Papaya Privacy: Empowering Web Users through AI-enabled Cookie Consent and Data Monetization Tools

Thesis submitted
In Fulfillment of the Requirements
for the Degree of Master of Arts in Computational Social Science

Ram M. Kripa
With Dr. Raul Castro Fernandez (Advisor)
Dr. Zhao Wang (Preceptor)

May 2025

Contents

Papaya Privacy: Empowering Web Users through AI-enabled Cookie Consent and Data Monetization Tools

Ram M. Kripa

May 1, 2025

Abstract

The current online advertising data ecosystem is broken. Web-users are repeatedly deceived by dark design patterns into clicking "Accept All" on cookie consent banners, unwittingly opting in to a myriad of unintended data flows. These data flows can have drastic consequences, ranging from targeted advertisement all the way to social oppression. To combat these problems, we present **Papaya Privacy**, a new way to empower web-users using AI-powered cookie consent and data monetization mechanisms. We designed and implemented two interventions, Agent Papaya and Papaya Payback. With **Agent Papaya**, our Agentic AI-powered chrome extension, we give users the ability to set their cookie preferences once and execute this seamlessly across websites, mitigating the effects of dark design patterns. With users now able to effectively opt-out of data collection, we conceptualize an alternative data market, **Papaya Payback**, based on the theories of Data Unions and Contextual Integrity. Through simulation, we show that by allowing users fine-grained control of who gets to access what data of theirs and for what purposes, the Papaya Payback market creates greater social welfare than the current online advertising data market. Together, Papaya Privacy's two interventions demonstrate a practical path toward a consent-enforcing, user-centered, and economically viable data economy for online advertising.

1 Introduction

"Sunlight is said to be the best disinfectant."
This quote by Justice Louis Brandeis captures the spirit of this thesis: bringing transparency and agency to a space long dominated by dark patterns and opaque data flows. In recent years, data privacy has become a growing concern for consumers, legislators, and technologists alike. One of the most visible manifestations of this concern is the cookie consent banner, such as the one shown below (Figure 1). These interfaces ask users to make privacy choices—typically between "Accept All" and some obscure path to opt out.

These banners, though ostensibly meant to facilitate user consent, routinely manipulate and mislead users. The stakes of this deception are high. Beyond annoyance or unwanted marketing emails, real harms have been documented:

In 2022, following the U.S. Supreme Court's Dobbs ruling, a data broker was caught selling the location data of individuals visiting abortion clinics and Planned Parenthood facilities "Data Broker Is Selling Location Data of People Who Visit Abortion Clinics", n.d. This data, harvested by so-called "abortion vigilantes," has been used to prevent people from crossing state lines for reproductive healthcare.

Figure 1: Example of a Cookie Banner Displayed on a University Website, demonstrating prevalent Privacy Choices Interfaces

Figure 2: Federal Trade Commission Report on Surveillance Pricing, illustrating Real-World Harms from Data Misuse

In another instance, a dual citizen of the U.S. and Russia was sentenced to 12 years in a Russian prison for sending a $52 online donation to a humanitarian organization supporting Ukraine "US-Russian citizen sentenced to 12 years for $52 donation to Ukrainian charity", 2024. While the exact mechanism of data acquisition remains unknown, it is likely the result of third-party data sharing or a privacy breach.

The U.S. Federal Trade Commission recently documented "Surveillance Pricing" "FTC Surveillance Pricing Study Indicates Wide Range of Personal Data Used to Set Individualized Consumer Prices", 2025, wherein retailers use cross-site tracking to dynamically price-discriminate against loyal customers (Figure 2). A user who frequently searches for a product may be charged more than a casual shopper, purely based on inferred behavior.

These cases illustrate a central challenge of the modern internet: while laws such as the General Data Protection Regulation (GDPR) in the EU and the California Consumer Privacy Act (CCPA) in the U.S. have established legal rights to data privacy, the mechanisms for exercising those rights—especially cookie consent banners—often fail in practice.

The banners themselves are plagued with *dark patterns*—user interface tricks that nudge users into surrendering control of their data. Despite formal opt-out rights, most users are effectively funneled toward "Accept All," undermining the spirit of the law and perpetuating the surveillance economy. This is often termed as the 'Privacy Paradox', the phenomenon where user concern about privacy is high according to surveys, yet users repeatedly choose less privacy-preserving options like 'Accept All'.

The core problem this thesis addresses is that users are presented with the illusion of choice while navigating cookie consent banners, but are structurally coerced into surrendering control over their data through deceptive design and legal loopholes. Although data privacy regulations like GDPR and CCPA provide users

with formal rights, the interface through which these rights are exercised—namely, the cookie banner—is riddled with dark patterns that exploit user fatigue, attentional blind spots, and behavioral defaults. This coercion results in participation in an inherently unfair and inefficient online data market.

Compounding this problem is the lack of viable alternatives. Rule-based tools like Privacy Badger and Consent-O-Matic fall short due to their limited scope or hardcoded logic, and more radical solutions like Solid or blockchain-based data platforms remain inaccessible to average users and incompatible with the existing data economy. Meanwhile, ethical frameworks such as Contextual Integrity and proposals like Data Unions have yet to be implemented in a way that meaningfully bridges user intention with usable infrastructure.

This thesis contends that the point of friction, the user's browser, is the most effective intervention site to address this gap. What is required is a mechanism that not only enforces consent reliably at scale but also empowers users to articulate and automate their data preferences via simple user-friendly interfaces, laying the foundation for a fairer and more efficient user-centered data market. Without such a system, privacy remains a right in theory but a fiction in practice.

More structural changes have been proposed, like **Solid** (Socially Linked Data), a completely alternative approach to user-sovereign identity, data collection, and storage. Other Decentralized Data Platforms have also been proposed, especially in the field of web3 and blockchain technology. These mechanisms have failed to gain traction due to their incompatibility with the existing online data ecosystem. More theoretical approaches like Data Unions and Contextual Integrity have been proposed, but lack appropriate implementation. The challenge to implementing Data Unions has been an inability to collect sufficient user data to meaningfully broker with platforms, while the challenge with Contextual integrity has been an inability to adequately elicit user preferences regarding the sharing of their data.

Recognizing all of this prior work, we propose a lightweight, user-friendly browser extension called **Agent Papaya**. This AI-powered tool automates cookie banner interactions to reflect user preferences—clicking "Reject All" when available, or navigating sub-menus to apply the most privacy-preserving settings. Our hypothesis is simple: if we can make privacy as easy as set-it-and-forget-it, users will choose it. They will not just choose it, they will express their true preferences, rather than the uninformed consent they're coerced into giving.

Beyond mitigating banner fatigue, Agent Papaya lays the groundwork for something deeper: a consent infrastructure that empowers users not only to say "no" to surveillance, but also to say "yes" their terms. To that end, this thesis introduces **Papaya Payback**, a model for ethical data monetization built on the theories of Data Unions and Contextual Integrity. By linking these two products, we solve the collection problem of Data Unions, as the extension lives in the browser. Payback also solves the preference elicitation problem of Contextual Integrity with out novel Data Preferences Statement (DPS) interface.

Together, Agent Papaya and Papaya Payback constitute a new socio-technical approach to online privacy. In the chapters that follow, we explore the design, implementation, and evaluation of these tools—and argue for a future in which web users can reclaim control over their data, not by opting out of the internet, but by participating in it more intentionally.

This thesis makes three core claims. **First**, that Large Language Models (LLMs), when embedded in a browser extension, can reliably detect and respond to cookie banners in a way that reflects user preferences and mitigates the effects of dark patterns. **Second**, that this ability of users to reliably express their preferences across websites will solve the privacy paradox, as more users will select their true preference to Opt-out of cookie based tracking. **Third**, that a consent-driven data sharing market, when mediated by an agent capable of executing users' data sharing preferences, can outperform traditional data markets in both platform utility and overall social welfare. We evaluate these claims by implementing two systems: **Agent Papaya**, a browser extension using LLMs to automate cookie consent across websites, tested on real browsing data from 229 users; and **Papaya Payback**, a simulated ethical data market where user preferences are operationalized via a Data Preferences Statement interface and evaluated through agent-based modeling.

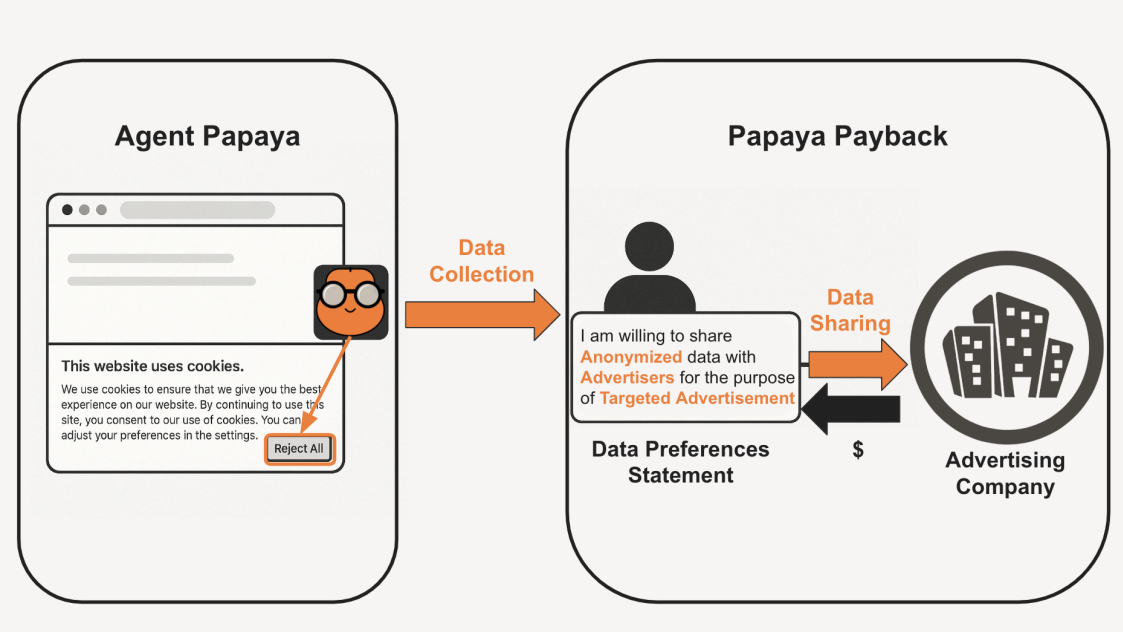


Figure 3: System Architecture of Papaya Privacy, showing both Agent Papaya and Papaya Payback (with novel interventions labeled in orange)

The results demonstrate that AI-based consent enforcement shifts user behavior away from mindless 'Accept All', and that a preference-respecting data intermediary improves social welfare compared to existing broker models.

Thesis statement: This thesis argues that when we empower users to enforce their privacy preferences, they change their behavior from the current state of passive acceptance, and this change in behavior results in a new online advertising data economy with increased social welfare and greater alignment with user intent than the current opaque data ecosystem.

2 Background and Context

A comprehensive understanding of the online advertising data economy requires examining the technical, regulatory, and behavioral dimensions that shape user consent. This section provides the necessary context by first explaining the role of cookies in enabling data exchange between web browsers and servers. It then outlines the emergence of regulatory frameworks such as the GDPR and CCPA, which mandate mechanisms for user consent, implemented through cookie banners. However, these banners frequently rely on dark patterns that subvert user autonomy and obscure the true nature of data collection. Drawing on empirical research, we examine the effectiveness of these consent interfaces and the resulting consumer behavior, characterized by privacy fatigue and coerced participation. This background establishes the foundation for the research questions and claims addressed in the remainder of the thesis.

2.1 Cookies

Cookies are text files that enable communication between web servers and web-users' browsers. For example, when a web-user logs into a website, the server needs some way of communicating with the browser to deliver to it personalized content based on the logged-in user's account information. These cookies can contain several types of information, including but not limited to the web user's name, address, items in their cart, information about which ads they are to be served, and so much more (Wagner, 2020). Since their invention by Lou Montulli in 1995 at Netscape, cookies have been used across almost every website in the pursuit of providing to internet-users the best possible experience. However, these cookies have become the tools of data monopolies to exploit users for advertising revenue. They have been used for data collection, fingerprinting, and cross-site tracking, invading users' privacy and holding other companies to ransom, as there does not exist a better way to target consumers and monitor their user analytics. Tracking scripts have been insidiously injected by these companies on other sights as well through the use of Google Analytics tags, and the Meta Pixel.

There are two ways of classifying cookies:

1. **Single and Multi Session Cookies:**
   (a) **Single Session Cookies:** As the name suggests, this type of cookie lasts only the duration of a 'session,' which varies depending on the website on which it is set. Typically, this period is between 24 and 72 hours.
   (b) **Multi Session Cookies:** These cookies persist between browsing sessions, and are typically set by organizations such as the Federal Trade Commission.
2. **First and Third Party Cookies:**
   (a) **First Party Cookies:** These are cookies set by the website which the web user is currently visiting. They typically track items such as preferences, shopping cart items, or identification information.
   (b) **Third Party Cookies:** Cookies of this variety are set by entities other than the one represented by the website currently being visited by the web user. For example, a Google Analytics cookie is set on my browser while I am visiting despite CNet being neither representative of nor a subsidiary of Google or Alphabet.

The most pertinent cause for privacy concern is Third Party Cookies. These trackers are set by "partnered advertising companies, analytics companies, and companies that monitor user behavior over time" (Wagner, 2020), allowing advertisers to gather data across multiple sessions, sites, and devices. Subsequently, advertisers profit from this data by selling it to Data Brokers, who leverage publicly available information (e.g. age, demographics) and transaction histories (acquired from partnering companies) of the web user to assign this individual to a segment. They then sell these 'segments', groups of people who exhibit similar characteristics or behaviors, back to advertisers and marketing departments for enhanced targeting. 'Department Store Moms' and 'Married Women and Children who have previously bought Men's Accessories' are only two of the segments curated by Acxiom, a leading Data broker, for their 'Father's Day Package'. These segments

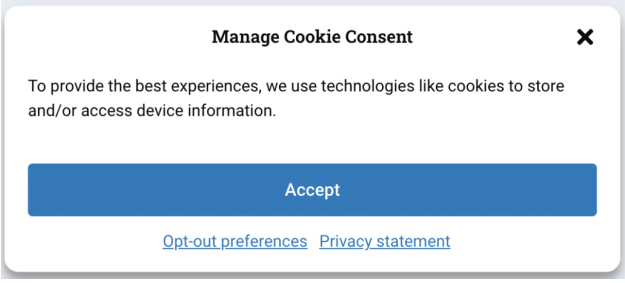


Figure 4: Standard Cookie Banner Layout Highlighting Common Elements

are additionally distributed to a variety of undisclosed entities including but not limited to U.S. and foreign governments.

Consider a web user, Anna, who visits the Babies R' Us website. Given this information, collected using a first party cookie, an advertiser might classify Anna as an expectant mother or a guest at someone's baby shower. Consequently, Anna is bombarded with advertisements for baby clothes and gifts, which she might find relevant, or even helpful. Now suppose that this website places a third-party cookie on Anna's browser that tracks her activity across several websites in the session, eventually learning that she subsequently visited Planned Parenthood, A College Prep and Study Website, and Baby Gap. This browsing data changes hands as described above, and is later combined with public records containing Anna's age, race, and zip code, ultimately leading to her segmentation as a 'Teenage High-School Mom.'

2.2 Regulation and the Rise of Cookie Banners

Privacy concerns evident in data flows such as those described above have led to the rise of various regulatory frameworks. Chief among them is the European Union's General Data Protection Regulation which came into effect in May 2018. It describes the current Data market as consisting of three types of entities ("What is GDPR, the EU's new data protection law?", 2018):

1. **Data Subjects:** Web users who engage with a particular platform or service, leading to the collection of their data.
2. **Data Controllers:** Owner or employee of an organization that decides why and how data subjects' personal data will be processed.
3. **Data Processors:** A third party that processes personal data on behalf of a data controller.

GDPR classifies the data collected by cookies as Personal Data. Hence, according to the law, companies are allowed to collect this data if and only if the data subject has provided "specific, unambiguous consent to process the data" ("What is GDPR, the EU's new data protection law?", 2018). This directly led to the proliferation of cookie consent banners and pop-ups (Wagner, 2020), such as the one in Figure 1. Similar comprehensive privacy regulation has been passed in several states in the United States, including California's CCPA, Colorado's CPA, Tenesee's TIPA, as well as legislation in Connecticut, Iowa, Indiana, Oregon, Montana, Texas, Delaware, Florida, New Jersey, and New Hampshire. Due to these state regulations, cookie banners have increased in number in the US.

Cookie Banners, however, have not empowered web users to wield control over their data and preferences. They are rife with Dark Patterns, 'user interface design choices that benefit an online service by coercing,steering, or deceiving users into making decisions that, if fully informed and capable of selecting alternatives, they might not make.' (Mathur et al., 2019) In their paper about Dark Patterns observed in advertising, sales, and data collection mechanisms on several top websites, Mathur et al highlight the consequences of these designs. These consequences range from annoyance and frustration to financial loss

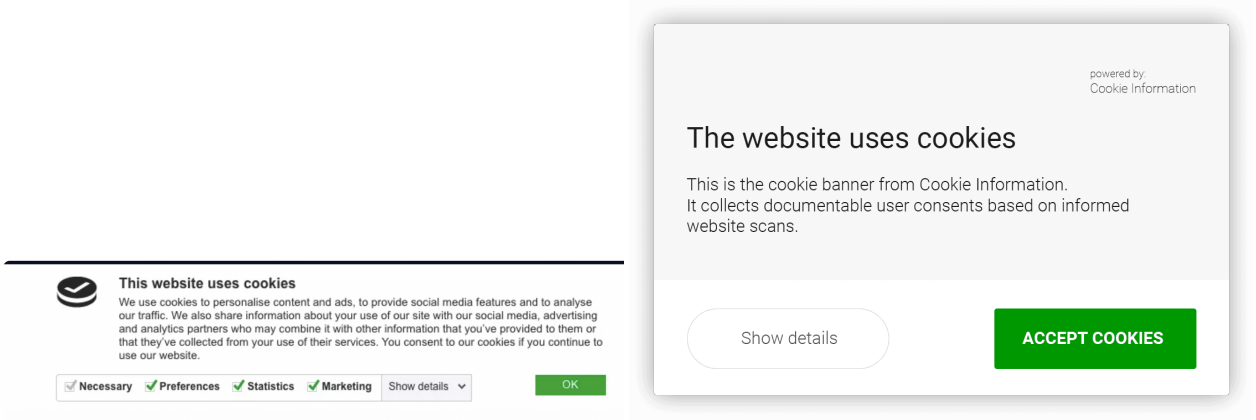


Figure 5: Examples of Dark Patterns (Manipulative Design Features) in Cookie Banners, *including pre-checked options and buried opt-out choice*

and misleading users into parting with troves of personal data.

2.3 Dark Patterns in Cookie Banners

A study of the top 10,000 websites in the United Kingdom revealed that only 11.8 percent of their cookie banners met the minimal requirements of GDPR. Among other stipulations, GDPR requires companies to make it as easy to opt-in as it is to opt-out of tracking and data collection, and as easy to withdraw their consent as it was to give it in the first place (Wagner, 2020).

In response to regulations, which pose "an existential threat to many companies in the online tracking and advertising industry," companies have chosen to implement cookie banners that "almost universally employ manipulative design to increase the likelihood of users consenting to tracking" (Narayanan et al., 2020). Some of the most common Dark Patterns in cookie banners include:

1. Using Pre-ticked Boxes (Figure 5)
2. Burying Decline/Reject buttons and requiring several clicks to make them appear. (Figure 5)
3. Tracking users before consent and even after pressing "Reject."

Similar patterns were found by Matte et al in their study "Do Cookie Banners Respect My Choice?" They implemented a web crawler using selenium-instructed Chromium to detect what cookies are stored in a user's browser based on their interaction with a cookie banner. For end users, they developed a browser extension, called 'Cookie Glasses' to show all the purposes and third party advertisers to whom consent is given by simply clicking 'Accept All' on cookie banners made by IAB Europe. Using this tool they found dark patterns like "pre-selected choices, consent stored before choice, no way to opt out, and non-respect of the user's choice" (Matte et al., 2020).

In the United States, companies like Onetrust and Trustarc have become market leaders in the field of Consent and Preference Management (the industry *nom du jour* for Cookie Banner merchants). However, they boldly advertise consent-rate optimization on their company websites, essentially meaning that they conduct extensive A/B testing to design cookie banners that make it most likely for users to opt-in to tracking and data collection (Nocera, 2022).

Cookie banners designed to bury the 'decline all' button, are leveraging techniques that manipulate user attention and decision-making. This manipulation could hinder the user's ability to provide truly informed

consent, as it exploits psychological vulnerabilities to guide choices in a predetermined direction, akin to the misdirection strategies used in magic to control audience perception, studied extensively in the field of Psychology. Hiding the 'Reject All' may be classified as "Perceptual - Non-Attentional - Masking" While Making the 'Accept All' button prominent may be classified as "Perceptual - Non-Attentional - Grouping" in the Taxonomy of misdirection (Kuhn et al., 2014).

The effectiveness of the aforementioned dark patterns on consent rates has been studied in the European Union. In "Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence", Nouwens et al examine this effect through an experiment conducted using 40 participants and the 8 most common cookie banners designs that they found after scraping the top 10,000 websites in the United Kingdom. They found that removing the opt-out (also displayed as 'reject all' or 'decline') button from the first page increased consent by 22–23 percentage points and that providing more granular controls (like necessary, marketing, performance, etc.) on the first page decreased consent by 8–20 percentage points (Nouwens et al., 2020). Other studies similarly suggest that when asked to opt-in rather than it being the default setting, less that 1 percent would do so (Narayanan et al., 2020).

A few of the granular cookie types, not covered by bulk options such as 'Accept All' and 'Reject All' include (Utz et al., 2019):

1. **Necessary:** Cookies that make the website work as intended.
2. **Personalization and Design:** Typically involving fonts and styles.
3. **Analytics:** Typically set by Google Analytics or similar entities to collect metrics on who is browsing the current website, how much traffic it generates, etc.
4. **Social Media:** Set by Facebook and Youtube.
5. **Marketing:** Set by advertising companies to track and measure engagement with advertisements.

2.4 Consumer Behavior toward Privacy and Cookie Banners

From these findings, it would seem that providing more options, or more information on cookie banners might yield the solution to the problem of informed consent. However, this opinion, and the studies carried out previously on cookie banner consent, ignore the fact that cookie banners pop up on almost every website. This annoyance, combined with the dark patterns, has led to current consumer behavior toward cookie banners. As Max Schrems, an Austrian privacy advocate states quite plainly, "No one reads cookie banners" (Nocera, 2022). Instead of tools of empowerment, cookie banners have become an almost useless exercise, which has led to a phenomenon termed 'Privacy Fatigue' (Choi et al., 2018). Web users are aware that their data is being collected and monetized, and that they are being manipulated, but feel powerless to stop it (Nocera, 2022).

Another study on Dark patterns and their effect, conducted by Luguri and Strahilevitz, suggests that aggressive dark patterns may lead to consumer backlash, while milder forms are more likely to cause users to provide uninformed consent. As early as 2005, it had been observed that "the ethical implications of internet marketing techniques necessitate a reevaluation of how relationships between businesses and consumers are established and maintained within a digital environment" (Palmer, 2005).

In a decidedly unscientific experiment, an artist, Risa Puno, stood in the middle of Times Square in New York City, and asked passers by whether they would trade some of their data in exchange for a cookie. 380 individuals gave up sensitive personal information — from fingerprints to partial Social Security numbers — for a sweet treat (Beckett, 2014). Privacy researcher Allesandro Acquisti commented on the event, stating that participants may have been more inclined to participate in the giveaway since it seemed fun and low-stakes, and unlikely that the data would be abused. This highlights the importance of transparency and trust in the data privacy space. Since Data brokers operate in an opaque, broadly unregulated market, there

are more privacy concerns associated with them.

To further understand web user behavior in terms of privacy, without the mitigating factor of cookie banners and the annoyance associated with clicking them, we turn to a study conducted by Acquisti in his 2013 study "What is privacy worth?" He conducted a field experiment in which mall shoppers were treated with 4 conditions based on two kinds of gift cards, one with 12 dollars but the purchase data would be collected, and one with 10 dollars, but the purchase would be anonymous. He found that subjects who had been given one of the cards and asked whether or not they would exchange it tended not to exchange it, which reinforces the power of defaults on privacy-related decision making. Furthermore, "the number of subjects willing to reject cash offers for their data was both significant in absolute terms and much larger in relative terms when they felt that their data would be, by default, protected than when they believed that their data would be, by default, revealed."

2.5 Research Questions and Claims

This thesis investigates several key questions in the area of data privacy and online data markets including:

1. **Can the use of modern Large Language Models mitigate the effects of Dark Patterns present in Cookie Banners today?**
2. **If provided with the tools to automatically manage cookie consent across all websites, do user preferences change from the current state (over 95% Accept All)?**

These questions are answered in Section 3 through the development of **Agent Papaya**, a browser extension that uses AI to scan and click cookie banners according to preset user preferences. This browser extension was implemented and released in late 2024, and we use real user-data to test these hypotheses. Next, we use Agent Papaya as a collection mechanism for an alternative data market, called **Papaya Payback**. Here, we claim that:

1. **The theory of Contextual Integrity can be operationalized through the development of a unified interface capturing users' data sharing preferences.**
2. **A data market with an intermediary that ensures users' data is shared according to their preferences produces greater utility and social welfare than the current data market without such an intermediary.**

These hypotheses are evaluated in Section 4 of this thesis using Agent-based modeling.

3 Enforcing Cookie Consent using Browser-based AI Agent

The objective of this section of the thesis is to answer the following research questions:

1. **Can the use of modern Large Language Models mitigate the effects of Dark Patterns present in Cookie Banners today?**
2. **If provided with the tools to automatically manage cookie consent across all websites, do user preferences change from the current state (over 95% Accept All)?**

First, we discuss the design of the solution, informed by user interviews and implemented as a chrome extension. Second, we discuss the implementation of the **Agent Papaya** chrome extension, with specific details on the use of Generative AI, as well as its evolution with available models. Lastly, we report our findings from data from over 229 real users, to see if this intervention produces the change we aim to create.

3.1 Design

3.1.1 Conceptualization

The main aim of this part of the Papaya project is to answer the broader question **'How does the provision of the facility for web users to manage and monetize their consent, preferences, and data affect their choices with respect to their data collection and privacy?'** Answering this question necessitates the creation of a mechanism for web-users to easily manage their cookie consent preferences across websites. This was accomplished by the creation of the **Agent Papaya Chrome Extension**.

First, user interviews were conducted to understand users' key frustrations and feelings toward cookie consent banners. The purpose of this process was to isolate the key user needs for the proposed extension and the functionalities the users wanted to automate. Although this tool was intended to serve all web users, there were low vision and dyslexic users in the interviews who had accessibility needs.

The following main user frustrations were identified, with quotes from our real interviewees:

1. **Bad Design & Dark Patterns**: The text in the cookie banners is too small, unclear, or overstimulating; Users want concise and easy-to-read information. Pop-ups can also be overstimulating or confusing, particularly if they appear abruptly and cover content. This leads to blindly accepting yes to get past it quickly.
   (a) *"I can't understand what it is asking me to do because it pops up and overstimulates me"*
   (b) *"I just blindly accept and move on"*
2. **Annoyance & Privacy Concern**: The users found the repeated occurrence of cookie banners very annoying, and would like for them to go away. However, they expressed concern at being informed that simply ignoring the cookie banner is considered as an opt-in and hence, clicking 'reject' or equivalent button was the only way to protect their privacy.

Based on these interviews, the following key user needs were identified:

1. **Set-it-and-forget-it Preference settings**: Users found it too annoying to set the same settings for each site, so try to create an overarching preference that allows them to apply it seamlessly across multiple sites *"Trying to change contrasts or colors to make it easier to read on individual websites is really obnoxious most of the times so I avoid it and will just leave the site."*
2. **Enable clear consent choices**: Provide a straightforward, keyboard-focus-able interface to set cookie preferences.

3. **Ensure Accessibility and Readability**: Incorporate large high-contrast text and clearly labeled controls.

4. **Offer Consistent, Non-Overstimulating Interaction**: Eliminate pop-ups and use simple layouts that don't overload users or cover the entire site.

Based on these user interviews, the idea of the **CookieMonster** extension, later renamed to **Agent Papaya** was born.

3.1.2 Novel Technical Architecture and Use of Agentic AI

I developed a browser extension that uses Agentic Artificial Intelligence to scan websites, detect whether a cookie banner is present on it or not, and if there is one, click the cookie banner with the user's preferences.

A Large Language Model (LLM) is prompted to identify buttons associated with consent preferences such as 'Accept All', 'Reject All', 'Manage Preferences', and more, so that the user's preferences can be applied effectively across every website.

I used one-shot learning as well as prompt-engineering to make sure that the LLM has as much context as I can provide it, so that it can reliably identify buttons. Here is the methodology that Agent Papaya uses, given that my preference is set to "Reject All".

1. First, the LLM is given a role using the prompt:

```
    "You are a friendly, helpful assistant specialized in detecting buttons
corresponding to options such as 'accept_all', 'reject_all', 'manage_preferences',
etc. in cookie consent banners."
```

2. Next, the external cookie banner HTML element on the target website is isolated using heuristics like whether it is a pop-up, div, or other element, as well as whether it contains certain keywords like onetrust, didomi, truste (names of popular CMPs), or even words like gdpr, cookie, and privacy.

3. This external banner is fed into the LLM, which is additionally prompted to identify buttons like Accept, Reject, Manage, Confirm, etc. The prompt to identify the buttons is shown here:

```
    Here is a cookie banner on a website:\n${cleanedBanner}\nWhich HTML element
corresponds to the following types:
    accept_all: <Description>
    reject_all: <Description>
    manage_my_preferences: <Description>
    Please return the output in JSON format with the following keys: 'text', 'id',
'class'
    Here is an example from a different website:
    <Example output JSON>
```

4. The results from the LLM are correctly formatted into JSON to produce a dictionary of buttons for the target website. This is stored for future use. Here is the sample output for the external cookie banner of www.chicagobooth.edu:

```
    'chicagobooth': {
        'external_buttons':
            {'manage_my_preferences':
                {'text': 'Manage My Cookies',
                'id': 'None',
                'class': 'toggle-modal'},
            'accept_all':
                {'text': 'Accept All Cookies',
                'id': 'None',
```

```
                'class': 'toggle-modal'},
            }
        }
```

5. Note that, in the current example, www.chicagobooth.edu has only a manage_my_preferences option and no direct reject_all. Hence, the logic dictates **Agent Papaya** to click the manage_my_preferences option.

6. Following this, Agent Papaya scans for an internal cookie banner, using heuristics again. The internal banner is found.

7. The internal banner is scanned using the LLM again, this time with a different prompt that looks like this:

```
    Here is the HTML element for the cookie preferences modal on a website:\n
        ${internalBannerHTML}\n
        Extract the UI elements associated with:
        - "reject_all" (Reject all cookies)
        - "accept_all" (Accept all cookies)
        - "confirm_my_preferences" (Confirm settings)
        - "marketing" (Marketing cookies toggle)
        - "performance" (Performance cookies toggle)

        Return as a JSON list of objects. For example:
        [{"option_name": "marketing", "element_type": "checkbox", "id": "option1",
"class": "toggle-checkbox"},
         {"option_name": "reject_all", "element_type": "button", "text": "Reject
All", "id": null, "class": "button"}]
```

8. The internal banner is scanned and the internal buttons are updated in the JSON object for www.chicagobooth.edu. Hence, the updated JSON object looks like this:

```
    "chicagobooth.edu": {
        "external_buttons": {
            "manage_my_preferences": {
                "text": "Manage My Cookies",
                "class": "toggle-modal"
            },
            "accept_all": {
                "text": "Accept Cookies"
            }
        },
        "internal_buttons": [
            {
                "option_name": "reject_all",
                "element_type": "button",
                "text": "Reject All Cookies",
                "id": null,
                "class": "toggle-modal btn-lg btn btn-maroon reject-cookies"
            },
            {
                "option_name": "accept_all",
                "element_type": "button",
                "text": "Accept All Cookies",
                "id": null,
                "class": "toggle-modal btn-lg btn btn-maroon accept-cookies"
            },
            {
                "option_name": "confirm_my_preferences",
                "element_type": "button",
                "text": "Confirm My Selections",
                "id": null,
                "class": "btn btn-maroon btn-lg toggle-modal accept-cookies"
            },
```

```
            {
                "option_name": "marketing",
                "element_type": "checkbox",
                "id": "targCook",
                "class": "switch"
            },
            {
                "option_name": "performance",
                "element_type": "checkbox",
                "id": "perfCook",
                "class": "switch"
            }
        ]
}
```


9. Following this, Agent Papaya clicks the reject_all option, thus executing my preference despite the Reject button being buried in the granular preferences modal.

Although the methodology for using AI has remained quite constant throughout the development of the project, the underlying LLM has not.

In this rapidly-evolving field, Large Language Models have become commodities, and hence, the choice of model on which the aforementioned process is left to the following *Implementation* section.

3.2 Implementation

3.2.1 Interface Evolution

The design of the Papaya browser extension evolved significantly over the course of its development, informed by iterative user feedback and real-world deployment. The original extension, **CookieMonster**, had a minimal dashboard and a simple cookie preference panel. Over time, both features were expanded and reimagined in the rebranded version, **Agent Papaya**, to increase usability, transparency, and accessibility.

In its first version, CookieMonster allowed users to set their cookie preferences using two basic checkboxes—one for marketing cookies and one for performance cookies. While this design was functional, it lacked clarity, and users often expressed uncertainty about what the options meant or how they impacted browsing behavior. In response, the cookie preference interface was redesigned to include four distinct, clearly labeled presets:

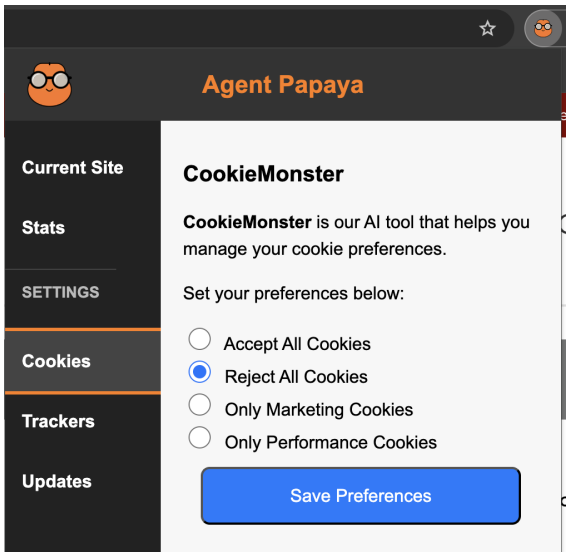
- **Accept All Cookies**
- **Reject All Cookies**
- **Only Marketing Cookies**
- **Only Performance Cookies**

This change allowed users to quickly and confidently express their preferences without needing to understand nuanced distinctions or legal language. The updated interface also improved contrast and accessibility, supporting keyboard navigation and better screen reader compatibility.

The dashboard also saw major improvements. Initially, it simply displayed the number of cookie banners handled and the number of unique websites visited. This limited view offered little insight into what actions had been taken or why. Based on user feedback—particularly from participants who wanted to "verify what the extension did on each site"—the dashboard was split into two views:

1. **Current Site** — displays the exact action taken on the cookie banner for the current website, such as "Reject All" or "Only Performance Cookies."
2. **Stats** — aggregates cookie-handling data across all websites, including total banners handled and distribution of chosen preferences.

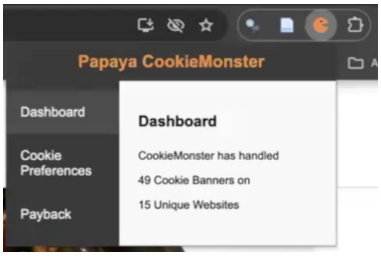
(a) CookieMonster: Early version with only two preference toggles



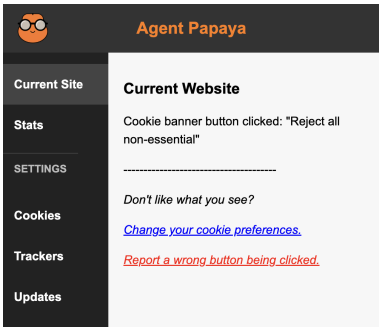
(b) Agent Papaya: Expanded preferences with four clear options

Figure 6: Cookie Preference Interface Evolution from CookieMonster to Agent Papaya, highlighting *Improved User Control Over Preferences*

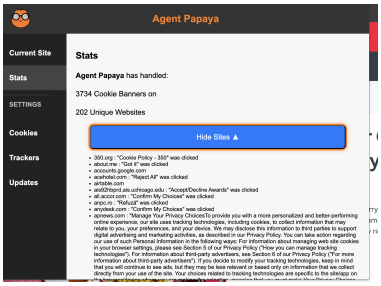
A third improvement was the introduction of a lightweight notification panel within the dashboard that communicates updates, including instructions for enabling in-browser AI capabilities. This turned the dashboard from a passive stats viewer into an interactive and informative user hub.



(a) CookieMonster: Basic stats-only dashboard



(b) Agent Papaya: Current Website Tab



(c) Agent Papaya: Stats tab

Figure 7: Dashboard Interface Evolution from CookieMonster to Agent Papaya, *from Basic Stats to Real-time Feedback*

All of these changes were grounded in feedback from real users during early deployments. For example, some users with low or no vision noted that they struggled to interpret the older UI, leading to improvements in navigability and contrast. Others requested "a way to check what it clicked," which directly inspired the separation of the dashboard into "Current Site" and "Stats" tabs. These iterative, user-informed improvements reflect Papaya's core philosophy: to give users clarity, control, and confidence in how their data is handled.

3.2.2 Technical Implementation Version 1: Lambda + GPT 3.5/4o

Initially, in May 2024, the cutting-edge LLM was OpenAI's GPT 3.5 accessible through an API. Hence, I set up an AWS Lambda function to scrape websites and store their cookie banner button data in AWS S3. I decided to use the majestic million dataset, a set of the 1 million most visited domains in the world as the source of websites. Later, with the release of GPT-4o, I switched to using this API to access the LLM.

I was able to acquire the cookie banner data of 644 websites, and this static data set was shipped with the extension, then called **CookieMonster**. However, there were some problems. I noticed that sometimes, it would find the incorrect button, and that data would be missing for some websites. These problems could only be solved by real-time AI-based cookie banner detection in the browser itself.

3.2.3 Technical Implementation Version 2: Gemini Nano

In late 2024, Google introduced a developer preview of Gemini Nano ("Built-in AI — AI on Chrome", n.d.), its on-device Large Language Model (LLM) for Chrome. This allowed developers to run AI models directly in the browser without requiring server-side infrastructure or internet connectivity. Recognizing the potential of this development, I signed up for the developer trial and participated in a hackathon ("Papaya CookieMonster", 2024) where developers were given early access to the Chrome Dev build with Gemini Nano support enabled.

I quickly adapted Agent Papaya's architecture to use Gemini Nano for real-time cookie banner analysis. This was a significant improvement over the previous server-based Lambda architecture, as it ensured better privacy (no server-side processing), reduced latency, and improved accuracy due to context-awareness on the client side. Agent Papaya now uses Gemini Nano as its default LLM.

This change has also eliminated the costs associated with repeatedly scraping websites and acquiring cookie banner button data, as Gemini Nano runs client-side, hence updating my growing dataset of cookie banners for free. The user's copy of the extension updates their local cache of cookie banners they have seen, scanned by Gemini Nano and periodically sends its updates to AWS S3 to update the global cookie banner dataset. Furthermore, the dataset is totally anonymized, with no user identifiers, ensuring total privacy.

3.2.4 A Brief Aside on Global Privacy Control

Global Privacy Control (GPC) is an emerging privacy standard that allows users to express a universal opt-out preference for data sale and tracking via a browser signal. Agent Papaya includes support for GPC by automatically sending the appropriate HTTP header ('Sec-GPC: 1') to all websites the user visits.

While GPC is not yet universally respected, some jurisdictions like California (under CPRA) and Colorado (under CPA) require companies to honor it. By integrating GPC, Agent Papaya ensures regulatory compliance for users in these regions and increases the likelihood that user preferences are respected even before interacting with a cookie banner.

In addition, GPC provides a legal framework that reinforces Agent Papaya's ethical position, offering users another line of defense against unauthorized tracking.

3.3 Evaluation

3.3.1 Efficacy of AI-based Cookie Banner Scanning

This evaluation was designed to answer the question: **Can the use of modern Large Language Models mitigate the effects of Dark Patterns present in Cookie Banners today?**. To provide a robust quantitative assessment of Agent Papaya's core AI capability, accurately identifying and classifying cookie banner elements, we constructed a dedicated test set of 87 distinct websites, sampled to represent a wide range of designs and underlying Consent Management Platforms.

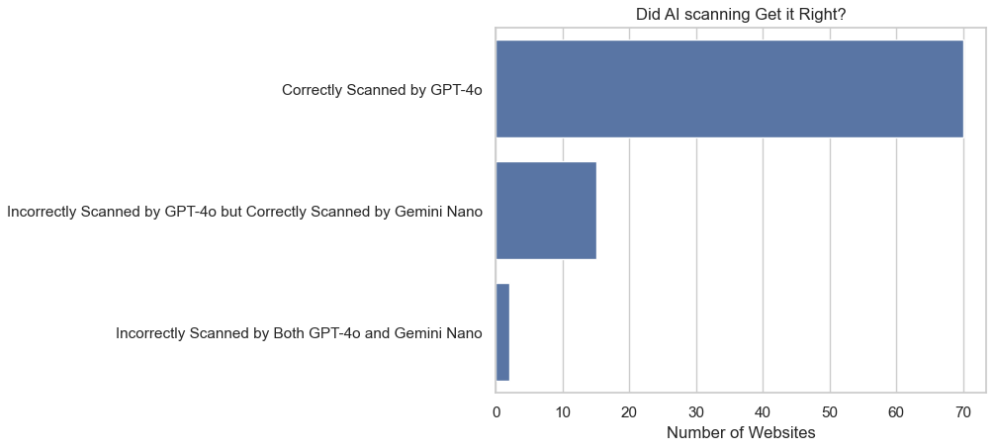


Figure 8: Accuracy of AI-Based Consent Enforcement on Randomly Sampled Websites

The accuracy of the static scrape (Version 1) using GPT-3.5 and GPT-4o was ascertained by manually verifying this sample of 87 websites. Since this data is packaged with the extension, we manually navigated to each website on this list and determined whether **Agent Papaya** clicked the correct button, or sequence of buttons, or not. Out of 87 websites, 70 websites were correctly clicked using data from the static scrape. A further 15 websites could have their static scrape data easily replaced by accurate scans by the real-time Gemini Nano AI. The button data for the remaining 2 websites remained inaccurate and was replaced by data manually filled, so as to improve the user experience of using the extension. Furthermore, since correctly clicking the banner once is insufficient to determine whether the scan has noted all buttons correctly, the study was repeated for each different preference available in the interface: Accept all, Reject all, Only marketing, Only performance.

From this, we can isolate a few key metrics.

Agent Papaya's dual AI-based Cookie Banner scanning approach, using the static GPT-4o scrape as well as Gemini Nano in real-time as a fallback, achieves 97.7% accuracy when scanning and identifying buttons on Cookie consent banners.

Within Agent Papaya's novel AI architecture, we see that the accuracy of our static scrape using GPT-4o, was (70/87) roughly 80 percent. Furthermore, the accuracy of Gemini Nano can be assumed to be at least (15/17) or 88 percent. However, this is a sample biased by the fact that GPT-4o model clearly got it wrong. Hence, the true accuracy of Gemini Nano is likely higher than 88 percent. Further complicating matters, Gemini Nano runs independently in each user's individual browser window. Hence, it does not consistently get the cookie banner scan right every time. For the same cookie banner passed into Gemini Nano 10 times, it would only result in the same output 9 times on average, from my evidence.

All of these metrics underscore the room for improvement in the context of quantifiable metrics for AI-based scanning. However, this is beyond the scope of this thesis and left for future work.

In addition to accuracy, latency is also a major factor in the deployment of real-time AI-based scanning of cookie banners. For all cookie banners which have already been scanned and stored within either the static dataset or the cached Gemini Nano results, the banner is clicked within 2 seconds of webpage load. This is remarkable, as it is faster than the average human can click a cookie banner, especially since they tend to appear significantly after the rest of the Document Object Model (DOM) loads. However, when the cookie banner is one that the extension has not "seen" before, it tends to slow down. If my preference is *reject_all* and there is a *reject_all* button on the main cookie banner, it still scans the external banner relatively quickly (inside 4 seconds). However, if there is no reject in the external banner, it clicks *manage_my_preferences* which causes the internal banner to load. Typically, these are large HTML elements with hundreds of char-

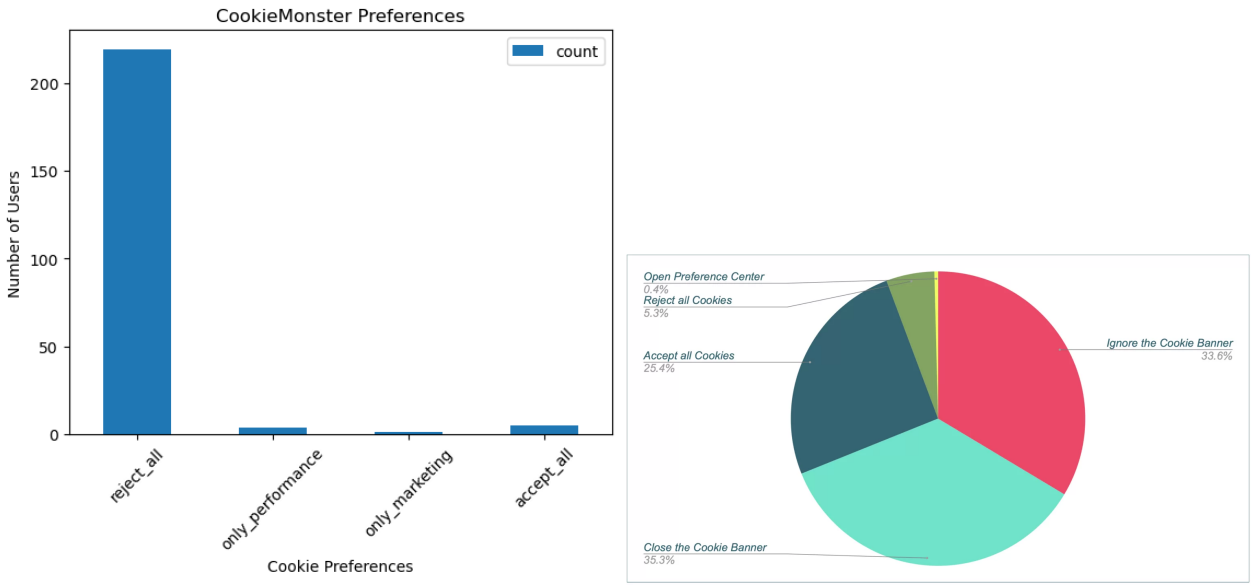


Figure 9: Agent Papaya User Preferences, showcasing *High rates of 'Reject All' Adoption*, and Industry Study showing *Low Opt-Out Rates* by Advance Metrics ("Cookie Behaviour Study – 5 years after GDPR", n.d.)

acters. This can really put a strain on Gemini Nano, as it sometimes requires upto 8 seconds to process the internal banner. In rare instances (the 2 websites in the manual verification study done above included), the internal banner has so much text that it exceeds Gemini Nano's context window, resulting in the user having to click the internal banner themselves.

3.3.2 User Data and Feedback

This evaluation was designed to answer the question: **If provided with the tools to automatically manage cookie consent across all websites, do user preferences change from the current state?**

We collected data on user preferences set in the extension in the following format:

User_SrNo	Marketing_cookies	Performance_cookies
000001	F	F

The user seen above would be considered as having a reject_all preference. Overall, 95.6% of **Agent Papaya** users set their preference to reject_all, with 2.2% each being accept_all or other granular preference. This is in stark contrast to prevalent cookie banner studies "Cookie Behaviour Study – 5 years after GDPR", n.d., which show that only 5.3% of users click Reject on cookie banners.

Since the extension was published on the Chrome Webstore in August 2024, it has acquired 229 total users, and 56 weekly active users. We've seen over 70% retention rates, which indicate that users really find value in using **Agent Papaya**. I anticipate that Agent Papaya will gain even more users when Google officially launches Gemini Nano in-browser AI for all chrome users. Given that it is still in developer preview, users must initiate a slightly cumbersome setup process to make sure they can access my developer preview of Gemini Nano.

Furthermore, user interviews have confirmed Agent Papaya's usefulness. All users stated that the extension positively impacted their experience with cookie banners after the setup was properly completed. This reflects how the extension was able to effectively combat cookie banners on sites that the users visited. Participants also stated how easy it was to change their cookie settings, and how convenient it was

to be able to set the settings and forget about it for the future. Specifically, the users with dyslexia and low-vision/cataract said that the banner did not create any visual stress. An older adult user who preferred keyboard-navigation also really liked that they could interact with the extension using their keyboard.

Some areas for improvement include knowing whether or not the extension was working correctly, and feedback mechanisms. For example, when a user navigated to Stack Overflow, the cookie banner did not show up at all. It was unclear whether this was because the extension had actually worked, or if the user had previously set cookie settings on this website and had forgotten about it, and so forth. This feedback has informed the continuous development of the tool.

3.4 Connection to the Broader Thesis

The Agent Papaya extension (formerly known as CookieMonster) offers a powerful yet simple tool for web users to automatically set and enforce their data-sharing preferences across websites. However, its utility extends beyond cookie consent management. Due to its privileged position in the user's browser, the extension can also serve as a privacy-preserving data collection mechanism.

Specifically, Agent Papaya can access users' browsing behavior—such as visited URLs, time spent on pages, and interaction patterns—which may be used to infer user preferences, interests, and behavioral traits. Crucially, if most users configure their cookie preferences to `reject_all`, and websites respect these preferences, the behavioral data accessible to Agent Papaya would be unavailable to third-party trackers. This results in a unique data asset, accessible only to the entity (a private company, nonprofit, or government) offering the extension.

This exclusivity addresses a key limitation of Data Unions discussed in the literature: their difficulty in acting as true intermediaries between users and data consumers. Typically, Data Unions struggle to offer data that cannot be obtained independently by companies. Agent Papaya changes this dynamic—by blocking conventional data flows while enabling voluntary, controlled data sharing, it empowers a Data Union model in which users retain control and derive value from their data.

This positions Agent Papaya not just as a consent tool, but as the infrastructural foundation for a new kind of data economy—one where privacy, exclusivity, and user agency are aligned.

4 User Consented Data Monetization

The objective of this section of the thesis is to assess the following claims:

1. **The theory of Contextual Integrity can be operationalized through the development of a unified interface capturing users' data sharing preferences.**
2. **A data market with an intermediary that ensures users' data is shared according to their preferences produces greater utility and social welfare than the current data market without such an intermediary.**

First, we discuss the design of the Data Preferences Statement interface to operationalize the constructs of contextual integrity as well as user research on data preferences into a concrete interface. Next, we discuss two economic models, one modeling the current online advertising data market, and the Papaya Payback market. Lastly, we run an agent-based simulation evaluating whether the Papaya Payback market, governed by Data Preference Statements, produces greater utility and social welfare than the current market.

4.1 Design

4.1.1 Conceptualization: The Ethical Data Broker

While Agent Papaya empowers users to reject data tracking, it also disrupts the data supply chain that powers the web economy. This raises a natural question: *What comes after "Reject All"?* **Papaya Payback** aims to provide the answer in the form of a new kind of ethical data broker that not only respects individual preferences but creates mutual value for users and platforms alike.

Unlike traditional brokers that profit from opaque and non-consensual data flows, Papaya Payback is focused on user consent and compensation. Drawing inspiration from **Data Unions** and **Contextual Integrity**, it serves as a trusted intermediary that allows web users to selectively share data in return for payment, with full control over *who* gets access to *what* data and *for what purpose*.

4.1.2 Development of "Data Preferences Statement" Interface

To operationalize consent in a user-friendly way, Payback introduces the "Data Preferences Statement" interface. Users construct simple, natural-language-like statements such as:

> "I am willing to share **anonymized** data with **retailers** for the purpose of **product recommendation**."

These statements are internally parsed into tuples of `(Platform, Purpose, Anonymization)`, which guide how and when Payback makes user data available to interested platforms. Users can submit multiple statements to fine-tune their preferences or opt out entirely.

This interface integrates seamlessly with Agent Papaya, so users can set these preferences at install time and later edit them from a lightweight dashboard.

Motivation

This interface was developed to test the claim that **the theory of Contextual Integrity can be operationalized through the development of a unified interface capturing users' data sharing preferences.**

As such, the design of this interface draws heavily from Helen Nissenbaum's theory of *Contextual Integrity* Nissenbaum, 2004, which posits that privacy is preserved when information flows conform to legitimate social norms defined by context-specific parameters such as actors, information types, and transmission principles. Rather than asking users to approve or reject every individual data transaction (which is cognitively burdensome and unsustainable at scale), we allow users to define *contextual norms* upfront via declarative statements.

This approach was also informed by findings from user interviews. Several participants voiced frustration with the excessive granularity of contextual-integrity inspired flows. One participant remarked, *"I don't want to make a decision every time. I just don't want advertisers to have my data."* Another shared, *"I'd be fine sharing if I knew what it was being used for."* These insights led to the decision to group platforms into broad categories (e.g., `Advertisers`, `Retailers`) and limit the scope of purposes (e.g., `Targeted Ads`, `Product Recommendation`) to a small, cognitively manageable set.

Design Principles

The design of this interface was guided by three key principles:

1. **Simplicity:** By bundling related purposes and companies into coarse-grained categories, users avoid decision fatigue while still retaining control.
2. **Expressiveness:** Users can express fine-grained preferences using combinations of `(A, P, R)` parameters (Anonymization, Platform Type, Reason for Use).
3. **Transparency:** Every choice is accompanied by plain-language explanations and examples of what data is shared and how it is used.

Interface Walkthrough

The Data Preferences Statement interface was developed as a multi-step onboarding flow linked from the Papaya browser extension. It allows users to create preferences in the following order:

Step 1: Choose the level of data anonymization. Users begin by selecting whether they wish to share *Anonymized* or *De-anonymized* data. Anonymized data excludes personal identifiers and is shared in aggregate form. De-anonymized data includes identifiable information such as name and email address. (Figure 11)

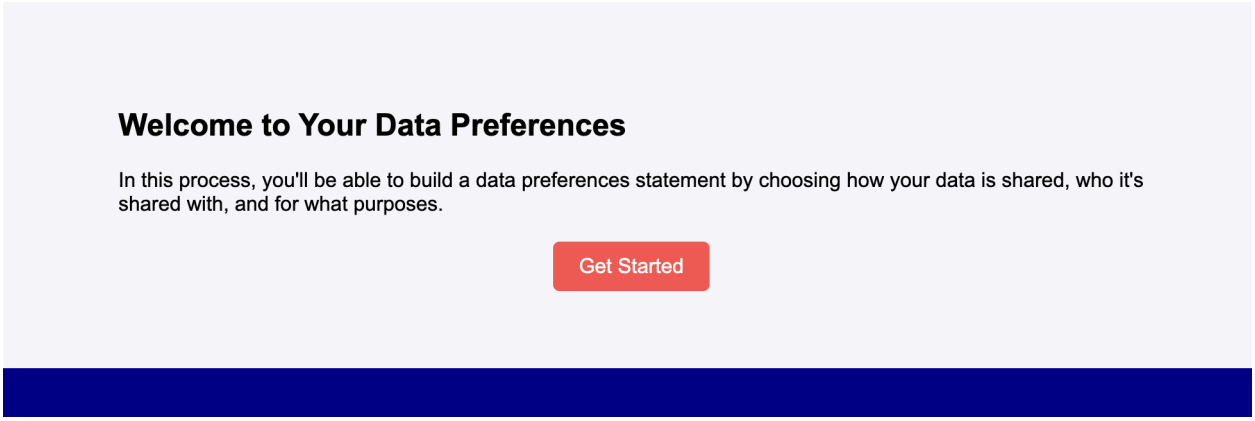


Figure 10: Introduction to the Data Preferences Statement (DPS) Interface for User Consented Data Brokerage

Step 2: Choose the platform types you are willing to share data with. Users then select from grouped categories of companies, such as `Advertisers` or `Retailers`. These categories reflect common distinctions in data usage across the ecosystem. (Figure 12)

Step 3: Choose the intended purposes for data use. Finally, users select the purposes for which their data may be used. Options include `Targeted Advertisement`, `Email Marketing`, `Market Research`, and `Product Recommendation`. Clicking "More Info" expands detailed descriptions and examples. (Figure 13)

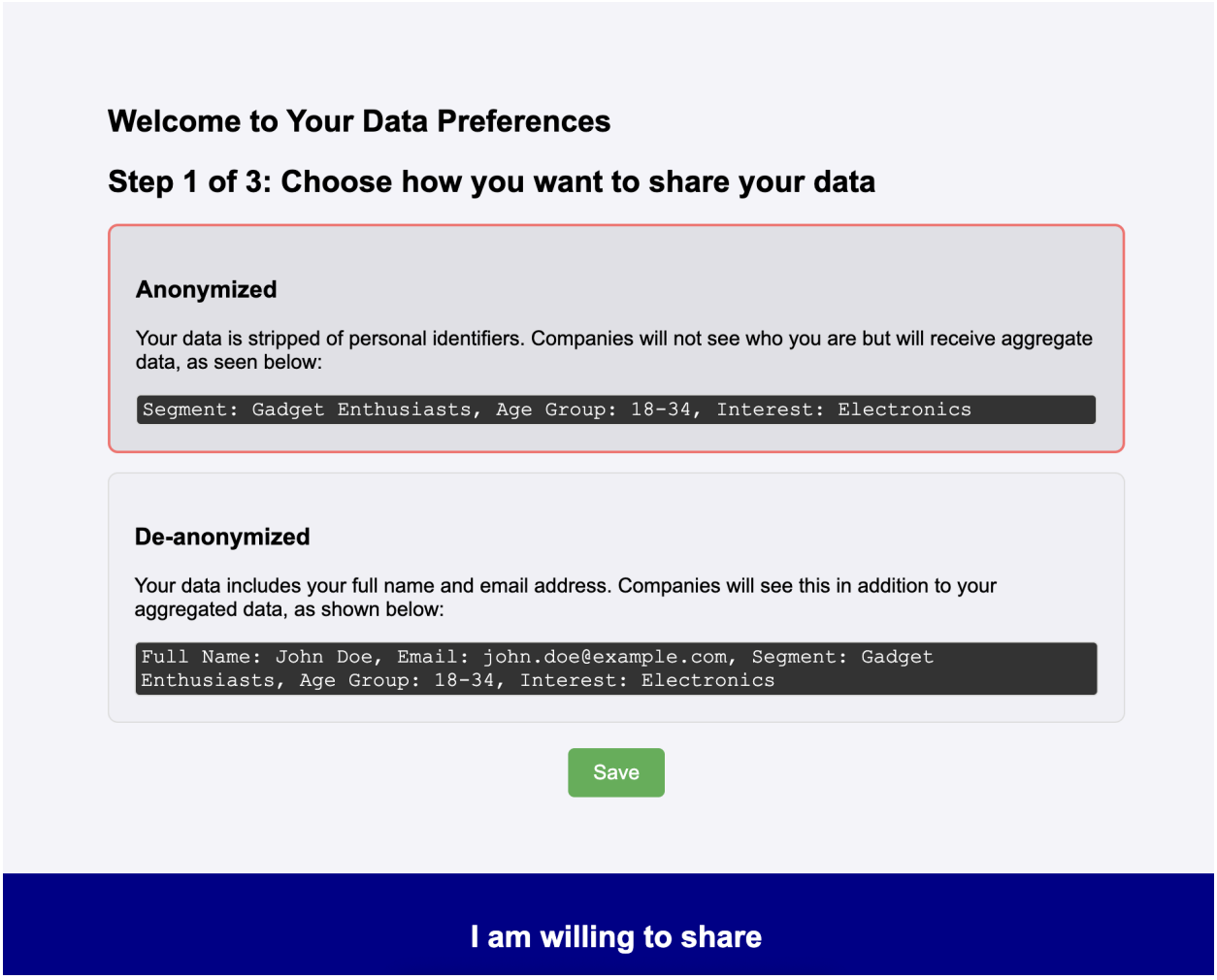


Figure 11: DPS Step 1: Choosing Anonymization Level

Generated Preference Statement: Based on the user's selections, a final preference statement is generated that explicitly states their sharing rules. For instance, a user might declare: *"I am willing to share anonymized data with advertising companies and retailers for the purposes of product recommendation and targeted advertisement."* (Figure 14)

Visualizing the Data Flow

To promote transparency, we also show users a visual walkthrough of how their data is collected, processed, and shared. This includes examples of what data is collected (e.g., browsing history), how it is transformed into segments, and what companies ultimately see. (Figure 15)

This interface feeds directly into our simulation and economic model, defining the constraints under which data sharing can take place in the Papaya Payback Market.

4.2 Implementation

4.2.1 Economic Model: IPDM

The default structure of today's data market can be modeled as an **Individual-Platform Data Market (IPDM)**. In this model:

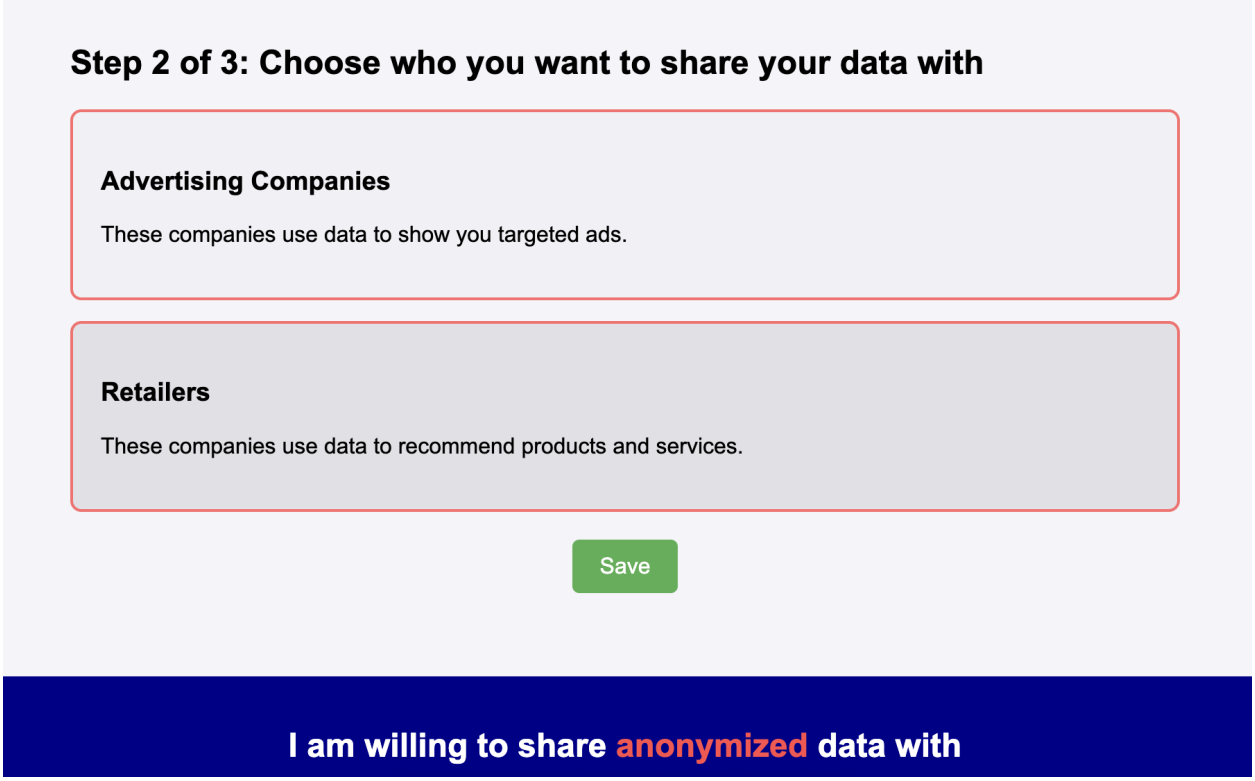


Figure 12: DPS Step 2: Selecting Target Platform Types

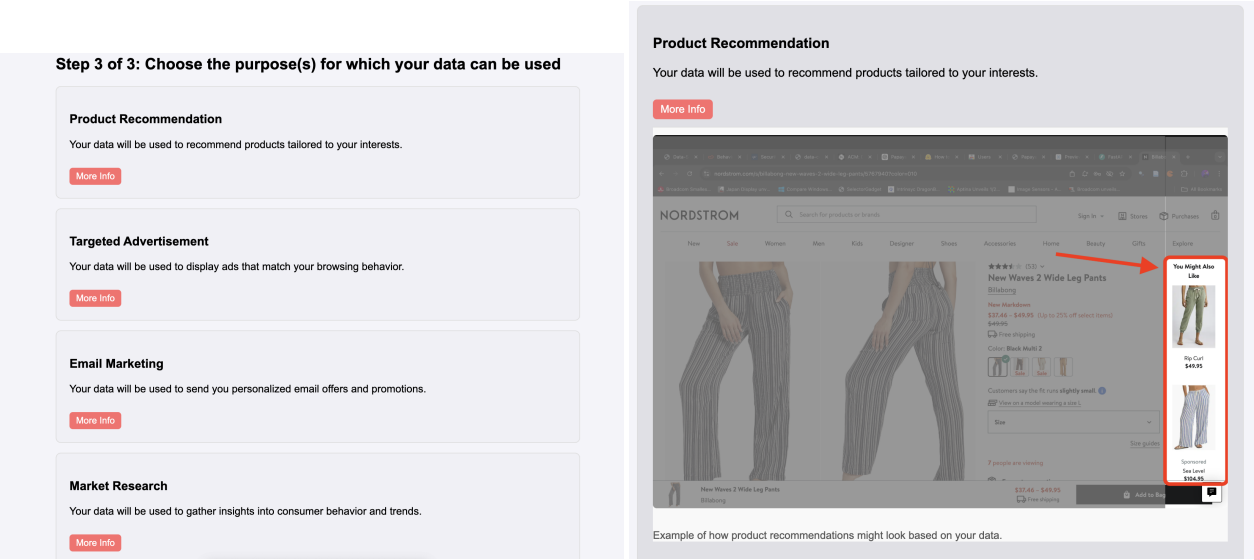


Figure 13: DPS Step 3: Choosing Purposes for Data Use

- Each user makes individual decisions about sharing data.
- Platforms offer a fixed budget for data and bid according to the expected utility of a data transaction.
- No notion of group leverage or collective bargaining exists.
- Privacy costs are borne entirely by the user; they are not internalized by the platform.

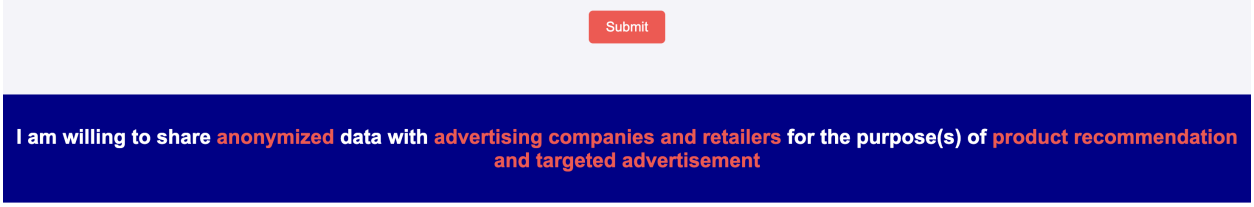


Figure 14: Final Output of DPS: Personalized Consent Statement

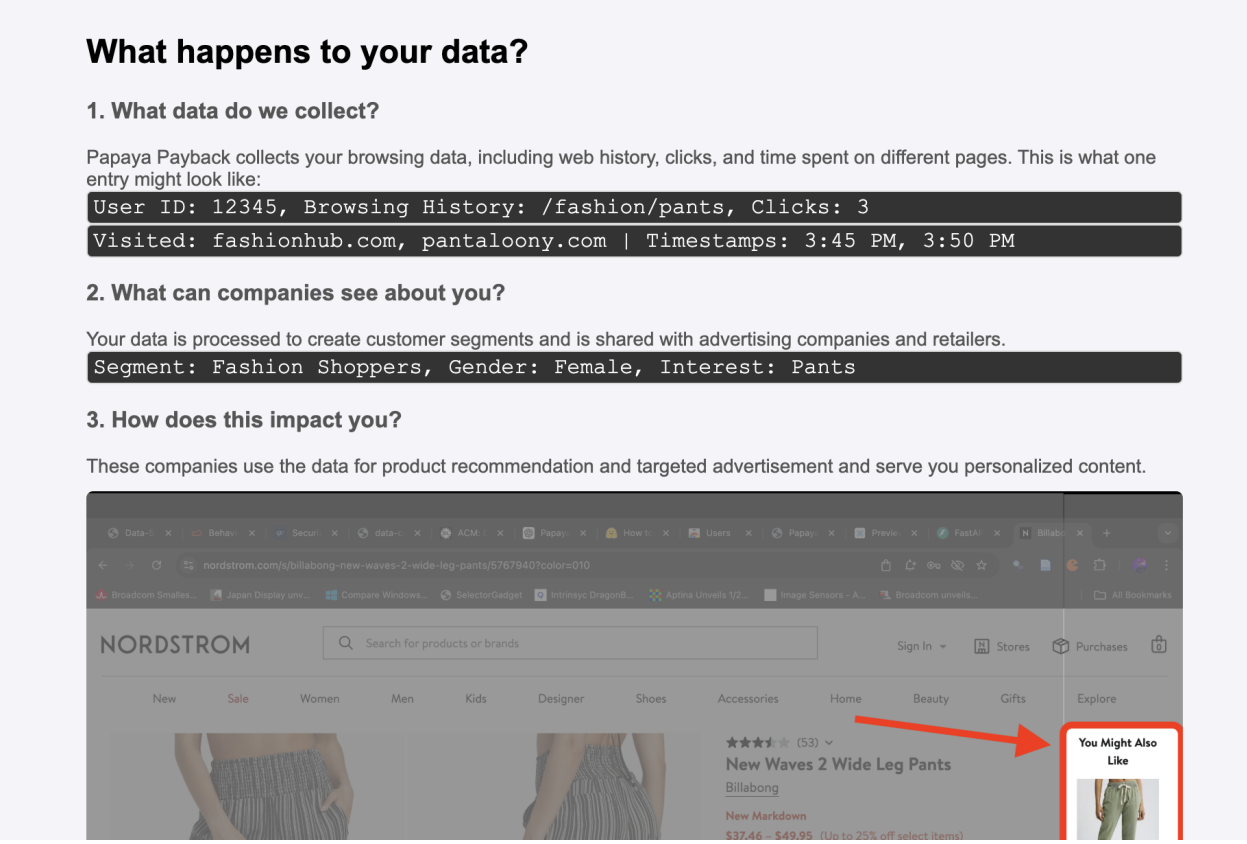


Figure 15: Visualization of User Data Flow from Agent Papaya to Ethical Marketplace

Simulated results in this regime show that without clear incentives, users may decline to participate, especially when empowered by Agent Papaya or GPC signals, leading to diminished data supply and reduced platform utility.

4.2.2 Economic Model: Union + CI

In contrast, the **Papaya Payback Market** introduces a broker that internalizes principles from both **Data Unions** (Hardjono and Pentland, 2019) and **Contextual Integrity** (Nissenbaum, 2004):

- **Data Union Effects:** When many users share data under similar conditions, platforms receive a multiplier incentive (a union effect), encouraging group-based access over piecemeal tracking.
- **Contextual Integrity:** Users define the context of data sharing via (P, R, A) preferences, ensuring that each transaction respects the expectations of the social setting in which it occurs.

The broker takes a cut (e.g., 10%) of each transaction to fund operations. However, this cut is offset by

greater platform utility due to targeted, context-aligned access, and by user willingness to participate due to higher perceived fairness and control.

4.3 Evaluation

This evaluation was designed to test the claim made earlier that **A data market with an intermediary that ensures users' data is shared according to their preferences produces greater utility and social welfare than the current data market without such an intermediary.**

4.3.1 Simulation

To assess the viability of the Papaya Market relative to the traditional IPDM model, we developed a detailed agent-based simulation. This simulator models 100 users and two types of platforms (Advertisers and Retailers), running across multiple rounds. It is designed to explore key outcomes such as user earnings, platform utility, and overall social welfare.

Simulation Overview

The simulation evaluates two models:

- **IPDM (Individual-Platform Data Market):** Based on Dr. Raul Castro Fernandez's theoretical market, where users make binary participation decisions per platform.
- **Papaya Market:** Extends IPDM with support for Data Unions and Contextual Integrity. Users can specify granular, purpose-based preferences.

Each user is initialized with one to three Data Preference Statements, each corresponding to a tuple of `(Platform, Purpose, Anonymization)`. The simulation proceeds in the following stages per round:

1. Privacy Sensitivity Calculation For each (P, R, A) tuple, we compute a privacy sensitivity score $v(P, R, A)$, which is higher for anonymized data and advertiser platforms, and lower when more purposes are approved.

2. Privacy Cost Function Each data-sharing option is also evaluated for its privacy cost $\rho(P, R, A)$. Costs are higher for de-anonymized data, advertisers, and invasive purposes such as targeted ads and email marketing.

3. Platform Bidding Platforms bid for access to user data, offering prices adjusted by purpose-specific profitability δ_R, the user's privacy sensitivity, and anonymization penalties.

4. User Participation In the IPDM model, if any one tuple for a platform fails the cost-benefit test (i.e., price $\leq$ cost), the user opts out from that platform entirely. In the Papaya Market, users make per-purpose sharing decisions, allowing them to selectively participate.

5. Transactions and Union Effects When a transaction occurs, users are paid the platform's price minus a broker commission. In the Papaya Market, a "union multiplier" is applied to data shared under popular (P, R, A) conditions, reflecting enhanced value through collective bargaining.

6. Payout and Utility Calculation Each round ends with an update to user earnings and remaining platform budgets. These are used to track per-round and cumulative metrics.

4.3.2 Results

The results of the simulation can be seen in Figure 16. Our simulation yields three core insights:

1. **IPDM fails to scale in a privacy-conscious world.** Users empowered by tools like Agent Papaya or GPC signals tend to reject all data requests unless contextually appropriate options are provided. This leads to low participation and severely diminished platform utility.

2. **Papaya Market increases participation and platform utility.** Because users can share data selectively for specific purposes with specific platforms, participation rises. Despite Papaya taking a commission, user earnings remain stable thanks to increased volume and union effects. Platform utility is also much higher, since data purchases reflect actual willingness and contextual appropriateness.

3. **Papaya enables high social welfare.** When we sum user earnings and platform utility to calculate total social welfare, Papaya consistently outperforms IPDM. Notably, in high-commission scenarios, Papaya continues to provide strong user earnings due to increased platform engagement.

Conclusion: In contrast to IPDM's binary, brittle model, Papaya's CI- and union-driven framework allows users to say yes on their own terms. In doing so, it preserves privacy, enables control, and still supports an economic model for the web. This suggests that the future of data markets is not binary opt-in or opt-out — it's programmable participation.

Figure 16: Simulated Comparison of Social Welfare and Participation Between Traditional Individual-Platform Data Market (IPDM) and Papaya Payback Data Market, highlighting **higher social welfare** and **higher participation** in the Papaya Market

Interpretation: These findings demonstrate the promise of Payback as a viable, privacy-preserving alternative to today's exploitative data economy. With Agent Papaya enforcing consent and Payback brokering ethical exchanges, users can finally say "no" to dark patterns and "yes" to purposeful data sharing on their own terms.

Crucially, if tools like Agent Papaya become widespread, the default will shift to *no tracking*. In that world, platforms cannot function without genuine user consent. Payback provides the mechanism for that consent to be earned, respected, and compensated, creating a win-win for users and platforms alike.

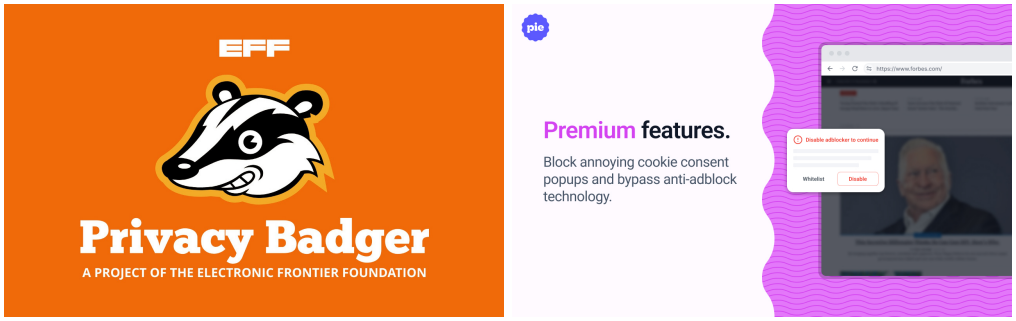


Figure 17: Comparative Overview of Existing Privacy Extensions and Their Limitations

5 Related Work

This research clearly exhibits the need for consumer-friendly services that keep web users' privacy as their main objective. However, as seen in other works, the tools and platforms that should be tasked with protecting privacy are deeply consumer-unfriendly and are designed to increase web users' consent rates. There are several solutions that have been proposed to alleviate the privacy concerns rife among web users.

5.1 Alternative Browser-based Privacy Tools

Consumers may use an array of browser extensions to protect their data privacy online.

Privacy Badger, created by the digital rights group, Electronic Frontier Foundation (EFF), is one of the most popular choices. It is even one of the recommended extensions on setup of Mozilla Firefox, the third-most popular browser in America. Privacy Badger operates at the network level, blocking cookies and certain advertisements by preventing a website from loading content from certain domain names, stored in a static list in the extension. This static list is updated periodically as the EFF maintains a separate repository to scan websites from time to time and discover new tracking codes. This approach has proven immensely popular, but lacks a certain dynamism.

Pie Adblock is a new player in the chrome extension arena, launching in November 2024 with a flurry of online advertisements. Their value proposition revolves around blocking advertisements, and paying users when they choose to see partner ads. However, from anecdotal evidence on the usage of the product by myself and some friends, we discovered that the revenue we could generate by allowing ourselves to be shown ads was minimal, and, while it claimed to block cookie pop-ups, it did not seem effective at doing so.

In fact, the sub-genre of extensions that interact with cookie banners on behalf of the user spans a rich and varied history. **I don't care about Cookies** was a solo developer project in the early days of GDPR to block or auto-accept cookie banners. It was since acquired by Avast for an undisclosed fee. This project, initially open-source spawned an open-source sequel called **I still don't care about Cookies**, which performed the same function. These tools were doubtlessly useful in their time. However, their value to protecting user privacy was minimal. After all, to allow users to protect their privacy, they would have to actively click "Reject All" or the equivalent option, since, in most jurisdictions, anything but an explicit opt-out is taken as an opt-in.

Consent-o-matic is a tool developed by researchers I've cited earlier from the Aarhus University Nouwens et al., 2020. While they were able to obtain more depth, clicking the "Reject All" on cookie banners produced by one of three major Consent Management providers (CMPs) (Companies like OneTrust), they sacrificed breadth as they are unable to handle custom cookie banner implementations as well as those created by other CMPs. **DuckDuckGo AutoConsent**, in a similar vein, is implemented as a Javascript API rather than an extension itself, allowing makers of chrome extensions to use their package to detect and

click cookie banners made by specific CMPs. Other tracker blockers like Disconnect, Ghostery, and uBlock Origin block banner ads or pop-up ads and depersonalize browsing sessions.

These extensions have failed to address the pressing user need of a tool that is able to automatically click "Reject All" or the button(s) appropriate to user preferences on every website seamlessly. The key problem they have failed to solve for is the variety and veracity of cookie banners at large. Cookie banners can look different, have different options, and be implemented by various different CMP companies, which makes them both difficult to detect and interact with using rule-based approaches.

5.1.1 Policy Methods

Several researchers in the papers I have reviewed above advocate for policy to solve the privacy problem. However, the prevailing hypothesis is that this is not in the interest of Data collection and brokerages who spend billions of dollars on lobbying to prevent this from happening. A simulation of different policy regimes that restricted different kinds of user information from being used for targeted advertising found that restricting more intrusive variables for targeting lowered ad effectiveness and led to fewer potential purchases (Aziz and Telang, 2016). Additionally, even if legislation were passed the question of enforcement remains unanswered. Many of the current Dark patterns violate existing regulation yet go unpunished (Luguri and Strahilevitz, 2021).

5.2 Alternative Data Markets

5.2.1 Data Unions

At the same time, researchers in several fields (Economics, Legal studies, Computer Science) have proposed ideas on an institution known as a Data Union, sometimes referred to as Data Collaborative, or Data Cooperative. One definition of the concept proposes that a "data cooperative refers to the voluntary collaborative pooling by individuals of their personal data for the benefit of the membership of the group or community" (Hardjono and Pentland, 2019). The idea is that if enough web users pool together their data in ways that obey both individualized and group consent, they would have greater leverage over the collection and use of their data by platforms, who would have to engage with the new entity, the Data Union, to gain access to the data of those web users who participate in it. In a similar way to credit unions, these entities could negotiate better services, or Data monetization benefits for their users.

5.2.2 Contextual Integrity

Another promising theoretical framework in privacy research is Contextual Integrity (CI), developed by Helen Nissenbaum Nissenbaum, 2004. CI challenges the notion that privacy is about secrecy or control in the abstract. Instead, it frames privacy as the appropriate flow of information within specific contexts, based on contextual norms and expectations. According to this framework, privacy violations occur not when data is shared per se, but when it is shared inappropriately — when the transmission principle (i.e., the conditions under which data flows from sender to recipient) does not match the expectations embedded in the social context of the exchange.

This model maps particularly well onto the current cookie banner ecosystem, which fails to communicate the full context in which data will be shared. Banners often present only vague categories (e.g., "marketing", "analytics") or bury these in sub-menus, creating a mismatch between user expectations and actual data flows. In contrast, a CI-compliant system would empower users to set fine-grained data preferences that align with the who (data recipient), the what (data type), and the why (purpose for data use).

Some attempts have been made to construct mechanisms for web-users to express their consent preferences in a more granular way that a blanket 'Accept' or 'Reject'. Kuru et al propose an Agent-based framework whereby preferences on very granular data points (Age, Occupation, Gender, etc.) are elicited from both users and enterprises, as well as their valuation of each of these data points, and then each of the agents deploy rule-based preference computational agents to engage in a negotiation to decide whether

the data point is shared or not (Kuru et al., 2024). However, this would be incredibly cumbersome for both web-users and enterprises. As seen previously in this literature review, it is difficult for web-users to value their data *a priori*. As mentioned in the theory of Contextual integrity, the context in which the data sharing request is made (as in, which platform is asking for this data) matters significantly. When users value their data, they are also making a value judgement on the trust they are placing in the company they are sharing their data with.

Furthermore, Ogunniye & Kökciyan allow software agents to negotiate or argue to resolve privacy preference conflicts (Ogunniye, 2023). In their system, each agent can reason about whether a proposed data transfer violates contextual norms and explain its decision. However, it is unclear from this work how such norms would be changed and updated through users' inherent preferences.

5.3 Gaps in these solutions

The above potential solutions are great pieces of work and seem to hold promise individually. However, there are some clear areas for improvement in relation to these previous solutions in the field of data privacy.

5.3.1 Browser Tools

It is clear from previous work in this space that Browser Extensions are a popular solution. However, they do not fully solve the problem of dark patterns in cookie consent banners due to the following problems:

1. **Ignore == Accept:** Most jurisdictions in the world today that have Privacy Law are opt-out jurisdictions. Users must explicitly opt-out of cookie-based tracking; otherwise, they will be regarded as having opted in. Agent Papaya makes sure it explicitly clicks the Opt-Out option, and maintains evidence of the same.
2. **Rule-based Heuristics are not enough:** There is enough variety in cookie banner design, options, CMP provider, that rules are insufficient to capture a significant percentage of all cookie banners online. A more advanced solution would be required to dynamically detect cookie banners, scan them to reveal their options, and click the one that corresponds to the user preference. Agent Papaya uses the latest Large Language Models to facilitate Cookie banner scanning, and is hence more likely to accurately capture cookie preference options.

5.3.2 Alternative Data Market: Unions and Contextual Integrity

Data Unions and Contextual Integrity both provide solid theoretical frameworks for how these interventions in the current Data Market could yield a fairer and more equitable data economy. However, neither of these ideas have been implemented into a real intermediary between web-users and platforms. This is because of two key gaps:

1. **Lack of Collection Mechanism:** Companies today collect and store user data at will. If an intermediary were to try to collect any user data, they would require the user to download all of their data from the company and upload it to their servers, which is a heavy burden to place on the user. Other solutions, like partnering with companies to link data to Authentication, the approach favored by Tim Berners-Lee's Solid project, require buy-in from companies who have no incentive to give up control of their data. There is a pressing need for a low-friction mechanism to collect vast quantities of user data across platforms. Agent Papaya delivers this via a chrome extension.
2. **Preference Elicitation:** There exists no unified interface for users to set their data sharing preferences in a simple and understandable manner. Cookie consent banners are deceptive and their options are unclear. Similarly, privacy policies are verbose and unintelligible. There is an urgent need for a simple

user interface that allows users to determine **who** gets access to **what data** of theirs for **which purposes**. Papaya Payback creates this system via the Data Preferences Statement Interface.

6 Key Contributions

6.1 Agent Papaya

Agent Papaya significantly advances the current state of automated consent tools by directly addressing two core problems identified in the literature:
(1) the deceptive and inconsistent nature of cookie banners, and
(2) the limited utility of existing autoconsent tools that lack transparency or customizability.
By using generative AI (first GPT-4, later Gemini Nano), Agent Papaya dynamically identifies consent options on banners across a wide range of formats, even when dark patterns or non-standard interfaces are used.

Unlike tools such as *I Don't Care About Cookies* or Consent-O-Matic, which rely on hard-coded rules and offer little visibility into what actions were taken, Agent Papaya supports real-time feedback through its "Current Site" tab and gives users the ability to pre-configure preferences using one of four preset options. It also surfaces stats over time, creating a transparent experience rather than a black-box one.

Compared to Consent Management Platforms (CMPs) like OneTrust, which dominate market share, Agent Papaya offers a more effective solution for users. While OneTrust may be widely deployed, their banners are often custom-implemented by different companies and may vary dramatically in design, behavior, and compliance. Agent Papaya, by comparison, works across CMPs, enabling seamless and scalable privacy enforcement.

6.1.1 Key Results

In a field deployment of the Agent Papaya extension (n = 229), we observe that

1. **95.6% of users selected 'Reject All'** as their default, contrasting starkly with benchmark studies showing ¡6% rejection in the wild.
2. **56+ weekly active** users found value in our extension, with **over 70% retention** rate, suggesting ongoing user engagement with set-it-and-forget-it privacy tools.
3. Agent Papaya's LLM based cookie banner scanning approach is capable of obtaining **97.7% accuracy** when correctly identifying and classifying the buttons of Cookie Consent banners.

6.2 Papaya Payback

Papaya Payback builds on the behavioral transparency of Agent Papaya to offer a new paradigm for ethical data brokerage. Rather than simply blocking tracking, Payback introduces a data economy grounded in user preferences, contextual integrity, and economic fairness. The simulation results demonstrate that a Payback-style model, especially when integrated with a Data Union and Contextual Integrity framework, can outperform traditional data broker logic (as exemplified by IPDM) on multiple dimensions—user utility, platform utility, and overall social welfare.

By introducing the notion of "privacy cost" and modeling trade-offs between purpose, anonymization, and platform type, Payback allows us to reason systematically about data flows and incentives in a way that most adtech models ignore.

6.2.1 Key Results

We simulated a market of 100 users and 2 platform types under two market regimes: Traditional IPDM (Individual-Platform Data Market) and the novel Papaya Payback market. The Papaya Market resulted in

1. **Higher participation**,

2. **Greater Platform Utility**, and

3. **Increased Social Welfare**

as compared to the Traditional IPDM.

This supports our claim that a user-consented, preference-based data sharing market can outperform coerced consent-based models.

7 Discussion and Conclusion

7.1 Integration: How Agent Papaya and Payback Fit Together

Together, Agent Papaya and Papaya Payback solve the full-stack problem of ethical data sharing:

- **Data Collection**: Agent Papaya is uniquely positioned in the browser to observe user browsing behavior in a privacy-preserving way.
- **Data Uniqueness**: Since Agent Papaya blocks conventional cookie-based tracking, the behavioral data it captures becomes exclusive to the Papaya platform.
- **Preference Elicitation**: Through a simplified onboarding and statement-based model, users express contextual preferences that drive data access policies.
- **Marketplace Creation**: Papaya Payback models a functioning data marketplace that uses these preferences to assign value and cost to user data in a transparent, ethically aligned way.

This integration is what makes Papaya Privacy more than just another privacy project. It is a complete rethinking of how users, platforms, and intermediaries interact.

7.2 Limitations

7.2.1 Limitations of Agent Papaya

While the design and functionality of Agent Papaya represent a step forward, the current user base remains limited. Early users likely exhibit selection bias—they are disproportionately privacy-conscious and tech-savvy. This may not represent the general population.

Further, Gemini Nano—the in-browser LLM powering Agent Papaya—is still in developer preview and not yet widely accessible. Until it becomes publicly available in Chrome, the extension's full capabilities are limited to a small subset of users.

Finally, the AI itself is not fully deterministic. On occasion, Gemini Nano has returned different button selections for the same banner across runs, likely due to non-determinism in its architecture or prompt conditioning. This poses challenges for reproducibility and debugging.

7.2.2 Limitations of Payback

The Payback simulation makes a number of simplifying assumptions. User preferences are treated as static and fixed, whereas in reality, preferences may evolve over time or in response to contextual shifts (e.g., news coverage, platform behavior).

Additionally, the values assigned to privacy sensitivity, purpose valuation, and anonymization cost are all heuristically chosen. While these provide useful intuition, they are not derived from large-scale empirical studies or real-world transactions. The simulation demonstrates feasibility and potential—not exact forecasts.

7.3 Potential Future Work

Several promising directions remain for future development:

- **Deploying Payback**: A real-world launch of Papaya Payback would allow us to study how users set their preferences, whether users choose to monetize their data, and how companies respond to this new paradigm.

- **User Studies on Preferences**: A formal qualitative or mixed-methods study could explore users' willingness to share data under different anonymization/purpose conditions, strengthening the input parameters of the simulation.
- **Wider Agent Papaya Release**: Once Gemini Nano is widely available in Chrome, Agent Papaya can be released to a broader audience with minimal server costs or privacy concerns.
- **Launching Papaya Privacy Co.**: The creation of a formal nonprofit or cooperative organization to serve as the Data Union / Contextual Integrity intermediary is a natural next step. I've begun pursuing this via participation in the Booth School's Social New Venture Challenge (SNVC), where this idea was received with early interest.

7.4 Conclusion

This thesis proposed and implemented a two-pronged solution to the challenges of online consent and ethical data use. Through Agent Papaya, users can now control their cookie preferences across websites automatically and transparently, reducing friction and dark pattern exploitation. Through Papaya Payback, users can express preferences for how and when their data should be monetized—building a market that respects privacy, context, and agency.

The findings from our simulation suggest that such a market can deliver better outcomes for both users and platforms than the current opaque, extractive status quo. In doing so, this work charts a path toward a more just and participatory data economy—one where users are not just subjects of surveillance but stakeholders in the value they help generate.

In the words of Justice Brandeis, "Sunlight is the best disinfectant." By surfacing how data flows, clarifying consent, and aligning incentives, Papaya Privacy Co. aims to bring sunlight to an otherwise shadowy system.

8 Data and Code Availability Statement

The code for this work is available in the Papayaverse GitHub Organization.

9 Bibliography

References

Aziz, A., & Telang, R. (2016, March 31). What is a digital cookie worth? https://doi.org/10.2139/ssrn.2757325

Beckett, L. (2014, October 1). *How much of your data would you trade for a free cookie?* [ProPublica]. Retrieved March 22, 2024, from https://www.propublica.org/article/how-much-of-your-data-would-you-trade-for-a-free-cookie

Built-in AI — AI on chrome [Chrome for developers]. (n.d.). Retrieved November 20, 2024, from https://developer.chrome.com/docs/ai/built-in

Choi, H., Park, J., & Jung, Y. (2018). The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior*, *81*, 42–51. https://doi.org/10.1016/j.chb.2017.12.001

Cookie behaviour study – 5 years after GDPR [Advance metrics]. (n.d.). Retrieved April 2, 2025, from https://www.advance-metrics.com/en/blog/cookie-behaviour-study/

Data broker is selling location data of people who visit abortion clinics. (n.d.). Retrieved April 5, 2025, from https://www.vice.com/en/article/location-data-abortion-clinics-safegraph-planned-parenthood/

FTC surveillance pricing study indicates wide range of personal data used to set individualized consumer prices [Federal trade commission]. (2025, January 17). Retrieved April 5, 2025, from https://www.ftc.gov/news-events/news/press-releases/2025/01/ftc-surveillance-pricing-study-indicates-wide-range-personal-data-used-set-individualized-consumer

Hardjono, T., & Pentland, A. (2019, May 21). Data cooperatives: Towards a foundation for decentralized personal data management. Retrieved May 13, 2024, from http://arxiv.org/abs/1905.08819

Kuhn, G., Caffaratti, H. A., Teszka, R., & Rensink, R. A. (2014). A psychologically-based taxonomy of misdirection [Publisher: Frontiers]. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01392

Kuru, A. E., Aydogan, R., Ozturk, P., & Razeghi, Y. (2024, October 14). Sharing personal data through agent-based negotiation framework: Preference modeling and empirical analysis. https://doi.org/10.2139/ssrn.4987022

Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis*, *13*(1), 43–109. https://doi.org/10.1093/jla/laaa006

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, *3*, 81:1–81:32. https://doi.org/10.1145/3359183

Matte, C., Bielova, N., & Santos, C. (2020). Do cookie banners respect my choice? : Measuring legal compliance of banners from IAB europe's transparency and consent framework [ISSN: 2375-1207]. *2020 IEEE Symposium on Security and Privacy (SP)*, 791–809. https://doi.org/10.1109/SP40000.2020.00076

Narayanan, A., Mathur, A., Chetty, M., & Kshirsagar, M. (2020). Dark patterns: Past, present, and future: The evolution of tricky user interfaces. *Queue*, *18*(2), Pages 10:67–Pages 10:92. https://doi.org/10.1145/3400899.3400901

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, *79*(1), 119. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10

Nocera, J. (2022). How cookie banners backfired. *The New York Times*. Retrieved March 26, 2024, from https://www.nytimes.com/2022/01/29/business/dealbook/how-cookie-banners-backfired.html

Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020). Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376321

Ogunniye, G. (2023). Contextual integrity for argumentation-based privacy reasoning.

Palmer, D. E. (2005). Pop-ups, cookies, and spam: Toward a deeper analysis of the ethical significance of internet marketing practices. *Journal of Business Ethics*, *58*(1), 271–280. https://doi.org/10.1007/s10551-005-1421-8

Papaya CookieMonster [Devpost]. (2024, December 2). Retrieved April 4, 2025, from https://devpost.com/software/papaya-cookiemonster

US-russian citizen sentenced to 12 years for $52 donation to ukrainian charity [France 24] [Section: europe]. (2024, August 15). Retrieved October 3, 2024, from https://www.france24.com/en/europe/20240815-russian-court-sentences-us-citizen-prison-52-donation-ukraine-charity

Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019). (un)informed consent: Studying GDPR consent notices in the field. *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 973–990. https://doi.org/10.1145/3319535.3354212

Wagner, P. (2020, December 8). Cookies: Privacy risks, attacks, and recommendations. https://doi.org/10.2139/ssrn.3761967

What is GDPR, the EU's new data protection law? [GDPR.eu] [Section: GDPR Overview]. (2018, November 7). Retrieved May 11, 2024, from https://gdpr.eu/what-is-gdpr/