clauses on "Right to Erasure" (aka, "Right to be Forgotten", Article 17) and "Right of Access by the Data Subject" (Article 15). Figure 1 provides an overview of the assignment, as well as the collection of key-value pairs that comprise the system (the full handout is at the link in the Open Science section, section 12).

Details of The Technical Problem. Students had to implement a function called `GDPRDelete` with the signature:

```
bool GDPRDelete(std::string& user_id);
```

You'll be working with a specialized key-value store for a new social media platform, Tweeter. On Tweeter, users can choose their usernames, write posts that appear on their profiles, and respond to other users' posts (which appear on both users' profiles). Tweeter has users who are in the EU, so it must comply with the GDPR's right to access and the right to be forgotten. In Tweeter's database, there are five kinds of key-value pairs:

Key Value Pair Structure	Example Return Value
user_id → username	"user_14" → "malte"
post_id → post content	"post_59" → "Hello, Tweeter!"
all_users → comma-separated list of user_ids	"all_users" → "user_13,user_14,user_160,"
user_id_posts → comma-separated list of post_ids that user has posted	"user_14_posts" → "post_59,post_1,"
post_id_replies → comma-separated list of post_ids that respond to post	"post_59_replies" → "post_60,post_61"

Even though there can be multiple users with the same usernames, every `user_id` *is unique. The same goes for posts: even though there can be multiple posts with the same text, every post has a unique* `post_id`*.*

You're in charge of making a decision on how to handle a particular user's request to exercise their right to access and their right to be forgotten.

Figure 1: Assignment excerpt showing key details

The function receives a single argument, which is the user ID of the data subject who is invoking the right to erasure. The function could use this argument to look up data related to the data subject in the key-value store (KVStore), or to find the user's identifier (e.g., "user_1") in other data.

The function does not receive information about the context in which the deletion happens (e.g., what other users' data the data subject might want to delete, or what the other users' views on this are). Students could extend the KVStore with auxiliary metadata that captures relevant context (e.g., special key-value pairs that indicate users who are of special interest, such as public figures), and have their `GDPRDelete` draw on this data. Using such metadata was not a requirement.

We told students that there are many potential implementations for `GDPRDelete`, with no hidden "right answer" for them to discover. Importantly, we did not expect them to come up with decisions that were actually GDPR-compliant, but merely to for their decisions to reflect a reasonable interpretation of their conflicting stakeholders' concerns. Students had to implement some sort of deletion in response to the user request (their implementation couldn't just ignore or reject the request), even though their implementation might not satisfy all parties; we designed the pairs such that there was no way to satisfy everyone. Students submitted a written description of their design, answering the following questions:

1. "Who was your stakeholder pair? (<1 sentence)"
2. "What kind of delete did you implement and why? Explain your decisions as if you were reporting to an ethical auditor who will determine whether your design is justified. Highlight and explain what you think is the most compelling reason that supports the specific kind of deletion you've implemented (1–2 short paragraphs)"
3. "What are the shortcomings of your implementation? Who are some short term and/or long term stakeholders (beyond the ones we've asked you to consider) who could be adversely affected by your decision? (1–2 short paragraphs)"
4. "How might your approach to this assignment change if you were asked to consider the interests of other stakeholders not mentioned in the scenario? (1–2 short paragraphs)"

Background information. We gave students links to the GDPR and CCPA, to refer to as needed to provide background information on the concept of a "right to be forgotten".

Stakeholder pairs. The assignment included five stakeholder pairs, shown in Figure 2. Each student was randomly assigned to design against one pair, but had to discuss their design in the context of some others (of their choosing) as part of a written response (see question 4 above). The Kirby pair concerned a congressperson with regrettable posts from their college days versus a citizen-advocacy group. The Breisand pair concerned a celebrity whose personal information was leaked online versus a researcher who studied her appeal. The Blimp pair concerned divorced spouses with a prior financial dispute. The Yoline pair concerned a film celebrity who had made an insensitive post versus a vengeful ex-spouse and a fan group. The Bleat pair concerned a collegiate athlete who was accused (but not charged) of sexual harassment versus a group working to improve campus climate.

The pairs were chosen to highlight specific values which (instrumentally) justify claims to privacy, such as physical safety (Breisand), the ability to redefine oneself and move beyond one's past views and actions (Kirby, Blimp), and the ability to protect one's reputation (Kirby, Yoline, Blimp, Bleat). The pairs put these in tension with competing values such as preservation for research and the integrity of historical records for public and private figures (Breisand, Blimp), transparency and accountability (Kirby, Yoline, Blimp), as well as questions around the presumption of innocence (Bleat). The stakeholder pairs were designed such that some of the value

conflicts involved well-known individuals who are already in the public eye, while others focused on ordinary people. We did this in order to prompt students reflect on whether individuals' claims to privacy and competing values differs depending on their public status. Section 7.1 reports on how students perceived the tensions within the pairs.

Grading. Grades were based on both the implementation and the analysis in the written response. The implementation had to match the written description (both in the code design and in the actual behavior when running the program). Written description grades considered whether students provided a comprehensive assessment of the context and competing claims, a clear justification of their design decisions against the assessment, and whether competing claims were weighed in a nuanced way. Specifically, students were told:

> *"A good response identifies the legitimate claims that each stakeholder may have, explains why those claims are important, and compares the importance of both claims. It provides concrete reasons for (fully or partially) prioritizing or rejecting individual stakeholders' claims and how those trade-offs are reflected in the chosen implementation of the right to be forgotten. A good response, importantly, also touches upon the limitations of those choices."*

The grades themselves do not add interesting context to our analysis, so we don't discuss them further. What students were told to expect about grading is relevant, however, for understanding students' approach to the assignment.

6 Research Design and Methods

This assignment was given in a second-year (sophomore level) introduction to systems course at a highly-selective university in the USA. Most students were in their first or second year (freshmen or sophomores), and nearly all were majoring in Computer Science (or perhaps Engineering). The assignment came towards the end of the course. The data in this paper are from the Spring 2023 offering, which had 230 students (of which data from 100 were analyzed for this paper). The course instructor is one of the authors on this paper. Different authors designed the stakeholder framework, and a third subset analyzed the data. IRB and ethics procedures (e.g., student consent to being quoted) are described in Section 11.

6.1 Research Questions

Our research questions explore students' work relative to the learning goals presented in Section 5 as well as the general skills described in Section 4.1. In addition, we were curious whether students would indeed implement different deletion strategies for different stakeholder pairs, which would indicate that the design was effective at getting at the interaction between system design and societal constraints. To that end, this experiment centered around the following research questions:

RQ1 To what extent did students exhibit the skills associated with stakeholder assignments (as defined in Section 4.1)?

RQ2 How did students' approaches to implement deletion vary across stakeholder pairs?

RQ3 How do students appear to understand privacy based on how they justify their decisions?

For RQ1, Section 7 presents our analysis, organized around the concrete skills from Section 4.1. We address RQ2 specifically in Section 7.2, and RQ3 in Section 8.

6.2 Analysis Methods

The course instructor anonymized students' assignment submissions before sharing them with the analysis team. That team uploaded students' de-identified responses into ATLAS.ti and proceeded to do manual (as opposed to AI-assisted) thematic analysis [3] on the data. Two of the authors read samples of the student responses and developed initial codes. These same two authors then randomly selected 20 students who had worked with each pair (for 100 students total) and coded the remaining data together. We chose to code together rather than to do separate coding with inter-rater reliability due to this being the first time we had worked with data from such an assignment, which led to rapid evolution of the codebook. Our codebook has more than 280 low-level codes across 10 high-level categories (the complete list of categories and codes appears in Appendix A[1]):

- *Implementation*: Actions students took on elements of the Tweeter data stored, such as deleting, obfuscating, or anonymizing (supports skills 3 and 4).
- *Tension*: Values or ideas students identified as conflicting with one another in their justifications, either explicitly or implicitly (supports skills 1, 2, and 3).
- *Values*: Values or ideas that students mentioned in justifying their implementation, but that had not been described as in tension with other values or ideas. These tags apply to all mentions of a value, whether endorsed, rejected, or just referenced (supports skills 2 and 3).
- *Implementation Changes*: Categories of changes students would make to their implementation given a different stakeholder pair or more time, resources, or skill (supports skill 4).
- *Privacy*: Tags applied based on students' descriptions of what privacy looks like or entails, e.g., "physical safety," "social media anonymity," or "avoiding reputational harm" (supports RQ3).
- *Privacy Spectrum*: Tags applied based on whether the students' belief in privacy and implementation privacy were contextual or absolute (supports RQ3).
- *Limitations*: Tags applied based on a subset of the implementation limitations identified by students, namely their

[1] a more detailed version of the codebook is on the project website: https://responsible.cs.brown.edu/research/stakeholder.html

Data Subject: Congressperson Kirby is currently the elected congressperson for Rhode Island's first district and is running for reelection in the current election cycle. Recently, some of their tweets from when they were in college have resurfaced. These tweets were written 10 years ago concerning a pandemic that occurred at the time. While the congressperson has issued apologies, this has not been enough to stop the ongoing discourse from users all across the political spectrum. Because of this, the congressperson has made a request to exercise their right to be forgotten. They wish to delete their account, which they hope will delete their original tweets and related public discourse (other users quoting or rewording the original posts) about the controversy.

Opposing stakeholder: The Freedom House Advocacy Group is concerned with government accountability and transparency. They oppose Kirby's attempt to remove evidence of their controversy from the past.

Data Subject: Sarsra Breisand is an American singer, actress and director. With a career spanning over six decades, she has achieved success in multiple fields of entertainment, and is among the few performers awarded an Emmy, Grammy, Oscar, and Tony (EGOT). Recently, a paparazzi reporter leaked information about where Sarsra Breisand lives. Despite her attempts to hide this information, she draws more attention to this leaked information. As a desperate final attempt to protect her privacy, Sarsra has made a request to exercise her right to be forgotten and delete her account. Sarsra understands that this will erase her social media profile, but hopes that this will put an end to the interest in her whereabouts.

Opposing stakeholder: Beth Abraham is a historian who has begun studying what she has named the Breisand Effect. She has released several research papers on this psychological and sociological phenomenon and is in the process of writing another.

Data Subject: Frank Blimp is a recent divorcee who doesn't have custody of his child from the marriage. Falling into hard times, Frank missed a couple of child support payments, for which his ex-wife, Marge, has called him out on Tweeter. Even though both Frank and Marge have private accounts, these posts are visible to Frank's friends and family. Since this tweet, Frank has been able to repay the missed payments and is financially stable enough to continue making future payments. He's asked Marge to take down her tweets shaming him, but she has refused. Frank now requests his data to be erased in the hope that this will also make posts mentioning him disappear.

Opposing stakeholder: Marge Blimp is Frank's ex-wife who has main custody of their child. After he had missed multiple child support payments, she resorted to using Tweeter to call him out. She wants others to be aware of Frank's past behavior.

Data Subject: Angel Yoline is an up-and-coming actress who is excitedly promoting her new film on Tweeter. However, a while ago she posted views on Tweeter that characterize working class people in a negative light. Her former partner, Brad Schmidt, has recently highlighted these problematic views as a means of getting revenge on Angel. She now wishes for her own post, Brad's post, and the public discourse about the controversy to be deleted.

Opposing stakeholder: Film enthusiasts have been excited about the new film that Angel Yoline is starring in. When an account exposes something that Angel Yoline said about working class people, there is a surge of new posts discussing her controversial opinions.

Data Subject: Matt Bleat is a senior in high school who has just committed to an Ivy League University after being recruited to their track and field team. Recently, there have been posts from Matt's high school classmates calling attention to former allegations against Matt for sexual harassment. The school determined that there was insufficient evidence to support those allegations and declined to take further action. Matt invokes the right to erasure, expecting that it will remove his posts, but also remove posts that mention hashtags related to the controversy.

Opposing stakeholder: The Center for Changing Our Campus Culture believe that posts highlighting the allegations against Matt should remain on Tweeter, and Matt should be prepared to deal with the possibility that the university could discover his tweet and rescind their offer of admission.

Figure 2: The five stakeholder pairs. Each has a data subject and an opposing stakeholder.

lack of ability to perform a *complete* deletion (supports skill 5).
- *Stakeholders*: Tags applied when students discussed non-human third-party stakeholders or facets of Tweeter itself as a stakeholder, indicating considerable depth of consideration in the stakeholder analysis section (supports skills 4 and 5).
- *Public Figures*: Tags applied when students argued that public figures (either celebrities, public servants/politicians, or public figures in general) should be handled differently when deleting Tweeter data (supports skill 3 and RQ3).
- *Second-order Effects*: Tags applied when students identified a second-order effect of their implementation (supports skills 3 and 5).

The code count is high because we tried to respect students' wording nuances. For example, there are separate values codes for "societal good" and "societal progress", as well as separate codes for "journalism", "research", and "preserving content". Tensions are pairs of values, leading to explosion in those code counts. For analysis, we focused on specific collections of codes, as described in Section 6.2.

We produced charts and summary statistics on codes using a combination of built-in features of ATLAS.ti and Google Sheets (after extracting the coded data to CSV). The two authors in charge of analysis reviewed each other's scripts and generated results to check for accuracy.

7 Evidence of Skill Development (RQ1)

We now review the skill development that the framework design is intended to support (from Section 4.1).

7.1 Identifying Tensions and Factors (Skills 1 and 2)

Figure 3 shows the values that students mentioned as being in tension when describing their implementations (due to the large numbers of individual tension: X-vs.-Y codes as shown in Appendix A, we show the high-level values without the specific contrasting value). As we had hoped, students raised different values across the pairs. The profile for Breisand (the singer with a leaked address) particularly stands out for its differences relative to the other pairs.

Figure 3 captures only those values that students explicitly placed in tension with another value when describing their implementations. Some students mentioned values without explicitly putting them in tension (despite the assignment instructions). Others raised other values or tensions when describing the shortcomings of their work or how things would change if they were asked to implement towards a different stakeholder pair. Figure 4 shows values that students raised in any part of the assignment, whether individually or in tension. Even there, the Breisand profile differs from the others. Certain values and ideas are also more strongly tied to some pairs (such as Accountability for Congressperson Kirby) or the low public interest in the Blimp family dispute.

Skill Summary. These two charts indicate that the stakeholder pairs influence students' thinking about when to honor a stakeholder's request to delete information from Tweeter. Overall, 89 (out of 100) students demonstrated the ability to detect the concrete conflict (skill 1). The other 11 did not put any values in tension when discussing their implementation (3 Kirby, 1 Breisand, 2 Yoline, 3 Blimp, and 2 Bleat). Fewer students than we had hoped, however, abstracted to the higher-level principles that we might have liked to see (skill 2). Across the pairs, 24 students cited only vague tensions like "Right to Delete vs. Free Speech"; the other 76, in contrast, raised values such as accountability, reputational harm, who controls data, and public discourse, aligning more closely with our hopes for skill 2.

7.2 Deletion Decisions (Skills 3 and 4, RQ2)

In the 2022 offering of the course—before we included stakeholder pairs—11 students (out of 121) did an optional assignment to implement `GDPRDelete`. Of those, one student deleted the posts but left the `user_id` intact, six deleted the `user_id` and left the posts intact, and four deleted both posts and `user_id`. These simple approaches showed little engagement with nuances of privacy.

In contrast, the students who worked with the stakeholder pairs pursued more nuanced approaches. Figure 5 summarizes how students' decisions about what to delete varied across the pairs. For each pair, the majority of students deleted the post as requested by the primary stakeholder, though this was least frequent in the case of Congressperson Kirby. Students were least likely to delete the stakeholder's ID from the system in the case of actress Yoline. Students were most willing to delete replies to the original post for Breisand's leaked address case, and least willing for the Congressperson. Nearly every student (with a single exception for some pairs) who didn't delete posts, IDs, or replies instead tried to obfuscate the offending contents by anonymizing names or replacing content with "redacted content" messages.

The differences in approaches across pairs suggests that the stakeholder framing succeeds at getting students to consider how technical decisions interact with contextual nuances. This was a major high-level goal of this assignment.

Students' justifications for their implementations raised various nuances. One student who worked on the Breisand pair remarked:

> *"This was an especially tricky pair to grapple with, because most of the content that Breisand would probably want to delete was content that came from other accounts (people leaking her address). I ended up just opting to delete her profile itself, as that's what she asked for specifically so as not to draw attention to herself. However, I did not decide to delete her posts, as she didn't explicitly request that and I assume that, although they would still be stored in the database, they wouldn't be publicly*

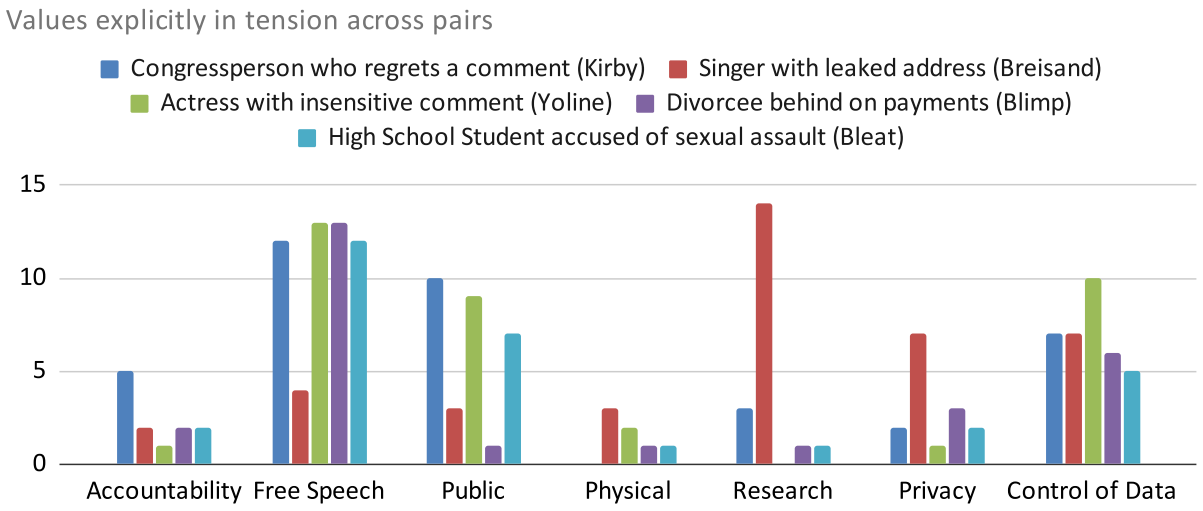

Figure 3: Concepts noted as in tension within each stakeholder pair.

> *viewable as there is no longer a profile for them to be attached to. However, they would still exist somewhere in Tweeter's database so that they might be accessible to the journalist for her research at a later time."*

This student notes the tension between Breisand's request and the data of concern being owned by another user. The student chooses to not delete the delete the concerning post, citing both a technicality ("she didn't explicitly request that"), and that the approach achieves the desired effect ("they wouldn't be publicly viewable"). The solution also preserves potential access for the researcher (termed "journalist" by the student) because the posts remain in the database.

Students who worked on the divorced spouses (Blimp) case generally saw it as a tension between one's right to delete or control their own data and free speech. Many students sympathized with Frank, sometimes wrestling with how to create a general policy while being fair to him, e.g.:

> *"While I personally stand with Frank, giving Frank the power to delete another person's post–even if he is mentioned in it–could have dangerous ramifications."*

Fourteen students discussed wanting to let Frank preserve his dignity or avoid reputational harm, especially since the issue appeared to be in the past. Unsurprisingly, these issues are not raised in the Breisand case. Curiously, three students cited preservation of the historical record as a reason to preserve Marge's posts, though none cited what sort of historical or research question might matter with the Blimps (unlike in other pairs, which featured public figures). We suspect that these students were viewing "history" as general record-keeping.

Students who worked with Congressperson Kirby's case generally cited the specific value of public accountability (whether or not they deleted the posts), while others spoke about the more abstract idea of maintaining public discourse. The idea of maintaining discourse also featured in students' discussions of whether *replies* to the posts should remain:

> *"In my initial implementation, I deleted the name of the user, their posts, as well as any references to the person found in comments or the wider site at large, but I quickly found out that this may end up accounting for complete censorship of any discussion about a political figure, which infringes heavily on both the freedom of speech and the press. With this knowledge, I backed down my implementation by a decent amount eventually deleting the name of the user and their own posts, but keeping the post IDs present. [...] I realized that while anonymizing replies to a post by removing any mention of the person's name may aid in preventing a previous post from being uncovered, it could also be abused, as people could write out a laundry list of people's names in a reply and whichever one got removed was the person that elected to delete their data."*

Such arguments illustrate how students view the same content through different values. While many students (across the scenarios) treated replies as data owned by others (and hence ineligible for deletion), some working the Kirby case instead valued replies as elements of public record. Some students navigated the public-record tension by limiting the set of posts that would be deleted to those made by Kirby themselves, which provided a right to be forgotten without sacrificing

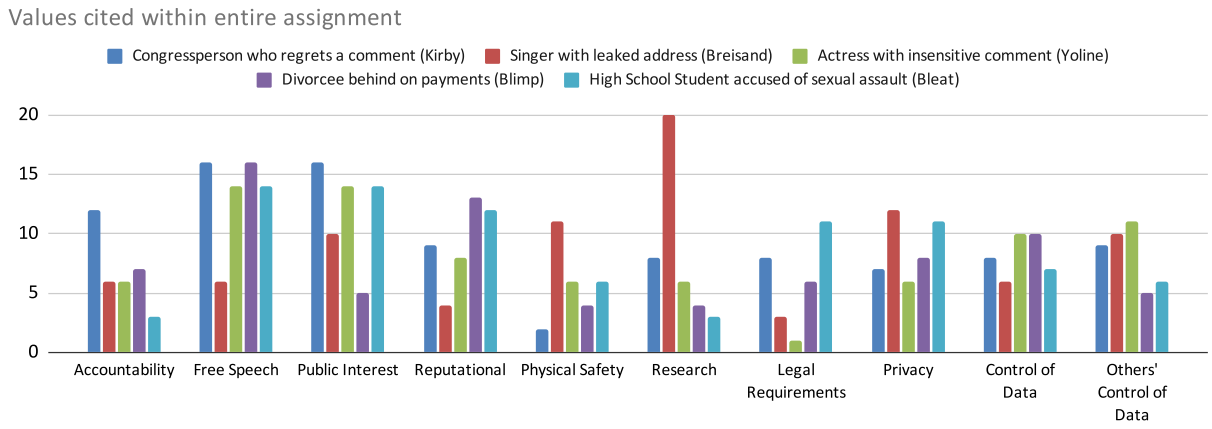

Figure 4: Values cited per pair (individual students may be counted under multiple values).

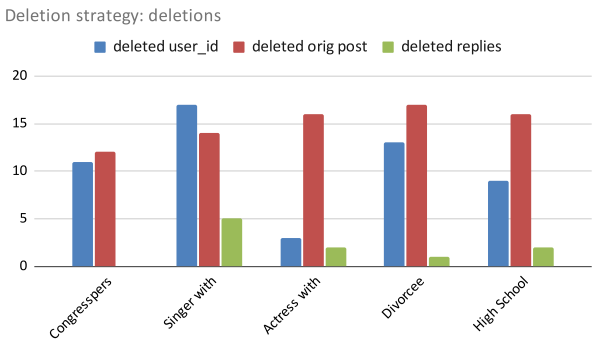

Figure 5: Which information was deleted, by stakeholder.

accountability.

Other students rejected this idea through an implicit value of transparency:

> *"I would not have allowed, for example, Congressman Kirby to only request that certain messages be deleted. I believe that when it comes to your digital footprint, you should either be all-in or all-out. After all, the spirit of the data protection laws is not to allow politicians to circumvent legitimate criticism from shady activities they may have engaged in."*

One student even suggested that deleting the account would not be good for Kirby:

> *"Furthermore, my design is not deleting the user's account, as having the account stay on Tweeter allows Congressperson Kirby to stay relevant in regards to any new information about his candidacy and to the media for publicity, which constitutes government accountability and transparency."*

Another student argued similarly for Frank Blimp:

> *"With this implementation, the friends and family who see Marge's tweet will still definitely know who it's about, even after Frank's data is deleted. The fact that Frank has repaid the missing payments and is now on schedule is totally lost, and most problematic of all, Frank, in hoping to remove Marge's tweet, has removed his tweet which clarifies the situation and defends him. He is arguably then in a worse place with this deletion of data. That is why it's important to clarify the nature of the deletion to the users, noting that posts mentioning them or replying to them will not be deleted."*

To the extent that students gave nuanced comments on the Yoline case, they tended to be about the rights of public figures, with some students arguing that being a public figure warranted more privacy and others arguing that it warranted less. With the Bleat case, nine students framed the discussion in terms of public interest or the public's right to know. Only three students cited Bleat's anonymity as a factor. One student felt that posts about Matt Bleat needed to be maintained for potential scientific research about Tweeter itself, which we see as a striking expansion of the idea of "research" as it interacts with individual privacy.

Considering Second Stakeholder Pairs. If students are taking the nuances of individual stakeholder pairs seriously, we might expect them to describe different design choices had they been assigned different stakeholders. The assignment specifically asked students about this, and indeed over a quarter of the students indicated that they would have changed some aspects of their designs. By and large, the sorts of differences that students described overlapped the decisions made by other students who had worked with the other stakeholders.

Most of the comments regarding changes referred either to having different standards for public figures or different standards for personal information that could lead to a user being personally harmed. In other words, considering a second pair encouraged students to *further refine* the values they originally cited based on context. We see this as particularly powerful since students did not necessarily cite those refinements when considering a pair for the first time.

7.2.1 Students Reference Free Speech Liberally

Across the pairs, 66 students explicitly invoked the principle of free speech when determining what data should be deleted. This principle usually arose when considering the opposing stakeholder or others who may have engaged with the Tweeter threads that contained disputed posts. Students who raised free speech generally expressed it as a right to express opinions, rather than as a guarantee that the *government* may not (pass laws to) restrict expression (the definition in the USA). Sample quotes include:

> *"This would also prevent potential infringement on free speech concerns for users in general who reference the event or the data subject."*

> *"I don't think it's reasonable for a complete deletion of other user's posts whenever one wants to delete their account, since this infringes on the free speech of other users."*

> *"My choice to remove/redact their name from other people's posts without removing the posts entirely were based on the reasoning that dropping the posts would infringe more on user's freedom of speech but just redacting a name from their post would still allow them to express important ideas while not allowing them to target or call out the individual requesting to have their data deleted."*

A few students were more careful in their wording, and instead used the term "free expression" rather than "free speech". While students may have been speaking loosely when using the term "free speech", we suspect that many actually aren't clear on the scope of this term, and instead are conflating it with a broad concept of "censorship" (a term which eight students used explicitly). Regardless of wording, these students observe a relevant aspect of the cost of privacy protection.

7.2.2 Students Value Preservation for Research

Students frequently cited the value of preserving data for research or the historical record when making design decisions. While the Breisand scenario raised this tension explicitly, it was implicit in the Kirby case (based on politics and accountability); students also cited it in the Blimp case in connection with personal accountability (or in one case, the potential that Frank Blimp might want to run for office someday!).

Nine students invoked the specific GDPR exception clause regarding the right to be forgotten as it pertains to research records when deciding whether to delete or merely hide posts. Legally, these exceptions have been interpreted as applying to data preservation by established organizations, not the mere potential that an individual researcher may someday want access to the data. This is not a nuance that students would know and not one we expected them to grasp through this assignment, but it may be an issue worth raising more directly in future iterations of the assignment.

Some students raised finer-grained nuances on values in deciding how rights to research should interact with privacy:

> *"While I do respect that Breisand's desire to maintain private information is valid, as a public figure and also now an important example in a scientificly [sic] studied phenomenon, she does not have further rights to data deletion."*

> *"While some might argue that the data is required for research purposes, I believe that preserving the data would not only harm Sarsra's privacy, but also goes against research ethics."*

> *"I couldn't make the determination that the adversarial party could not continue with their research without the data, or that their research on Sarsra Breisand was critical enough to warrant permitting some preservation of and privacy-infringing discussion directly surrounding the account and posts of Sarsra."*

7.2.3 Removing, Hiding, or Archiving Posts

Several students sought to hide posts, rather than remove them entirely, as a way to resolve tensions between privacy and future needs for access. Hiding posts typically meant that students would disassociate the posts from the `user_id` of their author. In this way, the user account could be deleted while the posts themselves remained accessible for researchers or law enforcement. In some cases, students argued that this provided value to potential victims of crimes:

> *"From a personal perspective (jumping out of the stakeholder's demand), I don't consider a removal on such extent is appropriate. First, despite the low possiblity [sic], there might be potential evidences that lead to serious misdemeanors, traces to unrealized crimes and other activities that demands public attention and judicial interference. The victims of publicly-unawared [sic] offences, therefore, could be seriously impacted."*

Some students considered the option to preserve data (e.g., as legally required to preserve evidence) in separate data stores, rather than leave them disassociated in the core KVStore.

"If the FBI had told me that they want to investigate Sarsra Breisand because she committed a crime, I would put the data into a file for them before I deleted it."

Only some students picked up on the temporal issue that some crimes or runs for political office could be future events, leaving a potential hole in their reasoning about the value of data for legal purposes.

Some students discussed whether hiding posts would meet the requirements of GDPR's right to be forgotten, though most didn't consider that question. If the assignment were intended to have students *understand* GDPR (as opposed to be aware of it beyond a surface level, as it was in our case), an additional activity that had students rate their strategy against the detailed GDPR requirements would likely be needed.

7.2.4 How Students Consider Third Party Stakeholders

Although the scenarios were designed around a single opposing stakeholder, students frequently invoked rights and responsibilities for third parties in justifying technical decisions. We both expected and hoped to see this: it shows engagement with the assignment and also demonstrates that students are connecting larger concerns to the assignment beyond its strict text. In these cases, however, students often stopped talking in terms of *values* as opposed to potential *experiences or impacts* that needed to be considered. Examples of the latter in the Breisand cases include neighbors (who could be harmed by undue attention to her home), future music historians "who would be baffled by the lack of a digital trail", and local services that might face additional demands from tourists. In the Blimp case, a student cited "the interests of authorities such as child welfare agencies."

We have noticed a related phenomenon in other studies we've done (not published) in which students emphasize technology's impacts on significant societal or natural systems, rather than focus on the context presented in a specific assignment. We did have a student who described the costs of data storage and its environmental impact as a potential factor (deemed not significant based on their analysis) in deciding which posts to retain for research purposes. Unlike in other assignment designs we've tried, however, this environmental factor was firmly tied to the specific implementation decision of `GDPRDelete`. We hypothesize that having students engage with values in the context of having to implement a *specific* technical feature forces students to ground their thinking on these larger concerns. Importantly, this doesn't imply that we want students to ignore abstract concerns like the environment or the global economy, but rather that intertwining stakeholder considerations with technical implementation prevents students from *only* invoking such big-picture ideas.

7.2.5 Alternate Approaches to `GDPRDelete`

For skill 4, students have to decide how to implement their policy in the specific data structures and constraints provided to them. For our students, this meant navigating the data organization as per the structures in Figure 1, with a `GDPRDelete` function that simply takes a `user_id` as input.

Most of the interesting comments here related to whether calling the `GDPRDelete` function is the only mechanism by which data could be removed from the database. Some of these were technical, while others were legal. For example, multiple students raised recourse to legal regulations beyond Tweeter's policies (e.g., libel laws) that could be pursued to handle exceptional cases that fell outside their own implemented deletion function. Some mentioned code modifications that might be needed if this were the only function that could delete from the KVStores:

"For one, if a user was maliciously hacked, I'm sure Tweeter would have a system to authenticate this in some way. If this information was passed in (say a boolean on whether the user is deleting their account because they were hacked or the `post_id` of the hacked post) I would consider deleting the surrounding discourse around that subsequent tweet. However, I'm not even sure that I would do this as it still perhaps infringes on other people's right to free speech to comment on an event they witnessed."

The assignment handout did not discuss whether their function would be the only mechanism for removing data, leaving students to make these interpretations for themselves. An instructor who wanted students to use one interpretation versus the other would need to specify this more clearly.

Skill Summary. All students met the letter of skill 3: they wrote a justification of their design decision. In 7 cases (4 Breisand, 2 Blimp, 1 Bleat), the justifications were vague (roughly, "I balanced the needs of both Frank and Marge"), but they were still present. Students invoked a variety of high-level concepts (Figure 4). For skill 4, students also made varied design choices (Figure 5) on what to delete, as well as varied decisions on whether to delete or obscure sensitive information in posts.

All in all, the charts in Figures 3 and 4 support our hypothesis that working with stakeholder pairs can influence which values students bring to bear on thinking about their implementations. Combining these with Figure 5 suggests that tying stakeholder values to implementation decisions within the assignment does impact how students approach the otherwise technical decision of which posts should be deleted.

7.3 Recognizing Limitations (Skill 5)

Students cited a variety of ways in which their strategies were limited, either in the policy design itself, the practicalities of the implementation, or the ability of a purely technical approach to address the problem. We assume that many of these insights stem from prior knowledge or personal experience, rather than anything inherent to the assignment design.

Deletion is an Illusion. Over half (57) of the students noted that deleting posts from Tweeter couldn't guarantee that the offending information would disappear. Students mentioned things akin to re-tweeting, people taking photos of Tweeter posts or copying the information to share on other platforms, and the media as ways in which the data in question could not be contained. Students also mentioned other ways for the concerning information to be shared or obtained, such as via interviews with people with knowledge of the situation.

> *"In effect, all of the efforts of the deletion could just be moot – someone could start an account and just post a screenshot of every single tweet that person made. This however is a greater question of legality that would probably need a lawyer to understand who exactly has the rights to those tweets. Were the rights to belong to the person who wrote the original tweets, I would probably extend the implementation to delete direct screenshots of those tweets as well, as in that case, tweeters would be using another person's content without their consent (stealing)."*

In general, we want students to understand that technology alone cannot manage societal values. These comments indicate that many students already think along these lines, though we would have liked to see this from more students.

Full Redaction is Expensive. We were pleased that some students considered whether the policies they defined could actually be implemented within reasonable resource constraints (computational or programmer time). For example:

> *"The only change that I can imagine being ethically valid would be changing or redating [sic] the location in the related tweets, thus protecting the specific aspect of her information. However, that would not be a generalizable deletion method and I believe it would place an undue burden on a hosting platform which the GDPR would not consider it liable for."*

> *"So, in certain cases, I would reject the request for deletion for the greater good of society. However, this would require extensive planning because there has to be a way to identify which one of the users on Tweeter who are requesting to delete their data is a criminal."*

> *"There could be replies to the replies, and maybe duplicates of Sarsra's posts and replies to those that also expose more information about Sarsra – unfortunately, there is just too much information on the internet for programmers to keep track of."*

Ideally, we would like to see such discussion from all students who proposed approaches that were more nuanced than removing entire key-value pairs. In our experience, students can be overly attuned to performance analysis, to the point that they value it over societal concerns. It would be interesting to know whether students skipped this part either because it didn't occur to them to do so or because we hadn't explicitly asked for it in the assignment.

Handling Temporal Matters. Students wrestled with how to handle personal opinions or behaviors that change over time, both in terms of values (skill 3) and in terms of technical implementation (skill 5). Some students proposed strategies that would consider the period in which a post was made:

> *"I implemented a selective delete strategy for the GDPR deletion request. Specifically, I allowed the admin to delete only the tweets that were listed 10 years ago about the pandemic, as these were the tweets causing controversy. However, instead of deleting all of the congressperson's data, I introduced a new field called "metadata" to keep a disclaimer about the deletion for logging purposes. [...] The admin is a new stakeholder that is introduced in this model and would need to decide on which tweets must be deleted on the topic."*

Some students who took such an approach recognized the limitations of blanket rules (which would need different stored information to implement):

> *"One of the shortcomings of my implementation is that it uses the timespan of ten years as an arbitrary measure of whether a communication is still of the public interest. Really, it was that the Congressman was not a figure of the public interest at the time, and was also in college and thus presumably still developing his worldview and education, that might lend him some protection under the GDPR. Thus, it would be more accurate to demarcate a point in time for each user that is a figure of the public interest where their communications became of the public interest. However, that was not possible in the scope of this project because there are not timestamps for each post, aside from the handout stating that his controversial posts were from ten years prior when he was in college."*

The need for metadata arose in other situations as well:

> *"People have the right to know about incriminating information, and our method should have a way of filtering out posts that could be used in legal proceedings. A useful indicator would be some sort of metadata attached to posts to label crucial pieces of information that should not be deleted."*

Corporate Interest. Some students raised new values when they were asked about limitations. In this case, a student cited how they might not actually delete posts in order to support corporate value:

> *"On the flipside, if I were asked to more strongly consider my own company's profit-motive, I might implement a less thorough form of delete so as to maintain the increased traffic from users flocking to the site to discuss Breisand and her address leak. For instance, while I would probably still delete her posts due to GDPR, I might not uncouple her user ID from her username. This would allow other users to continue tagging her account and being aware that this account is linked to Breisand, which would likely result in more traffic on the site and aid my company."*

Using Anonymity. Sometimes, students failed to mention a limitation that we would have liked them to raise. Most students who mentioned anonymizing data, for example, seemed to assume that redacting names would be sufficient to conceal identity.

> *"In order to help the data subject without violating the first ammendment [sic] rights of those who spread the revealing information, if the request were approved, from the post(s) in question, the name of the data subject would be **anonymized**, to protect their privacy."*

The few students who cited limitations to anonymization had worked on the Blimp case. They recognized that the posts might only be read in a small group who would know who Marge was referring to even if Frank's name were redacted.

Skill Summary. As instructed, all 100 students answered the question about limitations in their design. All but 2 pointed out actual limitations, as opposed to saying something like "this is fine". Eleven (11) students specifically raised second-order effects within their limitations. Twenty-three (23) raised non-human third-party stakeholders—such as companies or the entertainment industry—or even the environment due to the cooling needs for additional storage for deleted posts. All in all, students performed better on this task than on skill 2, though we expect this is because they can rely on personal experience and outside knowledge.

8 How Students View Privacy (RQ3)

As described in Section 5, having students reflect on privacy as a construct is one of the goals of the assignment. Half of the students (50 out of 100) explicitly mentioned privacy in their responses (the others framed comments as "Right to Delete" or "Right to be Forgotten"). A few said something more specific about what privacy means or controls. These comments typically invoked one of anonymity, ability to control data, maintaining or recovering reputation, physical safety, sensitive information, or protecting oneself on social media. Physical safety was the most commonly linked to privacy (by 15 students). While only 10 students explicitly linked reputation to privacy, 43 additional students raised reputation as a value that figures into designing a delete function for Tweeter. As such, we expect that more students had conceptions of privacy than those who used the term explicitly.

Six students talked about privacy less as a value, but rather as a legal requirement (via GDPR); four of these students didn't reference any tensions in values when describing their designs. This suggests that some students may be trying to think in terms of procedural requirements rather than engage in the values-based analysis that the assignment tried to encourage. Keeping an eye out for this trend in future studies seems important, especially when working with students with a technical orientation.

We were curious whether students seemed to view privacy as an intrinsic value as opposed to a contextual one. We identified only two students among the 100 who described it as intrinsic (both of them worked on the Breisand pair, though that could be coincidence). An additional four students justified their designs as supporting intrinsic privacy (two each on the Kirby and Bleat pairs). Nearly all of the rest (94 students) made statements that suggested that they believed privacy or the right to be forgotten to be contextual. For the remaining four students, we were unable to determine whether they held intrinsic or contextual views.

9 Lessons Learned and Future Work

Overall, we deem our stakeholder activity a success. It resulted in a range of deletion designs and got students to connect conflicting abstract values to design decisions. It also got students to reflect on the challenges of value-based privacy design in practice: many students noted issues such as the difficulty of locating (much less redacting) all privacy-violating posts, and various mitigating factors (such as the passage of time when controlling for reputational harm) that do not lend themselves to simple technical solutions. We believe that these kinds of learning experiences are what the computing community needs long term if we are to build robust, fair, and privacy-preserving socio-technical systems.

That said, our analysis suggests some limitations to how the stakeholder framework shapes students understanding of privacy-related questions.

Skills over knowledge. Students vary in terms of how much they know about the values embedded in the stakeholder design, and the assignment itself did not provide explicit explanations of those values. Given this deliberate lack of specific input, it may not come as a surprise that students sometimes used concepts imprecisely or drew overly general conclusions. A case in point is the way in which some of them interpreted any form of redaction as an infringement on freedom of speech, as discussed in section 7.2.1. The stakeholder framework scaffolds reasoning skills more than it builds knowledge: students practice justifying privacy-related decisions, taking into account both value-based and practical constraints. While our results indicate that the stakeholder framework helps students understand that privacy is a complex value embedded in

a broader system of related and conflicting values, it does not necessarily convey *systematic or holistic* knowledge about the finer details of any of those values.

For future work, it would be interesting to add a component which challenges students to discover those nuances: e.g., moving beyond an all-or-nothing understanding of freedom of expression and requiring students to reflect more closely on the extent of individuals' entitlement to preserve the literal and contextual integrity of what they wrote. This would add nuance to context-specific trade-offs with conflicting privacy-claims. A multi-step assignment in which students discuss their designs with others (TAs or peers) might help.

Prioritizes instrumental understanding of privacy. Students articulated a variety of reasons why protecting privacy is important. These reasons, such as personal safety and respect for autonomy (ability to develop as a person unburdened by the past), all portray privacy as an *instrument* for protecting or promoting other values. That's a result of the narrative structure built into the stakeholder framework: the stakeholders all care about something for which privacy is either useful or detrimental. Thus, the framework is not intuitively suited to conveying the idea that privacy might be valuable intrinsically—that individuals may have a valid claim to have their privacy protected irrespective of whether that is good in terms of any other values they care about.

This could be rectified naively by adding stakeholders who care about their privacy for its own sake, but we worry that it doesn't provide a very compelling conflict for students who don't recognize privacy as intrinsically valuable in the first place. Given the fact that the intrinsic status of privacy is controversial in the wider literature as well (see Section 2), we don't expect this idea to be as intuitively compelling to students in the same way as the other values the stakeholder framework invokes. This is an important limitation: our anecdotal experience teaching privacy-related concepts explicitly in the setting of technology ethics courses suggests that many students struggle to understand and articulate how privacy could be intrinsically valuable. In the long run, we would want to ensure that the stakeholder framework doesn't reinforce an already one-sided narrative about the foundations of the value of privacy. This suggests a direction for further research: how does exposure to the stakeholder framework change students' pre-existing notions of privacy?

Entanglement with GDPR compliance. In retrospect, a weakness of our assignment framing is that it conflates the general notion of privacy and the right to be forgotten with a specific regulation (the GDPR). This runs the risk of suggesting that the assignment teaches students how to correctly comply with the GDPR, which is not its goal. For example, some of the answers students provided—such as merely removing top-level user accounts but leaving the posts in place (Section 7.2)—are valid for the learning goals of our assignment, but are in contradiction with the GDPR's requirement to remove all data identifiably associated with a user.

The GDPR framing came from an evolution of the assignment over time: it started as an assignment (without stakeholders) to teach students about the GDPR, but we turned it into a more general assignment on privacy. It would have been wise to explicitly state the relationship with GDPR, which merely provides a well-known reference to the right to be forgotten that motivates the analysis. In the next iteration of the assignment, we will reduce the GDPR-related framing, and more explicitly tell students that their solutions do not need to and may not be compliant from a regulatory perspective.

Redesign to Improve Student Justifications. Under the current design, students may not automatically and uniformly recognize and consider the *finer* gradations of conflicting values; they need only consider the position of a stakeholder that reflects those values. We see this in cases where students juxtapose research needs against privacy (Section 7.2.2). It would be interesting to revise the assignment design and ask students to explicitly point out why a given conflict is sufficiently significant to warrant their proposed trade-off with privacy-related concerns or vice versa.

Important Learning Happens through Limitations. Asking students to discuss limitations of their approach (skill 5) suggested impacts that we hadn't expected to see. Observing how students dealt with time-related limitations of their implementations (Section 7.3) provided insights into how students see the nature of the value-conflicts they considered: some thought that they had to choose a side (while acknowledging the value conflict). By contrast, others took a more nuanced stance and balanced different values which in turn required them to *qualify* those values. This tells us that asking students to think about the limitations of their implementation choices may be successful at helping them see privacy as a value in relation and in conflict with other related values. The current design does not uniformly get students to realize that values are complex and that their relative importance can be qualified, rather than requiring uniform responses.

Building Transferable Skills. Ideally, the stakeholder framework will help develop students' skills at considering societal values during technical design. This will not happen, however, after a single assignment. There is a long and robust education literature on the concept of transfer, and what it takes for people to take a concept learned in one problem or context and apply it to another [2]. Transferable learning starts with concrete examples over which learners can abstract general principles about how to approach a particular type of problem. We believe that our stakeholder framework provides a template for designing concrete assignments that support learning socio-technical design. To build from this to transferable knowledge, students will need to do multiple exercises of this style, in multiple contexts, each time connecting to general design principles for socio-technical systems. Consequently, the authors have embarked on a larger effort

to deploy the stakeholder framework across multiple courses. Doing this, and aggregating the findings across instances, is our next step.

10 Conclusion

We presented a stakeholder framework designed to help university students in computer science reason concretely about nuances of socio-technical concerns such as privacy, and to integrate their reasoning with technical implementation work. Our case study with an assignment motivated by the "right to be forgotten" in recent privacy legislation indicates that the stakeholder framework succeeds at making students recognize concrete instances of tensions between high-level, abstract values at play, and to relate them to technical design decisions.

This effort lays the groundwork for a broader investigation into good assignment design around the stakeholder framework and questions about effective deployment of this framework across courses and socio-technical topics.

11 Ethics Considerations

IRB Review. As our study was conducted using data from a homework assignment that was already part of an existing course, it did not require IRB review under our university's rules. We nonetheless took several steps to protect identity and respect students' data.

Data De-identification and Handling. Only the course instructor saw student identifying information; other authors worked with de-identified copies of the submitted work tagged with an anonymous id (e.g., "student12"). The de-identified materials for the study are maintained in our university-managed Google Drive system, in folders that are access-restricted to members of the project team. Only de-identified data were uploaded to Atlas.ti.

Permission to Quote. The instructor sought and received explicit permission (via email) from each individual student who we have directly quoted in this paper.

12 Open Science

This research is based on a course assignment that is publicly available [5]. Appendix A contains the codebook we used for our data analysis. We are unable to release the set of original student work due to privacy considerations.

Releasable materials related to this work (to include future assignments that use the stakeholder framework) are available at: **https://responsible.cs.brown.edu/research/stakeholder.html**.

Acknowledgments

This research was supported by NSF Award 2335625.

References

[1] Edward J. Bloustein. Privacy as an aspect of human dignity: An answer to Dean Prosser. *New York University Law Review*, 39:156–202, 1984.

[2] J.D. Bransford and D. Schwartz. Rethinking transfer: A simple proposal with multiple implications. In *Review of Research in Education*, volume 24, pages 61–100. American Educational Research Association, 1999.

[3] V. Braun and V. Clarke. Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, and K. J. Sher, editors, *APA Handbook of Research Methods in Psychology, Vol. 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological*, pages 57–71. American Psychological Association, 2012.

[4] Nathan Bronson, Zach Amsden, George Cabrera, Prasad Chakka, Peter Dimov, Hui Ding, Jack Ferris, Anthony Giardullo, Sachin Kulkarni, Harry Li, Mark Marchukov, Dmitri Petrov, Lovro Puzar, Yee Jiun Song, and Venkat Venkataramani. Tao: Facebook's distributed data store for the social graph. In *Proceedings of the USENIX Annual Technical Conference*, pages 49–60, June 2013.

[5] Brown CS 300 staff. Project 5b: Privacy-Compliant KVStore. Retrieved Feb 08, 2024 from https://cs.brown.edu/courses/csci0300/2023/assign/projects/project6.html, May 2023.

[6] Ann Cavoukian. Privacy by Design: The 7 Foundational Principles, 2011. Retrieved Feb 08, 2024 from https://iapp.org/media/pdf/resource_center/pbd_implement_7found_principles.pdf.

[7] California Civil Code. 1.81.5. California Consumer Privacy Act of 2018 [1798.100-1798.199.100]. Retrieved Feb 08, 2024 from https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5, 2018.

[8] Serge Egelman, Julia Bernd, Gerald Friedland, and Dan Garcia. The teaching privacy curriculum. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, page 591–596, 2016.

[9] E. R. Fyfe, N. M. McNeil, J. Y. Son, and R. L. Goldstone. Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, 26(1):9—25, 2014.

[10] Ruth Gavison. Privacy and the limits of law. *The Yale Law Journal*, 89(3):421–471, 1980.

[11] Barbara J. Grosz, David Gray Grant, Kate Vredenburgh, Jeff Behrends, Lily Hu, Alison Simmons, and Jim Waldo. Embedded EthiCS: Integrating Ethics across CS Education. *Commun. ACM*, 62(8):54–61, jul 2019.

[12] Tadayoshi Kohno, Yasemin Acar, and Wulf Loh. Ethical frameworks and computer security trolley problems: Foundations for conversations. In *Proceedings of the 32nd USENIX Security Symposium*, 2023.

[13] Priya C Kumar and Virginia L. Byrne. The 5ds of privacy literacy: a framework for privacy education. *Information and Learning Science*, 123(7/8):445–461, 2022.

[14] C. Dianne Martin, Chuck Huff, Donald Gotterbarn, and Keith Miller. A framework for implementing and teaching the social and ethical impact of computing. *Education and Information Technologies*, 1(2):101–122, June 1996.

[15] Bryce Clayton Newell, Cheryl A. Metoyer, and Adam D. Moore. Privacy in the family. In *Social Dimensions of Privacy: Interdisciplinary Perspectives*. Cambridge University Press, 2015.

[16] Rajesh Nishtala, Hans Fugal, Steven Grimm, Marc Kwiatkowski, Herman Lee, Harry C. Li, Ryan McElroy, Mike Paleczny, Daniel Peek, Paul Saab, David Stafford, Tony Tung, and Venkateshwaran Venkataramani. Scaling Memcache at Facebook. In *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (NSDI)*, pages 385–398, April 2013.

[17] European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). Retrieved Feb 8, 2024 from https://eur-lex.europa.eu/eli/reg/2016/679/oj, 2016.

[18] Richard A. Posner. The economics of privacy. *The American Economic Review*, 71(2):405–409, 1981.

[19] Beate Roessler. *The Value of Privacy*. Polity Press, 2005.

[20] Garrett Smith, Kirsten Chapman, Zainab Agha, Janet Ruppert, Spring Cullen, Sushmita Khan, Bart Knijnenburg, Jessica Vitak, Priya C. Kumar, Pamela J. Wisniewski, and Xinru Page. Privacy interventions and education (PIE): Encouraging privacy protective behavioral change online. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023.

[21] Wouter Steijn. A developmental perspective regarding the behaviour of adolescents, young adults, and adults on social network sites. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 8, 07 2014.

[22] Daniel Susser. Information privacy and social self-authorship. *Techné: Research in Philosophy and Technology*, 20(3):216–239, 2016.

[23] Mohammad Tahaei, Kami Vaniea, and Awais Rashid. Embedding privacy into design through software developers: Challenges and solutions. *IEEE Security & Privacy*, 21(1):49–57, 2023.

[24] Ying Tang, Morgan L. Brockman, and Sameer Patil. Promoting privacy considerations in real-world projects in capstone courses with ideation cards. *ACM Trans. Comput. Educ.*, 21(4), oct 2021.

[25] Judith Jarvis Thomson. The right to privacy. *Philosophy & Public Affairs*, 4(4):295–314, 1975.

[26] Samuel D. Warren and Louis D. Brandeis. The right to privacy. *Harvard Law Review*, 4(5):193–220, 1890.

[27] J. Whittle. Is your software valueless? *IEEE Software*, 36(3):112–115, 2019.

A Codebook Overview

This section gives more information on the collection of codes that we identified and used during data analysis. A spreadsheet version—complete with all of the low-level codes, descriptions, and examples of when we applied them—is on the project website.[2]

A.1 Implementation Details (of students' decisions)

Codes for student implementation captured which data were modified, the nature of the modification (or deletion), and (if applicable) how data were obfuscated. The low-level codes are constructed from the options in A.1.1 through A.1.3.

A.1.1 Tweeter Data Categories

- **Req post(s):** Post(s) made by the Tweeter user requesting GDPR deletion.
- **Other post(s):** Post(s) made by Tweeter users other than the requester.
- **Req replies:** Replies to the requester's post(s). Reply posts are included in the superset of the requester's posts and other users' posts, but these tags were only applied when the student employed a deletion strategy specifically for replies.
- **Req IDs:** The requester's Tweeter user ID/username.

A.1.2 Actions

- **Deletion:** The student fully deleted the data from the Tweeter DB.

[2] https://responsible.cs.brown.edu/research/stakeholder.html

- **Obfuscation:** The student left the username or post key-value pair but redacted some or all of the username or post content (e.g., replacing "username123" with "[Deleted User]" or "I dislike @MattBleat" with "[Deleted Post].") Modifiers from A.1.3 were also applied in the case of obfuscation.
- **Disassociation:** When the post content is fully retained but disassociated from the user/ID who posted it.
- **No deletion:** No deletion or modification to the data.
- **Other:** Something not captured by one of these, which in all cases was the student using multiple deletion strategies for one category of data. In those cases, this tag was applied in addition to the most destructive action employed (Deletion > Obfuscation > Disassociation).

A.1.3 Obfuscation Categories

- **Specific content:** Redacted/obfuscated only specific portions of the post(s), e.g., replacing "I dislike @MattBleat" with "I dislike @Deleted."
- **Whole post:** Redacted/obfuscated only specific portions of the post(s), e.g., replacing "I dislike @MattBleat" with "I dislike @MattBleat" with "[Deleted Post]."

A.2 Tensions

Pairs of values or ideas that students cited as in conflict in their justifications, either explicitly or implicitly.

- **Privacy vs. Free Speech (explicit):** "This method strikes a balance between the right to privacy and the right to freedom of expression."
- **Right to Delete vs. Research (explicit):** "However, if people have the right to delete all of their data, individuals who do so can no longer be studied by historians and researchers."
- **Right to Delete vs. Accountability (implicit):** "By deleting the original tweets, our implementation could potentially limit the public's ability to hold political figures accountable for their past actions."

A.3 Values

Values or ideas that students mentioned in justifying their implementation that were not in tension with other values or ideas. These codes were applied to *cited* values and ideas, regardless of whether or not the student actually holds them. Samples of values include:

- **Research:** "Academics like Beth should be able to conduct their research."
- **Public Interest:** "To benefit the public interest, there would need to be a 'public figure' flag for each user, which would decide between a hard deletion (for normal users) and soft deletion (for public figures)."

A.4 Implementation Changes

Specific categories of changes students said they would make to their implementation given a different stakeholder pair or more time, resources, or ability. For example:

- **Additional moderation:** "I would have thought more about checking if certain posts should be deleted for other reasons or not."
- **Add more tech:** "I would have done sentiment analysis on the posts using natural language processing."
- **DB Modifications:** "I would've added a 'public figure' tag to every user identifying if they were a public figure or not."
- **Separate DB/archive access:** "If I was considering pair 1, I would've added a special archive of deleted posts by politicians."
- **Different delete system:** "Even if the post doesn't merit GDPR-based deletion, the user should still be able to delete their posts using a different deletion tool."
  Human moderation: "Tweeter would need to employ moderators to decide which posts should be deleted in order to handle each case fairly."

A.5 Privacy

Tags applied based on students' descriptions of what privacy looks like or entails, e.g., "physical safety," "social media anonymity," or "avoiding reputational harm." Every instance of the "Privacy" value tag and tension tags involving privacy will have one or more of these tags applied.

Of particular note for tagging concerns are the "previous reputational harm" and "future ability to recover" reputation tags: the former is applied in cases where the student was attempting to mitigate previous reputational harm, while the latter was applied when student was concerned with the subject's future ability to recover their reputation.

Additionally, the undeserved disadvantages tag is applied when students argue that not supporting privacy would result in undue disadvantages to other stakeholders in the situation, e.g. the requester's family or sports team. Examples include:

- **Anonymity:** "This protects the user's privacy by allowing them to 'drop off the map."'
- **Social media anonymity:** "To further protect her privacy, we uncoupled the link between her ID and username so people couldn't track down her content on social media."
- **Future ability to recover reputation:** "If private info about Matt leaked from the investigation, that would make it hard for him to make friends or find a job in the future."
- **Previous reputation harm:** "Deleting his post content will help protect his privacy and undo some of the damage done to his reputation."

- **Undeserved disadvantages:** "Not deleting the Blimps' posts would cause undue harm to their family and kids and deny them privacy in a complicated chapter of their lives."

A.6 Privacy Spectrum

Tags applied based on whether the student's viewed privacy as contextual or as absolute, both in their implementation decisions and in any stated beliefs. For the implementation spectrum, we looked at statements made about privacy related to the student's specific implementation; belief spectrum tags were based on statements made anywhere in the document. In many cases, it was unclear what the spectrum was for either or both of these components, as not all students made privacy-specific justifications of their implementations. Specifically:

- **impl contextual, belief contextual:** Applied when the student made claims about privacy being contextual, both in their actual implementation and in their overall beliefs about how to approach any scenario.
- **impl absolute, belief contextual:** Applied when the student made absolute claims about privacy in their implementation such as "this deletion tool allows anyone to delete their data to protect their privacy," but then in later sections walks back their absolute belief by saying the right to privacy might be contextual in some cases, like "for public figures, I would consider implementing an archive DB to allow researchers access," indicating that the right to privacy isn't universal.

A.7 Limitations

Tags applied based on a subset of the implementation limitations identified by students. Not all identified limitations were tagged. We separated these from the values category specifically because they were identified as limitations of the students' implementations and carry that negative connotation. For example:

- **Deletion process is overly-broad:** "I would have thought more about checking if certain posts should be deleted for other reasons or not."
- **Implementation doesn't do all of requester's wish/desire:** "One shortcoming of my implementation is that it doesn't."
- **No true deletion:** "I would've added a 'public figure' tag to every user identifying if they were a public figure or not."

A.8 Stakeholders

Tags applied when students discussed non-human third-party stakeholders or facets of Tweeter as a stakeholder, respectively, indicating considerable depth of consideration in the stakeholder analysis section.

- **Tweeter or facets of Tweeter:** "Doing this would lead to a conflict between the staff of Tweeter and its users."
- **Non-human third-party stakeholders:** "Deleting these posts could jeopardize academic understanding of the Breisand effect and research as a whole."

A.9 Public Figures

Tags applied when students argued that public figures (either celebrities, public servants/politicians, or public figures in general) should be handled differently when deleting Tweeter data. These tags were applied whether the student actually implemented different deletion strategies for public figures or just said they would when considering alternate implementations.

- **Public figures treated differently:** "If Matt becomes a public figure, it might be in the public's interest to retain these posts rather than deleting them."
- **Celebrities treated differently:** "Being a celebrity inherently means they have a lower expectation of privacy."
- **Politicians treated differently:** "I would have thought more about checking if certain posts should be deleted for other reasons or not."

A.10 Second-Order Effects

A general tag applied when students identified a second-order effect of their implementation. An example quote is "Not deleting her address from posts could put increased pressure on local services due to an increased tourist presence, disrupting the peace for her neighbors and community."