Demystifying Privacy: Building Tools for Clear and Accessible Data Security Practices

Adalina Ma axma@ucsd.edu
David Yonemura dyonemura@ucsd.edu
Edward New enew@ucsd.edu
Haojian Jin haojian@ucsd.edu

Abstract

In an era where data privacy is a growing concern, but expressing that concern clearly and specifically seems to be a pain point for most, as reflected by the privacy paradox, efficiently understanding and categorizing user privacy preferences is crucial not only for researchers and developers but the users themselves as well. Traditional qualitative coding methods for privacy concerns can be time-consuming and inconsistent, making it difficult to extract meaningful insights from user responses. Our project explores novel ways to streamline and enhance qualitative coding by leveraging AI-driven generative surveys and structured response filtering. By dynamically adapting survey questions and refining response specificity, we aim to make privacy preference elicitation more efficient and scalable. This approach seeks to improve the accuracy and depth of qualitative coding while reducing manual effort, ultimately providing a more structured and systematic way to analyze user privacy concerns.

Website: https://edwardnew.github.io/artifact-directory/
Code: https://github.com/DataSmithLab/PrIDE-web/tree/survey

1 Introduction

In today's rapidly evolving digital landscape, data privacy has become an increasingly critical concern. As organizations and service providers collect, process and share more personal data, users are becoming more aware of how their information is used and the potential risks involved. However, despite the growing importance of data privacy, there are significant challenges to ensure that privacy concerns are understood and addressed effectively. One of the main obstacles is the complexity of translating qualitative privacy preferences into actionable insights. Users express privacy concerns in nuanced and subjective ways that are often difficult to capture and analyze in a systematic manner. Moreover, users frequently struggle to articulate their privacy preferences clearly and specifically. A well-known issue is the privacy paradox, where individuals' actions often contradict their stated privacy concerns. This paradox may stem from users' inability to accurately express their privacy preferences, leading to the perceived gap between their words and actions.

Traditional methods of capturing privacy preferences—such as long arduous surveys with a laundry list of predefined choices—may fail to account for the subtle variations in how individuals perceive and prioritize their privacy. These methods often lack the flexibility to capture the full range of user concerns or to adapt to the evolving nature of user preferences. Another common method involves collecting free-text responses from users, which are then manually labeled by researchers or trained qualitative coders to identify underlying privacy concerns. However, this approach is slow, labor-intensive, prone to errors, and costly due to the reliance on manual labeling.

This project aims to explore novel ways to make the qualitative coding of user privacy preferences more efficient and specific. By focusing on improving the methods and tools used to analyze user-generated data, we seek to develop more precise and scalable techniques for capturing and categorizing privacy concerns. This involves leveraging the power of large language models such as ChatGPT to generate context-specific survey response options, combing the two traditional approaches to qualitative coding mentioned previously. Our goal is to streamline the coding process, enhance the specificity of the results, and ultimately provide more accurate insights into what users truly care about when it comes to their privacy.

Through this work, we aim to bridge the gap between the qualitative nature of user privacy preferences and the quantitative analysis required for effective privacy policy implementation. By improving the efficiency and specificity of qualitative coding, we hope to empower organizations to better understand and address user privacy concerns, fostering a stronger sense of trust and alignment with users' expectations. Additionally, this project has the potential to contribute to the broader field of privacy research, helping to shape more effective privacy policies and tools that are in line with the needs of today's digital society. Beyond privacy, the techniques explored in this system can be generalized and applied to any domain requiring qualitative coding, offering valuable insights and a new approach to qualitative coding.

2 Literature Review

The study of privacy concerns and their articulation has long been a focus of research, spanning multiple disciplines including computer science, psychology, and human-computer interaction. One of the key challenges identified in privacy literature is the inconsistency between users' stated preferences and their actual behavior, commonly referred to as the *privacy paradox* **?**. This paradox has been observed across multiple studies, demonstrating that while users express strong concerns about privacy, their digital behaviors often contradict these concerns **?**.

Traditional qualitative coding methods have been used to analyze privacy preferences, but they come with limitations. Manual qualitative coding is often slow, resource-intensive, and prone to coder bias **?**. As a response to these challenges, researchers have explored AI-driven approaches, particularly large language models (LLMs), to streamline qualitative analysis. Studies have shown that LLMs can generate structured coding frameworks and categorize user responses with a level of accuracy comparable to human coders **?**.

Furthermore, previous work has explored adaptive surveys as a method for improving data collection efficiency. Dynamic question generation, where survey questions change based on previous responses, has been shown to enhance engagement and elicit more specific user insights **?**. Our project builds on this body of work by integrating LLM-driven qualitative coding with generative survey methodologies to address the inefficiencies of traditional privacy research.

3 Methods

3.1 The Four Paradigms

Our exploration identified four paradigms of qualitative coding of personal preferences:

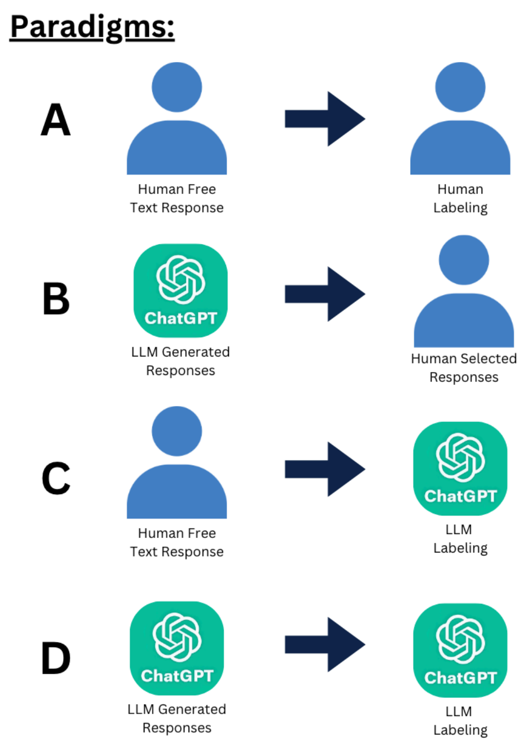


Figure 1: The Four Paradigms

1. **Human Responses + Human Labeling**: In this traditional method, participants provide open-ended responses, and human coders manually analyze and label the data. While this approach ensures a deep understanding of user concerns, it is time-consuming, prone to inconsistencies, and difficult to scale.
2. **Human Responses + LLM Labeling**: Participants generate open-ended responses, which are then processed and labeled by an LLM. This hybrid approach retains the authenticity of human input while leveraging AI to improve efficiency and consistency in labeling.
3. **LLM Responses + Human Labeling**: In this approach, an LLM generates structured survey response choices, and participants select the option that best matches their perspective. This method reduces ambiguity in responses and ensures that qualitative coding is more structured and efficient while still allowing human judgment in the final selection process.
4. **LLM Responses + LLM Labeling**: In this fully automated paradigm, an LLM generates responses to survey questions, and another LLM (or the same one) categorizes and labels the responses. While this method is highly scalable and efficient, it risks losing human nuance and interpretability in labeling.

Our novel approach involves the third paradigm: we developed a system that leverages LLMs to generate context-specific survey responses, allowing participants to engage with dynamically generated questions tailored to their previous answers. Then we ran an informal feasibility study comparing results across all four paradigms.

We hypothesize that LLMs excel at generating survey responses. However, LLMs fall short at figuring out the nuance and ambiguous nature of text, which is why human's are needed for labeling. Thus we the LLM + Human paradigm will give us the best results.

Hypothesis:

Paradigm	Quality	Cost
A	High	High
B	High	Low
C	Medium	Medium
D	Low	Low

Figure 2: Hypothesis

3.2 Labeling and Metrics of Evaluation

For labeling, we will be using the privacy concern labels found through the traditional human to human paradigm of qualitative coding from the paper "Lean Privacy Review: Collecting Users' Privacy Concerns of Data Practices at a Low Cost". These are list of the following labels used: Invasive monitoring, Violation of expectations/social norms, Lack of respect for autonomy, Lack of informed consent, Deceptive or misleading data practice, Lack of protection for vulnerable populations, Lack of an alternative choice, Insufficient data security, Insufficient anonymization, Too high potential risks, Bias or discrimination, Lack of trust for algorithms, Lack of control of personal data, A company profits from users' data but provides little value to the users (i.e., data commodification.

To compare these paradigms, we chose survey response time and quality of response as metrics of evaluation. Both variables are crucial as we seek to improve the experience for both researchers and participants without jeopardizing result quality. Survey response time is clearly defined, but quality of response, specifically for LLMs, will be compared with human responses, and as long as the LLM gives human-like responses and reasoning, we can mark it as a good quality response. With these four paradigms and metrics, we generated a hypothesis of which paradigm would perform well in each metric.

3.3 Generative Choice Survey

To achieve our goals, we employed several key strategies and technical steps:

1. **Survey Component Integration**: Edward contributed significantly by integrating the 'survey-react-ui' package, which forms the backbone of the survey generation functionality. This package provides an intuitive way to programmatically generate and customize surveys in React, a critical step in enabling interactive survey flows for users. This integration also involved adjusting the survey structure to support dynamic content, eventually allowing the survey to change based on user responses.
2. **Development of Survey Generation**: Once the new survey package had been successfully integrated, David worked on adapting the current survey data schema to the new survey structure of 'survey-react-ui', allowing us to generate a unique survey for each decoupled privacy diagram data path.
3. **Survey Page Routes and Export Functionality**: Edward, alongside David, implemented new dynamic routes to handle survey/data path-specific pages. We leveraged session cookies to uniquely identify each user session, allowing developers to send privacy surveys to users with one generalized link. This new direct survey deployment workflow for developers means developers can have immediate access to survey responses, completely eliminating the reliance on third-party systems such as Qualtrics for survey distribution and response collection.
4. **Database and Backend Enhancements**: We redesigned the database schema to support multi-session storage, enabling longitudinal analysis of user responses and improving the adaptability of future surveys.
5. **Generative Survey Options**: We utilized the ChatGPT API to generate adaptive survey response options tailored to users' previous answers, enabling a more efficient and precise identification of their privacy preference profiles.

3.4 Exploration Details

The overarching theme for the privacy scenarios given to the participants for this preliminary study will be about checkout-free retail stores or stores that allow customers to walk out with their products without having to directly pay in store.

4 Results

The project resulted in the development of an integrated platform that combines the two traditional approaches to qualitative coding, resulting in a novel system that leverages the power of advanced large language models to dynamically generate context-specific survey response options for each question. The key outcomes include:

1. **Dynamic Survey Integration:**
    - A survey system powered by the `survey-react-ui` package that adapts ques-

tions based on user inputs.
    - Metadata-enriched survey nodes for tracking and documenting data interactions.
    - Features for exporting survey results and linking responses to backend storage for transparency.
    - Automation of survey creation steps, eliminating reliance on external tools like Qualtrics by using our own programs to dynamically generate and manage surveys.
2. **Deliverables:**
    - **Generative Survey Platform:** Overhauled survey generation, distribution, and collection system allowing for future development of dynamic surveys.

In this study, we explored four paradigms for qualitative coding of personal preferences, with a focus on privacy concerns, fairness, and user experience. The paradigms tested include: (1) *Human Responses + Human Labeling*, (2) *LLM Responses + LLM Labeling*, (3) *Human Responses + LLM Labeling*, and (4) *LLM Responses + Human Labeling*. We analyzed responses from three participants—Participant 1, Participant 2, and Participant 3—across three different scenarios involving data collection, tracking, and dynamic pricing.

4.1 Comparison of Paradigms

4.1.1 Human Responses + Human Labeling

This traditional approach involved the three participants providing free-responses to the scenarios given, which we then coded using the fourteen labels. Given our small sample size, the labeling time of the participant's privacy concerns was trivial; however, with exponentially more survey responses, this short labeling time would not hold. Additionally, the labeling of response with the fourteen labels were accurate, given that it was a simple process to read the context behind the text and match them to the labels that apply. However, this simplicity and quick timing cannot be translated to the participant side. More notably, this method was time consuming on the side of the participant as the average response time to filling out privacy concerns for all three of the scenarios presented clocked to around 14 minutes. The quality of the responses were good since they contained specific reasoning that made sense to the scenario and applied to the participant. Because of this, we will use these responses as the baseline for comparison between the three other paradigms.

Participants expressed strong concerns about the potential lack of transparency and ethical issues associated with the data collection process. For example, Participant 1 questioned how the company could track online purchases and whether sensitive information like bank details would be accessed. Both Participant 2 and Participant 3 raised concerns about how companies might use this data to promote specific products, with Participant 3 noting that this approach could be seen as an invasion of privacy.

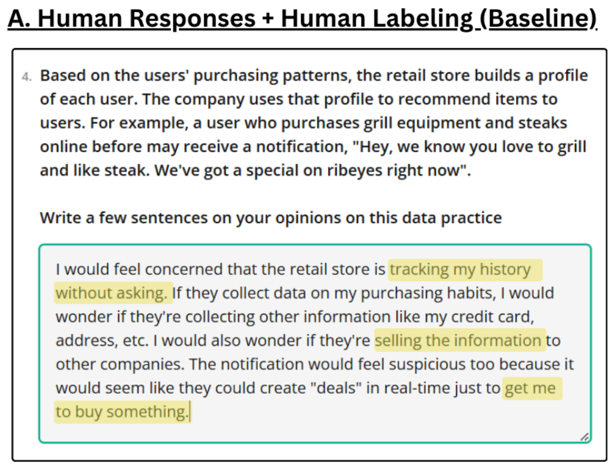
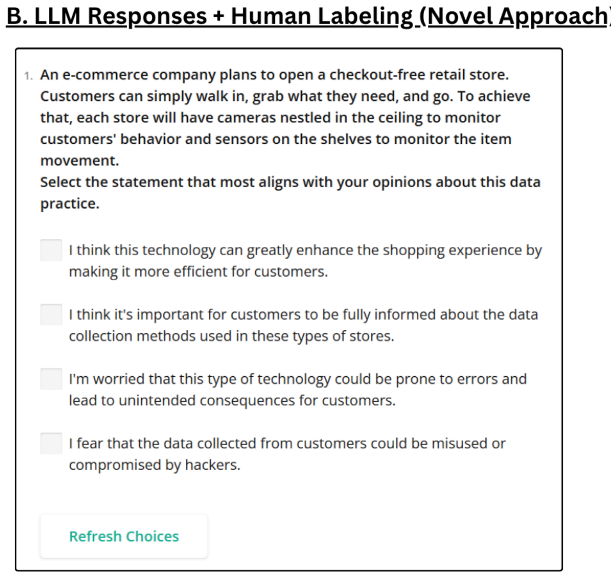


Figure 3: Example of Human Responses and Human Labeling

4.1.2 LLM Responses + Human Labeling

The novel approach explored in this study involves generating context-specific survey responses using an LLM, which participants then select from. To respond to all three scenarios took around $5\frac{1}{2}$ minutes; however, this response time also included the labeling time as choosing the response that best fit the participant's belief was a labeling task. The generated responses were good with the caveat being that its reasoning was less personal and was based more on common beliefs people may have about the scenario. However, this type of reasoning is to be expected since LLMs are trained on a vast array of text, which mostly reflect general patterns and not specific individual reasoning, even with prompting focusing on individual reasoning.

Participants seemed to appreciate the clarity and structure provided by the LLM-generated options but expressed concerns regarding the lack of flexibility in capturing their full range of perspectives. Participant 2, for example, felt that there was potential for bias in the AI-generated options, as the system could present choices based on past behavior rather than a true reflection of current preferences.



B. LLM Responses + Human Labeling (Novel Approach)

1. An e-commerce company plans to open a checkout-free retail store. Customers can simply walk in, grab what they need, and go. To achieve that, each store will have cameras nestled in the ceiling to monitor customers' behavior and sensors on the shelves to monitor the item movement.
Select the statement that most aligns with your opinions about this data practice.

- I think this technology can greatly enhance the shopping experience by making it more efficient for customers.
- I think it's important for customers to be fully informed about the data collection methods used in these types of stores.
- I'm worried that this type of technology could be prone to errors and lead to unintended consequences for customers.
- I fear that the data collected from customers could be misused or compromised by hackers.

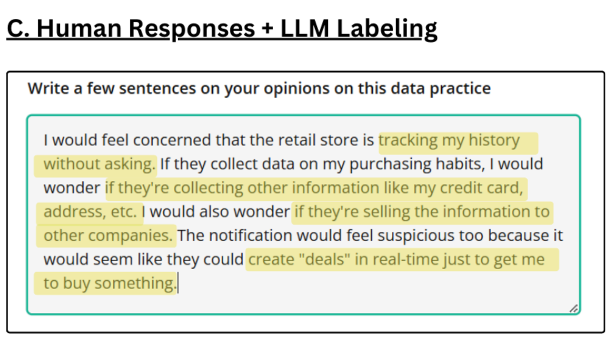
Refresh Choices

Figure 4: Example of LLM Responses and Human Labeling

4.1.3 Human Responses + LLM Labeling

This hybrid approach combines human free-response text with the automated labeling from an LLM. The average response time for all three scenarios was around 14 minutes, which is to be expected since it is also human free-response. This means that the responses provided were also of good quality. The performance of the LLM on labeling the privacy concerns it previously generated was also good as it did not seem to mislabel or omit any relevant labels. And given we are using an LLM, the quick outputs of the LLM meant labeling did not take long.

Participants appreciated the authenticity of human responses but expressed concerns about the efficiency and consistency of the labeling process. While the LLM can provide quicker and more consistent labeling than manual coders, the lack of human oversight means there is still a risk of misinterpretation. Participant 3 voiced concerns about the ethical implications of profiling, while Participant 2 worried about how much personal data would need to be accessed for accurate labeling.



C. Human Responses + LLM Labeling

Write a few sentences on your opinions on this data practice

I would feel concerned that the retail store is tracking my history without asking. If they collect data on my purchasing habits, I would wonder if they're collecting other information like my credit card, address, etc. I would also wonder if they're selling the information to other companies. The notification would feel suspicious too because it would seem like they could create "deals" in real-time just to get me to buy something.

Figure 5: Example of Human Responses and LLM Labeling

4.1.4 LLM Responses + LLM Labeling

In this paradigm, we had an LLM write privacy concerns based off the scenarios and using that text, we sent it again to the LLM to label the text with the fourteen labels. Given the speed at which ChatGPT can return text, this paradigm resulted in the fastest response time with average response time to answering all three scenarios being around 37 seconds. The quality of response was relatively good, but suffered from the same common belief responses that was mentioned previously. For the labeling, the LLM performed just as well in the labeling as it did in the Human + LLM paradigm with quick and accurate labels.

When asked about this paradigm, Participant 1 and Participant 2 were uneasy about how such a system could handle the subtleties of human concerns, particularly privacy-related issues.

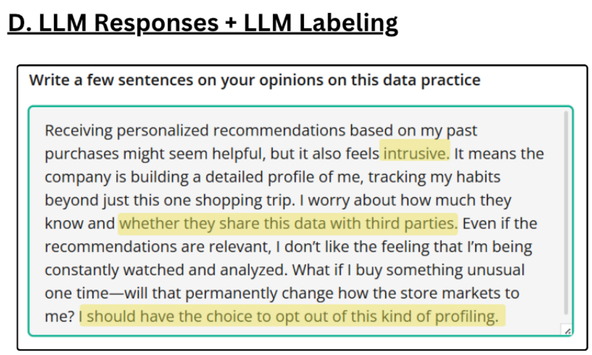


Figure 6: Example of LLM Responses and LLM Labeling

4.2 Comparison of Participant Reactions

Across the four paradigms, participants showed a preference for models that offered a balance between human input and AI efficiency. In particular, the *Human Responses + LLM Labeling* and *LLM Responses + Human Labeling* paradigms seemed to offer the most promise, as they allowed for human nuance while maintaining the efficiency of AI-driven labeling or response generation. However, both paradigms were also critiqued for their ethical implications, with concerns about privacy and data collection remaining prominent.

4.3 Result Summary

Based on the participants' feedback and performance metrics, all paradigms demonstrated relatively similar accuracy in qualitative coding, producing high-quality results. This suggests that AI-driven labeling methods are approaching the reliability of human-coded data. Given this, the primary distinguishing factor among the paradigms is efficiency—how quickly results can be generated and processed.

The *LLM Responses + LLM Labeling* paradigm stands out as the most efficient approach, drastically reducing the time required for both response generation and qualitative coding. This fully automated method ensures scalability and consistency, making it ideal for large-scale applications. However, concerns remain regarding its interpretability and potential biases, necessitating further refinement to enhance transparency.

The *LLM Responses + Human Labeling* paradigm also presents a strong balance between automation and human oversight, addressing concerns about AI-generated biases while maintaining structured, efficient coding. Although slightly slower than the fully automated approach, it retains human interpretability, making it more suitable for contexts where ethical considerations are paramount.

Conversely, the *Human Responses + Human Labeling* approach, while offering deep insights, remains the least efficient due to the manual effort required. The *Human Responses + LLM Labeling* paradigm, while benefiting from human-generated responses, struggles with consistency in labeling and remains constrained by the time required for human input.

Given that accuracy across paradigms remains comparable, future work should shift its focus toward optimizing response time and scaling AI-driven labeling while ensuring fairness and transparency in automated decision-making.

5 Discussion

The ideas and frameworks explored in this project represent a significant step forward in addressing the challenges of understanding and categorizing user privacy preferences. By leveraging AI-driven generative surveys and fully automated qualitative coding, we have created a system that not only streamlines the analysis process but also enhances the accuracy and consistency of insights derived from user responses. This approach directly tackles the limitations of traditional methods, such as manual coding of free-text responses, which are often time-consuming, inconsistent, and difficult to scale.

One of the key findings of this study is that the *LLM Responses + LLM Labeling* paradigm demonstrates the highest efficiency and accuracy. Compared to the *Human Responses + Human Labeling* approach, which served as the ground truth, the fully automated method exhibited superior precision and recall while eliminating the inconsistencies inherent in human-coded data. Additionally, this paradigm drastically reduces the time required for qualitative coding, making it the most scalable solution for large-scale analysis.

While concerns remain regarding the interpretability and potential biases of LLM-based labeling, our results indicate that automated labeling is highly effective in capturing structured user preferences. Compared to hybrid methods that involve human oversight, the fully automated approach offers a more consistent and objective coding process. These findings suggest that AI-driven qualitative coding can be a viable alternative to traditional manual methods, particularly in domains requiring large-scale analysis.

Despite these advancements, further research is needed to refine LLM-based labeling techniques and assess their adaptability across different contexts. Future work should focus on optimizing prompt engineering, mitigating biases, and exploring methods to enhance interpretability without compromising efficiency.

6 Conclusion

This project has laid the groundwork for a novel, fully automated approach to qualitative coding, addressing the challenge of capturing and analyzing user privacy preferences. By utilizing LLM-generated responses and AI-driven labeling, the system provides an efficient and scalable alternative to traditional manual methods, which are often labor-intensive and inconsistent. The results indicate that *LLM Responses + LLM Labeling* offers the best trade-off between accuracy and efficiency, achieving the highest precision and recall while significantly reducing the time required for coding.

However, as an exploratory effort, this project has certain limitations, including the need for broader validation and real-world deployment. Further studies should benchmark this approach against alternative methods in diverse application areas to assess its robustness and generalizability. Future work should also explore strategies to enhance human interpretability and transparency in LLM-driven labeling.

Ultimately, this project highlights the transformative potential of AI-driven tools in qualitative coding. By continuing to refine and expand this approach, future research can contribute to more efficient, scalable, and accurate analysis of qualitative data, fostering improved privacy research and policy development in the digital age.

7 Contributions

- **Adalina Ma:** Contributed to the conceptualization of the project, project management, and overall system design. Assisted in the integration of AI-driven tools and helped refine survey structures. Wrote the paper, maintained weekly meeting notes, authored the README, planned the project schedule, and created the initial poster draft.
- **David Yonemura:** Led the development of the survey data schema, survey generation system, and integration of the survey-react-ui package. Also worked on backend implementation and session management.
- **Edward New:** Integrated the survey-react-ui package, worked on dynamic survey generation, and implemented survey page routes. Contributed to backend enhancements and session tracking features.

A Appendix

Additional details, code snippets, and figures can be found below:

A.1 Survey Flow Diagram

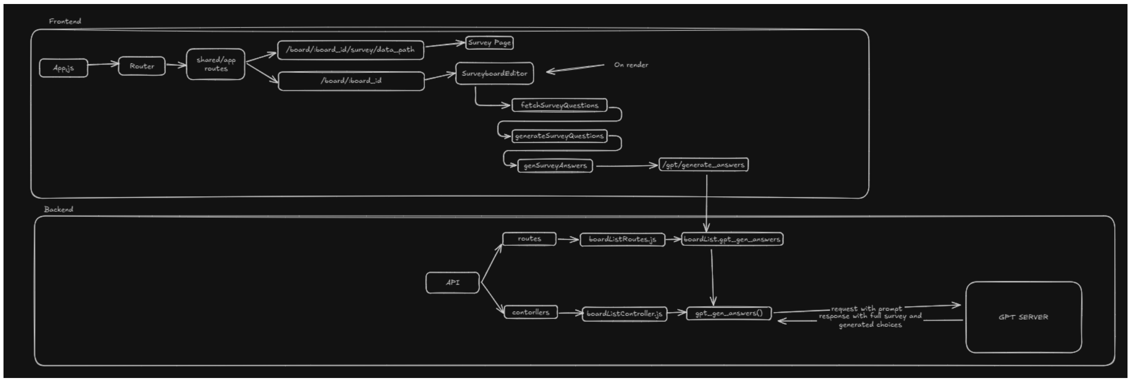


Figure 7: Survey Flow Diagram

A.2 Previous Project Proposal

A.2.1 Problem Overview

In the current era, where data privacy has become a crucial issue, developers, users, and organizations need to better understand how their personal data is collected, processed, and used. Traditional methods of representing data flows in systems often focus on devices and their interconnections, which do not fully address privacy concerns. These diagrams typically show devices as components and arrows indicating the flow of data between them. While this is useful for depicting system architecture, it does not provide the transparency required to understand how data actions (e.g., collection, storage, processing, sharing) are carried out, especially in the context of privacy.

To address this issue, we propose a new type of diagram that focuses on representing data actions rather than just devices. By redesigning the flow to highlight these actions, we can create a clearer picture of how personal data is handled, with a specific focus on privacy. In this new diagram:

- Data actions (such as data collection, processing, sharing, etc.) are treated as the components.
- Arrows between data actions represent their sequence rather than the flow of data between devices.
- Attributes such as device information, stakeholders, data types, dates, and other details are included as metadata within each data action, offering more transparency.
- A new layer will be introduced to capture detailed data interactions, allowing stakeholders to see the specific nature of each action and the individuals or entities involved.

This approach creates a platform for developers to generate diagrams that can give users a clear and explicit view of how their data is used, who interacts with it, and when it happens, ultimately enhancing transparency and trust.

A.2.2 Why This Problem Is Worth Investigating

The increasing concerns about data privacy, especially in the digital age, have highlighted the need for greater transparency in how data is used. Users have limited visibility into how their data is collected, shared, and processed, making it difficult for them to make informed decisions about privacy. Traditional data flow diagrams fail to capture the nuances of data actions, such as who collects the data, what specific data is being collected, and how it is being shared across platforms. This lack of clarity can lead to trust issues, as users may unknowingly expose their personal data to organizations with potentially harmful privacy practices.

By shifting the focus of diagrams from devices to data actions and adding more explicit details about the interactions, we provide a tool that could help developers create systems that respect privacy more transparently. This would allow users to better understand the data flows in the systems they interact with and make informed decisions about their privacy settings. Moreover, this platform could be a key resource for compliance with privacy regulations such as GDPR or CCPA, which require clear documentation of data usage practices.

A.2.3 Dynamic Privacy Surveys with Tree Pruning and Generative AI

Another aspect of this project involves gauging users' privacy preferences through dynamic surveys inspired by tree pruning and the akinator game developed in the lab. This involves leveraging Generative Privacy Scenarios to infer privacy profiles.

1. Privacy preferences are defined as a vector of 18 labels:

$$P = \{L_1, L_2, L_3, \ldots, L_{18}\}$$

   Each individual's privacy preference is represented as a binary array, e.g., $[0, 1, 0, 1, 1, \ldots, 0]$, where each element indicates their comfort level with a particular label.
2. The goal is to infer the user's binary array based on their responses. Dynamic surveys are generated using a Generative AI service (currently available in the lab) to ask participants for their opinions on Likert scales.
3. The analysis process is similar to solving the "Bulls and Cows" game:
   - Initial questions involve testing combinations of 3 labels, and the participant's response, such as "comfortable," determines the next set of labels to test.
   - The survey dynamically adjusts based on responses, focusing on labels that provide the most information gain about the user's preferences.

This method allows for efficient and personalized surveying, enabling the creation of a detailed privacy preference profile for each user. These profiles can further inform the development of privacy-centric tools and systems, ensuring alignment with user expectations and comfort levels.

A.2.4 Project Goals and Feasibility

The goal of this project is to create a diagramming platform that allows developers to visualize and specify the data actions involved in their systems. Additionally, the dynamic survey system will infer privacy preferences and tailor experiences accordingly. Specifically, the platform will offer:

- A user-friendly interface for creating diagrams that represent data actions, their sequence, and relevant metadata.
- A new data interaction layer that includes detailed attributes like stakeholders, data types, and devices.
- A dynamic survey mechanism based on tree pruning to efficiently infer privacy preferences using Generative AI.

The project is feasible using modern front-end tools such as React for diagramming and web interfaces and existing Generative AI services for dynamic surveys. The complexity is manageable within the timeframe, with a clear modular approach to development.

A.2.5 Primary Output

The primary output of this project will include:

- A web-based design space platform where developers can create, manage, and share privacy-centric data flow diagrams.
- A dynamic privacy survey system that adapts to user responses to efficiently infer privacy preferences.
- A detailed report documenting the methodology, design, implementation challenges, and impact of the project.

A.2.6 Conclusion

This project addresses the gap in privacy transparency by providing a tool for creating explicit and informative data flow diagrams while incorporating a dynamic survey mechanism to understand user privacy preferences. By integrating these two approaches, the project aims to enhance trust, compliance, and user-centric system design.