computers

MDPI

Article

A User-Centered Privacy Policy Management System for Automatic Consent on Cookie Banners [†]

Lorenzo Porcelli *, Michele Mastroianni, Massimo Ficco and Francesco Palmieri

Department of Computer Science, University of Salerno, 84084 Fisciano, Italy;
mmastroianni@unisa.it (M.M.); mficco@unisa.it (M.F.); fpalmieri@unisa.it (F.P.)
* Correspondence: lporcelli@unisa.it

[†] This paper is an extended version of our paper published in the 23rd International Conference on Computational Science and Its Applications (ICCSA 2023), Athens, Greece, 3–6 July 2023.

Abstract: Despite growing concerns about privacy and an evolution in laws protecting users' rights, there remains a gap between how industries manage data and how users can express their preferences. This imbalance often favors industries, forcing users to repeatedly define their privacy preferences each time they access a new website. This process contributes to the privacy paradox. We propose a user support tool named the User Privacy Preference Management System (UPPMS) that eliminates the need for users to handle intricate banners or deceptive patterns. We have set up a process to guide even a non-expert user in creating a standardized personal privacy policy, which is automatically applied to every visited website by interacting with cookie banners. The process of generating actions to apply the user's policy leverages customized Large Language Models. Experiments demonstrate the feasibility of analyzing HTML code to understand and automatically interact with cookie banners, even implementing complex policies. Our proposal aims to address the privacy paradox related to cookie banners by reducing information overload and decision fatigue for users. It also simplifies user navigation by eliminating the need to repeatedly declare preferences in intricate cookie banners on every visited website, while protecting users from deceptive patterns.

Keywords: Personal Information Management System (PIMS); User Privacy Policy Management System (UPPMS); privacy paradox; Large Language Model (LLM); Generative Pre-trained Transformer (GPT); cookie banner; Transparency and Consent Framework (TCF); General Data Protection Regulation (EU GDPR)

check for updates

Citation: Porcelli, L.; Mastroianni, M.; Ficco, M.; Palmieri, F. A User-Centered Privacy Policy Management System for Automatic Consent on Cookie Banners. *Computers* **2024**, *13*, 43. https://doi.org/10.3390/computers13020043

Academic Editor: Paolo Bellavista

Received: 28 December 2023
Revised: 26 January 2024
Accepted: 28 January 2024
Published: 1 February 2024

1. Introduction

The introduction of the General Data Protection Regulation (GDPR) has had a significant impact on the online advertising industry, giving users more control over their personal data [1]. A key feature of the GDPR is its broad applicability, extending beyond European companies to all global companies that process the data of EU citizens. The GDPR strengthens consumers' rights to control their data, but its implementation, particularly in relation to cookie banners, presents challenges for companies and requires significant investments in legal, privacy and web development resources [2].

The IAB Europe's Transparency and Consent Framework (TCF: https://iabeurope.eu/transparency-consent-framework/, accessed on 18 December 2023) was developed to address GDPR complexities, particularly in managing user data processing permissions. The TCF provides standardization guidelines, but lacks specific directions for cookie banner implementation, leading users through a complex consent process. Companies commonly employ tactics like deceptive dark patterns and cookie paywalls to encourage cookie acceptance. These include making the "Accept all" option more prominent in cookie banners and implementing complex procedures for users opting out. Cookie paywalls pose another indirect influence, where access to content is conditional on data processing consent or subscription payment [3,4].

Despite the GDPR's efforts, users often face tedious decision-making processes with cookie banners. Providing comprehensive information can lead to unappealing banners, conflicting with users' desire for quick website access. The time-consuming nature of understanding and responding to these banners contributes to the privacy paradox, where privacy-concerned users often consent to all tracking cookies [5].

Personal Information Management Systems (PIMSs) (https://edps.europa.eu/data-protection/our-work/subjects/personal-information-management-system_en, accessed on 18 December 2023), also referred to as Personal Information Management Services, have emerged as a solution to these challenges, providing centralized tools for organizing personal information and facilitating informed decision making [1]. A PIMS ideally automates the updating and communication of user preferences to data controllers and processors, but faces technical and regulatory obstacles, including establishing connections and consensus on a common technical language. Currently, PIMS implementations are very basic, mainly as browser extensions for Google Chrome and Mozilla Firefox. These extensions manage consent decisions by blocking cookie banners, opting out of data processing activities, or automatically responding to data processing permission requests. However, the use of extensions that block all cookie banners may prevent consent for all data processing, potentially affecting website functionality.

In our preceding work [6], we presented an initial proof of concept illustrating how a PIMS, augmented by a Large Language Model (LLM), can analyze the contents of cookie banners and formulate interaction rules. Building upon this foundation, the current study advances the concept by delineating the specifications for a User Privacy Policy Management System (UPPMS). The UPPMS, a specialized derivative of a PIMS, is designed to aid users in crafting, generating, and implementing their personal privacy policies.

We introduce a standardized methodology for users to define their cookie policy preferences. This has enabled the development of a framework for rule generation that incorporates customized LLMs. We delineate a systematic process for the application of these rules. This approach streamlines the application of cookie policy preferences and helps to avoid deceptive patterns. Additionally, we conducted an experiment to test user behavior when visiting a website with cookie banners for the first time. We compared the time it takes for humans and the UPPMS to express preferences and close the banner.

The remainder of this article is structured as follows. Section 2 introduces Large Language Models and examines the context related to user privacy signals and the privacy paradox. Section 3 discusses related works, providing insights into existing research and developments in the field. Section 4 describes the requirements and how the UPPMS, leveraging a Large Language Model, enables an end-to-end process from defining the user's policy to its application during navigation. The experiments are described in Section 5, and the results are presented in Section 6. Section 7 discusses the implications of the study. Finally, Section 8 draws conclusions and proposes future research directions.

2. Background

2.1. Large Language Models

Large Language Models (LLMs) represent an advanced class of machine learning models, particularly in the field of Natural Language Processing (NLP). They are built upon deep neural architectures, such as Transformers, which have demonstrated remarkable capabilities in understanding, generating, and manipulating natural language [7].

The concept of Transformers, forming the foundation of LLMs, was introduced by Vaswani et al. [8]. This work laid the groundwork for the development of large-scale language models such as GPT (Generative Pre-trained Transformer), which undergoes extensive unsupervised pre-training followed by supervised fine-tuning on specific tasks [9]. Training these models involves extensive exposure to vast text corpora, allowing them to acquire a profound grasp of language subtleties and a broad general knowledge. These models excel in generating coherent and contextually relevant texts, addressing queries, translating languages, and executing text comprehension tasks. For instance, in [10], it has been demonstrated that LLMs effectively comprehend HTML, significantly enhancing accuracy and efficiency in tasks such as semantic classification of HTML elements, description generation for HTML inputs, and autonomous web navigation.

In our work, we employ customized LLMs to perform tasks that necessitate a syntactic and semantic understanding of a web page. We fine-tune LLM models to identify cookie banners within a web page and, through text comprehension, discern the buttons on the cookie banners and the types of actions they enable for accepting or rejecting cookies.

2.2. Cookies

Cookies are small pieces of data that are stored in a user's web browser. They are also referred to as web cookies, browser cookies, Internet cookies, or HTTP cookies. While different taxonomies for cookies exist, especially from a technical perspective, a detailed discussion of cookie categorization is beyond the scope of this article. What is relevant to our discussion, however, is the differentiation between technical cookies and profiling cookies established by applicable laws regarding cookies. Technical cookies primarily ensure the appropriate functioning of a website and are also referred to as "strictly necessary". In contrast, profiling cookies are cookies that gather non-anonymized information that enables the tracking of a person's browsing activity even across various devices. They are not treated equally from a regulatory standpoint. According to the EU's GDPR, strictly necessary cookies can be stored without the user's explicit consent as long as a notice is displayed. In practice, the level of detail in the categories of cookies presented to users can vary. Table 1 shows cookies categorized by the most common purposes. With "strictly necessary" cookies, we refer to all cookies that are essential for the basic functioning of the website. Without these cookies, requests cannot be properly delivered. For instance, they enable mechanisms such as authentication and load balancing of requests. Functional cookies, on the other hand, involve other website functionalities that are not essential but can improve the user experience. For example, they enable the personalization of user interactions by remembering language preferences or location to customize content delivery.

Table 1. Cookies grouped by purpose.

Cookie Category	Description
Strictly necessary	Essential for basic website functionality, e.g., secure user authentication.
Functional	User customization, e.g., set language preferences.
Performance	Measure website traffic and performance, e.g., count visitors for optimizing server load.
Analytics	Track user behavior, e.g., analyze user interactions with a website to enhance user experience.
Targeting	Customize ads based on user habits, e.g., show ads related to the last searched product.
Unclassified	Other cookies that do not fall into any of the preceding categories.

The other main types of non-essential cookies are performance, analytics, and targeting cookies. Performance cookies anonymously collect user interaction data, aggregating it for performance analysis. They contribute to optimizing aspects such as website visit duration and loading speeds, thereby aiding in performance enhancement. Analytics cookies delve into user behavior, providing website owners with insights to comprehend and improve website interactions. Primarily first-party, these cookies focus on enhancing user experience by analyzing and optimizing user engagement with the website. Targeting cookies, predominantly third-party, are designed to deliver relevant advertisements based on user profiles. They follow users across different websites to enable targeted advertising. This classification also recognizes the existence of cookies that may defy these categories or have yet to be classified. Such cookies are often labeled as "Other" or "Unclassified" in consent banners.

2.3. Cookie Banners

To store types of cookiesther than strictly necessary, publishers must explicitly request users' consent for their use. In most cases, this explicit consent is implemented by displaying a cookie consent banner, informing users of the types of cookies that may be stored, and giving them the possibility to accept or reject them according to their preferences. It is well known that cookie banners are not always impartial [11]. Often, these banners incorporate design elements that subtly encourage users to accept all cookies through deceptive patterns [12]. These nudges are also known as dark patterns [13].

Deceptive patterns highlight the buttons that grant permission for all purposes specified by the publisher, making them more visible to users. This design may include prominently presenting an "accept all" option on the first layer of the cookie banner, which is the part that the user sees immediately. In contrast, denying permission for data processing is often less straightforward, as many cookie banners do not provide a "deny all" button, making the process more complex for users who wish to opt out. This discrepancy in design between accepting and denying cookies serves as a nudge towards easier acceptance of all cookies, potentially influencing user behavior towards less privacy-conscious decisions [14,15].

Cookie paywalls represent another notable behavior by companies, perceived as a subtle means to influence user decisions. This practice has been adopted by some publishers as a response to the limitations imposed by the GDPR. Cookie paywalls operate on the principle that users can access the publisher's content if they either provide permission for their personal data to be processed for online advertising purposes, which often includes profiling and targeting or pay a subscription fee to the publisher, often including an ad-free experience.

Even though there are no explicit patterns attempting to influence user behavior, the decision-making process could become tedious [16]. The banner must provide comprehensive information, but this requirement can lead to boring banners that conflict with the user's desire for quick access to the website's content. It has been calculated in [1] that a user visits an average of 2.49 new publishers per day. In a worst-case scenario, it would take approximately 79.13 min per day to make all possible decisions regarding permission to process data. Furthermore, opting out is much more complicated than opting in, and in particular, changing or revoking previously accepted settings is very difficult [17].

The time and effort required to read and understand the information on cookie banners, coupled with the need to make multiple decisions, can be burdensome for users. This may explain why even users who claim to be concerned about their privacy accept all tracking cookies at the same time—this is the so-called privacy paradox.

2.4. Privacy Paradox

The privacy paradox, as discussed in the literature, refers to the contradiction between users' expressed concerns regarding online privacy and their actual behavior, which often does not align with these concerns [18]. This phenomenon is observable in various contexts. For instance, Iacono et al. [19] highlight this trend in recent years, where there is a growing concern about cybercrime but a reduced inclination to take measures to protect against such risks. The cause of the privacy paradox is a subject of much debate in the academic literature. Most researchers identify decision biases, lack of experience, and the illusion of control as possible explanations for the privacy paradox. Users may have a distorted perception of their ability to control personal information online, leading to behavior inconsistent with their privacy concerns [5]. A recent empirical study also suggests that user behavior may vary depending on age and medical conditions [20].

On the other hand, some works argue that the privacy paradox is a myth created by flawed logic. They contend that people's everyday behavior is not an accurate indicator of their preferences because it is distorted by biases, heuristics, manipulations, and other factors [21]. This perspective suggests that regulations focused on how information is used, retained, and transferred could have a more significant impact on this issue. Additionally, in [22], it has been highlighted that consumers fundamentally care about online privacy and provide evidence of numerous actions they take to protect it. However, in some cases, achieving desired levels of privacy is prohibitive or even undesirable through individual actions, necessitating political intervention for privacy protection.

We think that the privacy paradox may be rooted in the way cookie banner interfaces are designed and the process required to express preferences. Possible reasons behind the privacy paradox in this case could include decision fatigue, information overload, and the influence of dark patterns. Under these assumptions, even users who are attentive to their online privacy may passively accept cookies.

A company adopting a standard like the Transparency and Consent Framework (TCF), which is GDPR-compliant, is not obliged to explicitly follow any cookie banner template. Currently, the TCF is one of the most widely used standards as it is built by the industry for the industry (https://iabeurope.eu/transparency-consent-framework/, accessed on 18 December 2023).

2.5. Transparency and Consent Framework

Privacy preference signals are digital representations of user choices for processing personal data online. These signals have evolved in response to changing data privacy concerns and regulations. Hils et al. [23] discern two waves of privacy preference signals. The initial wave featured early standardization attempts, exemplified by platforms like the Platform for Privacy Preferences (P3P) and initiatives like Do Not Track (DNT) and the Network Advertising Initiative's (NAI) opt-out standards. These were primarily browser-centric and aimed at establishing universal settings across web interactions. The subsequent wave, influenced by legal frameworks such as the California Consumer Privacy Act (CCPA) and the General Data Protection Regulation (GDPR), introduced signals like the Global Privacy Control (GPC) and the Transparency and Consent Framework (TCF). Unlike their predecessors, GPC and TCF are more industry-driven, with a specific focus on user consent within the realms of online advertising and data processing. The TCF, aligning with GDPR principles, facilitates the communication of users' consent choices to publishers and advertisers, marking a notable departure from the earlier challenges faced by the browser-centric signals of the first wave. TCF is currently the most widely adopted specification, and projections suggest that this trend will continue in the future [23]. The TCF includes several actors including Publishers, Vendors, Consent Management Platforms (CMPs), Global Vendor List (GVL), and the Interactive Advertising Bureau Europe (IAB Europe) itself (https://iabeurope.eu/iab-europe-transparency-consent-framework-policies/, accessed on 18 December 2023). Publishers, the digital property operators, bear the responsibility of presenting the framework's user interface to consumers and establishing legal bases, including user consent, for vendors processing personal data derived from user visits to their content. Vendors, in this context, refer to companies involved in digital advertising delivery or related online activities within a publisher's digital domain. These entities may access end-user devices or process the personal data of visiting users, with their roles under GDPR varying as either controllers, processors, or both, depending on specific scenarios.

Consent Management Platforms (CMPs) play a crucial role as mediators, centralizing and overseeing end-user transparency, consent, and objections. These platforms, which can be either privately operated by publishers for their purposes or commercially available for others, manage the legal basis statuses of vendors listed on the Global Vendor List (GVL). They facilitate the establishment of legal bases for processing, acquire necessary user consent, manage user objections, and communicate these statuses within the digital ecosystem.

Overseeing the TCF is the Interactive Advertising Bureau Europe (IAB Europe), which administers and governs the Framework, including its policies, specifications, and the GVL. IAB Europe's role is dynamic, involving periodic policy updates to ensure the Framework's continued efficacy. The Global Vendor List, maintained by IAB Europe, catalogs vendors registered to participate in the framework. This list is pivotal for CMPs, publishers, and individual vendors, its structure and content being defined by the framework's specifications. Collectively, these actors form the backbone of the TCF, ensuring compliance and facilitating a transparent digital advertising environment.

To understand what happens under the TCF, we describe a concrete scenario in which User X visits the website of Publisher A for the first time. This process can be abstractly described through the following sequence of actions.

1. User X arrives at Publisher A's website.
2. Publisher A contacts its Consent Management Platform (CMP), CMP A, using a CMP tag—a JavaScript tag added to the website.
3. CMP A's code runs on the page and checks if there is a Transparency Consent (TC) string in User X's local storage corresponding to Publisher A.
4. Since it is User X's first visit, no TC string is found. Then, CMP A displays a cookie banner.
5. User X makes consent choices for each purpose listed on the cookie banner.
6. CMP A creates a TC string for User X and Publisher A by encoding the consent information according to TCF standards.
7. Publisher A stores the TC string locally on User X's device.
8. Publisher A uses its CMP API to decode the TC string, understanding which purposes User X allows it to pursue.

Subsequently, when User X revisits Publisher A, steps 1–3 are repeated to check the existing consent preferences, but steps 4–7 will be skipped, proceeding directly to step 8.

Our proposal integrates into the Transparency and Consent Framework (TCF), taking the place of the user in the interaction with cookie banners, as illustrated in Figure 1. Once the user has defined their privacy policy within the UPPMS, the system intercepts the CMP request and subsequently responds on behalf of the user according to the established policy.

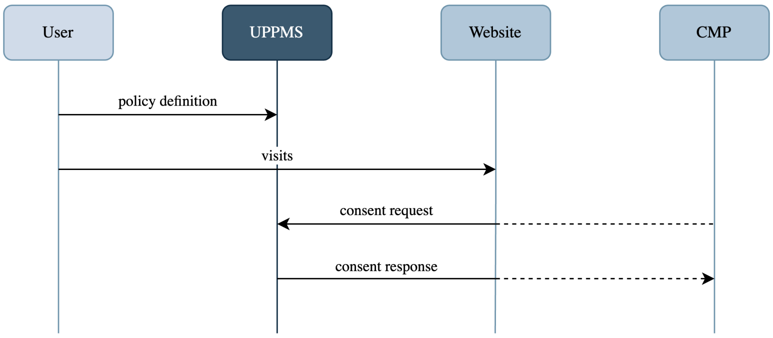


Figure 1. The diagram illustrates how the proposed UPPMS integrates and interacts within the Transparency and Consent Framework. After a user has defined his policy, the UPPMS intercepts consent requests from CMPs and responds in a manner consistent with the user's policy.

3. Related Work

Most of the systems presented in the literature that aim to assist users with privacy issues using automated techniques focus on privacy policies rather than privacy preference signals. In [24], automatic assessment of privacy policies using machine learning, natural language processing, and manual annotation is proposed. The use of these technologies for the same purpose is explored in [25]. Deficiencies of policies in terms of readability and ambiguity are evaluated in [26], while [27] investigates how language technologies can support users in better understanding these policies. In [28], machine learning-based strategies for automating GDPR compliance checks in data processing agreements are introduced.

The Platform for Privacy Preferences (P3P: https://www.w3.org/TR/P3P/, accessed on 18 December 2023) was the first attempt to standardize privacy preferences in a machine-readable format. The P3P facilitated the development of the first user agents for managing privacy preference signals. One example was Privacy Bird (http://www.privacybird.org/, accessed on 18 December 2023), a browser add-on that analyzes the privacy policies of websites and compares them to the user's personal preferences. Other more rigid approaches include the Mozilla Firefox add-on Targeted Advertising Cookie Opt-Out (TACO), which provided an updated opt-out cookie list but deleted all current cookies. However, the P3P protocol became obsolete because of limited adoption by AdTech vendors, making P3P-based solutions outdated. Additionally, privacy advocates criticized P3P for failing to impose consequences for false privacy practice reports (https://archive.epic.org/reports/prettypoorprivacy.html, accessed on 18 December 2023).

Following P3P's obsolescence, browsers like Mozilla Firefox incorporated alternative measures like Do Not Track (DNT). However, DNT faced similar challenges as P3P, particularly when major AdTech vendors chose to ignore DNT signals. This decision was influenced by Microsoft's default activation of DNT in Internet Explorer (https://www.iab.com/news/do-not-track-set-to-on-by-default-in-internet-explorer-10iab-response/, accessed on 18 December 2023). Over time, even pro-privacy browsers like Firefox ceased supporting DNT (https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/DNT, accessed on 18 December 2023). New browsers and add-ons, such as Brave (https://brave.com/web-standards-at-brave/4-global-privacy-control/, accessed on 18 December 2023) and Privacy Badger (https://www.eff.org/gpc-privacy-badger, accessed on 18 December 2023), have adopted the Global Privacy Control (GPC) signal, although skepticism remains about its widespread adoption by AdTech vendors [23].

The General Data Protection Regulation (GDPR) necessitated compliance strategies for websites, leading to the emergence of Consent Management Platforms (CMPs) and "Consent as a Service" solutions. Quantcast (https://www.quantcast.com/, accessed on 18 December 2023) and similar companies have facilitated TCF adoption, paralleling the impact of Let's Encrypt (https://letsencrypt.org/, accessed on 18 December 2023) on HTTPS adoption [23]. In this landscape, PIMSs have emerged, offering centralized management of personal data [1]. However, current PIMSs offer limited functionality, primarily focused on managing consent decisions. These tools, generally available as browser extensions, face challenges in addressing the variability of consent requirements across different websites.

Two of the most popular browser extensions are "I Don't Care About Cookies" (https://www.i-dont-care-about-cookies.eu/, accessed on 18 December 2023), which simplifies the cookie consent process by granting permission for all or only necessary cookies, and "Never-Consent" (https://www.ghostery.com/blog/never-consent-by-ghostery, accessed on 18 December 2023) by Ghostery, which removes cookie pop-ups and expresses disagreement with online tracking. The technical specification for Advanced Data Protection Control (ADPC: https://www.dataprotectioncontrol.org/, accessed on 18 December 2023) provides an alternative approach that permits users to establish general or specific consent indications, while authorizing publishers to solicit consent via TCF or specific requests. While ADPC is still in its developmental stages, a prototype is accessible for Firefox and Chromium-based browsers.

The history of privacy preference signals indicates that websites and AdTech vendors have been hesitant to adopt a unified signal. Indeed, it is expected that investment in TCF signals will continue [23]. Within this future perspective, our proposed User Privacy Policy Management System stands out as a valuable tool in supporting users. This system automatically handles consent requests using a user's policy within the TCF framework, reducing the time users spend declaring their preferences and minimizing their exposure to unfair practices by publishers.

4. User Privacy Policy Management System

A User Privacy Policy Management System (UPPMS) must provide users with three fundamental functionalities:

1. Offer the user a direct and understandable means to define their privacy policy;
2. Generate rules based on the user's preferences to implement the specified policy;
3. Automatically enforce the user's policy during navigation.

Figure 2 provides an abstract representation of the logical elements interacting with the UPPMS.

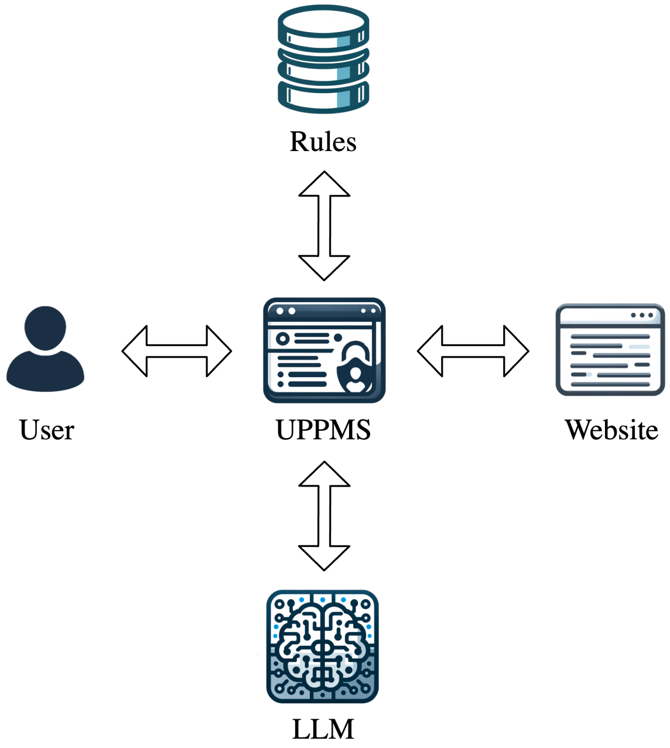


Figure 2. Logical elements interacting with the UPPMS.

The UPPMS generates various rules corresponding to different sequences of actions required to apply diverse types of policies on a specific website, utilizing an LLM. These rules are stored in a distributed database. Users declare their policies within the UPPMS. Subsequently, as the user visits each page, the UPPMS retrieves and applies the necessary rules to express the user's preferences.

Given a list of websites, there is a finite number of policies that may be applied to a particular website. Rules for implementing these policies could be generated at any time. If a user intends to apply a policy for which there are no generated rules, these rules could be generated lazily and added later to the distributed database accessible to all users. The list of policies that are not applicable is also stored in the distributed rule database. In the event that a user with a particular policy accesses a website where the policy cannot be applied (for example, there is no sequence of actions that enables policy application), the user is informed that the website does not support policy application.

The remainder of this section describes how users declare their policies, how rules are generated, and how rules are applied.

4.1. User Policy Definition

User privacy policies should be standardized by establishing a default setting system that reflects different levels of privacy and data usage. This will provide users with a simplified decision-making process. A list of user privacy policies based on the categories of cookies described above is presented in Table 2.

Table 2. List of user privacy policies.

	Cookie					
Policy ID	**Strictly Necessary**	**Analytics (A)**	**Functional (F)**	**Performance (P)**	**Targeting (T)**	**Unclassified**
REJECT ALL	✓					
ACCEPT A	✓	✓				
ACCEPT F	✓		✓			
ACCEPT P	✓			✓		
ACCEPT T	✓				✓	
ACCEPT A-F	✓	✓	✓			
ACCEPT A-P	✓	✓		✓		
ACCEPT A-T	✓	✓			✓	
ACCEPT F-P	✓		✓	✓		
ACCEPT F-T	✓		✓		✓	
ACCEPT P-T	✓			✓	✓	
ACCEPT A-F-P	✓	✓	✓	✓		
ACCEPT A-F-T	✓	✓	✓		✓	
ACCEPT A-P-T	✓	✓		✓	✓	
ACCEPT F-P-T	✓		✓	✓	✓	
ACCEPT ALL	✓	✓	✓	✓	✓	✓

The REJECT ALL policy is the most restrictive, permitting only strictly necessary cookies. Conversely, the ACCEPT ALL policy is the most permissive and allows for any type of cookie. Between these policies are varying combinations of acceptance regarding analytics, functional, performance, and targeting cookies. For example, the policy labeled ACCEPT P provides authorization solely for performance cookies, whereas the policy labeled ACCEPT F-P permits functional and performance cookies but denies analytics and advertising cookies.

A standardized schema for defining user privacy policies would allow users to easily select a privacy level that matches their personal preferences. However, not all users have the computer literacy to directly define their policies. To bring all user types to common ground, a step-by-step approach must be used to guide them through the process of defining privacy policies. Figure 3 illustrates a policy definition procedure that assists the user.

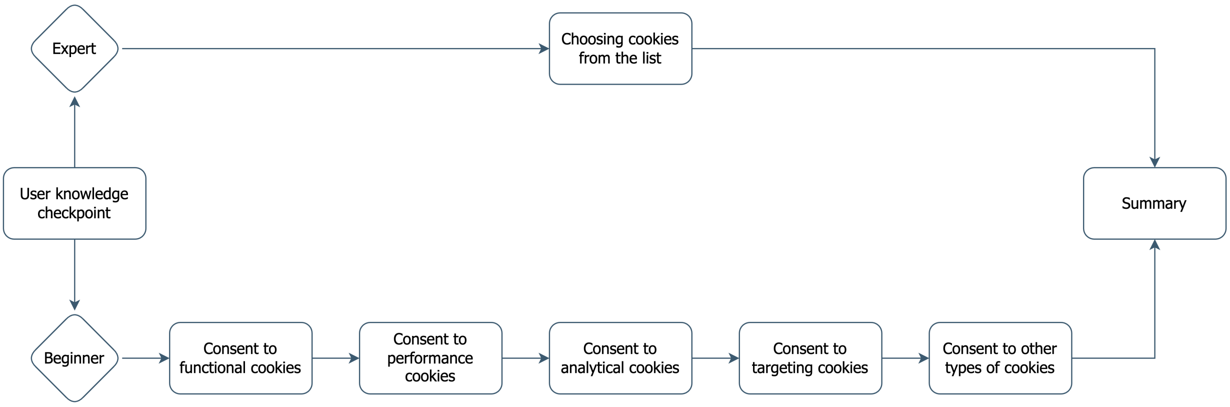

Expert
Choosing cookies from the list
User knowledge checkpoint
Summary
Beginner
Consent to functional cookies
Consent to performance cookies
Consent to analytical cookies
Consent to targeting cookies
Consent to other types of cookies

Figure 3. User-guided privacy policy definition process.

This process offers two options: a quick route for experienced users to select the policy directly and a step-by-step wizard to guide other users. See Appendix A for an example of how to present the wizard to users.

4.2. Rules Generation

To apply a specific user policy, it is necessary to execute a sequence of actions (the rules) that must be applied to a banner. The rule generation process is thus an iterative process aimed at identifying the sequence of actions necessary to close a cookie banner while expressing the user's preferences.

The process for obtaining the list of actions, as depicted in Figure 4, involves several steps. Initially, the source code of the targeted web page is extracted (step 1). This code is then input into an LLM to identify the selector for locating the cookie banner (step 2). Once the source code of the cookie banner is extracted (step 3), it is fed into another LLM, along with the user's policy. This specialized LLM interprets the text and the semantic meaning of the banner elements, determining the action required to progress toward the goal of applying the policy. It returns the selector for the element that needs to be clicked to execute the correct action (step 4). This selector is added to a list of actions that will be provided as output upon completion of the process (step 5). The web page is then revisited, executing all the actions (step 6). The state of the page after performing all the actions is captured, and the iterative process continues with the next iteration (step 7).

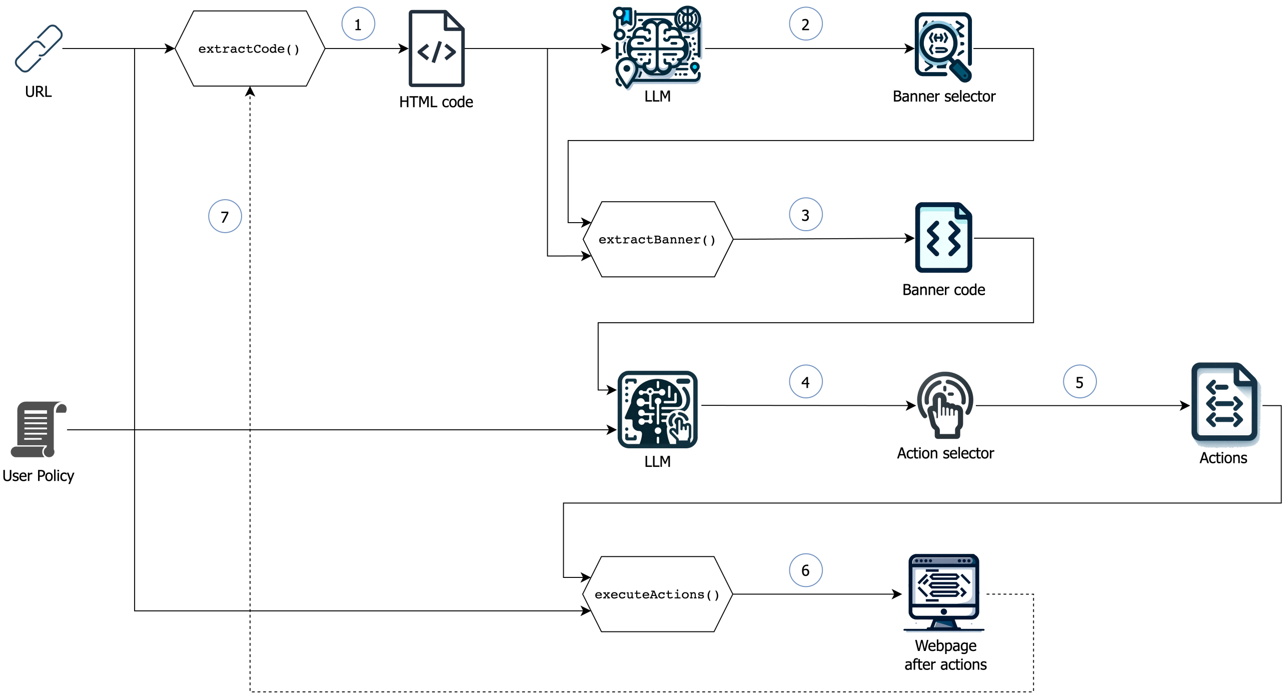


Figure 4. Rule generation process.

This iterative procedure allows for step-by-step navigation through various banner screens until the objective is achieved, if possible. In the case of success, the list of actions required to apply the specific policy to the particular website is returned. This information is stored in a distributed database and made available to users upon request.

The inability to achieve the goal primarily occurs in two cases, as identified by the LLMs. The first scenario involves the absence of the cookie banner. The second is related to the fact that the cookie banner does not allow the application of the user's policy. For example, there might be no buttons for making choices, or it could be a cookie paywall. In both cases, this information will be stored in the database, and the user will be notified of the scenario when visiting the specific website.

4.3. Enforcement of the Rules

The process of applying a policy is less complex than the process of generating one. The UPPMS, after guiding the user through generating their policy as previously described, will track the user's policy through an identifier called the policy ID. For instance, in Listing 1, an example is provided of how a user's policy might be represented, where the user intends to reject all cookies.

Listing 1. Example of a user policy rejecting all cookies.

```
{
  "version": "1.0.0",
  "user_policy_id": "REJECT_ALL"
}
```

When a user visits a web page, the UPPMS will conduct a direct search within the rules database using the website URL and the user's policy as keys. At this point, the sequence of rules to apply to enforce the user's policy will be retrieved.

Two particular cases need consideration during the policy application, concerning the inapplicability of the policy and the absence of a rule in the database. The first case occurs when the policy is identified in the database but is not applicable to the specific website due to the website's limitations. In this scenario, based on the user's preferences, the most restrictive policy, the most permissive policy, or no choice at all might be applied. It could be left to the user to interact with the banner.

In the second case, if the rule being attempted for application is no longer valid (e.g., due to changes in the website's banners or the presence of a cookie paywall) or if the rule has never been generated for that specific website, a notification will be shown to the user. The user will be informed that UPPMS cannot be used for a specific reason, and a request will be sent to a remote server responsible for generating the rule. Even though the current user may not benefit from UPPMS support in this instance, the generated rule will be made available for future users.

5. Experiments

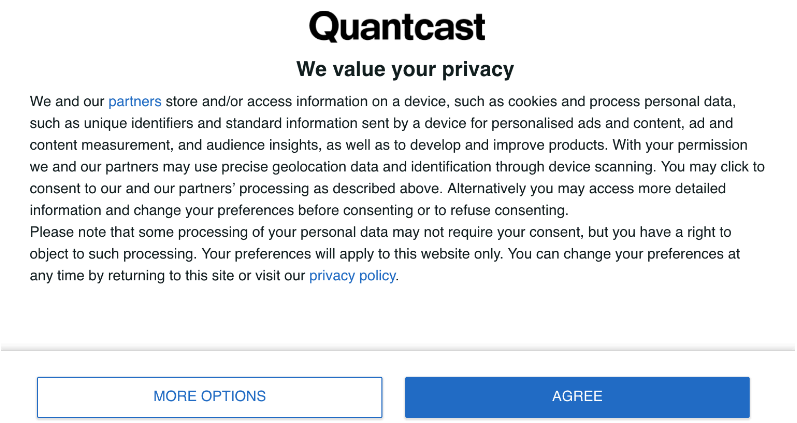
The main goals of the experimental phase were to observe the behavior of users during their first visit to a website that implements a cookie banner compliant with the TCF, and to measure the average time users take to implement a policy (i.e., a set of actions indicating a preference) on the cookie banner. The time taken by users to express their policy is then compared to the time taken by UPPMS.

The Quantcast website was chosen as the target website for this experiment for the following reasons:

- The cookie banner presents a detailed initial description, and users cannot proceed with navigation until they interact with the banner;
- On the cookie banner, there are only two buttons: "Agree" and "More options", with the "Agree" button emphasized, as shown in Figure 5. As a result, all other options, including "Reject all", are on another screen;
- This website is somewhat representative because Quantcast, as a company, provides a free product to other businesses called Quantcast Choice that enables the asking of consumers for consent regarding their data. At the time of writing, Quantcast Choice is one of the first CMPs on the market to comply with TCF v2.0, with over 25,000 clients in 25 countries (https://help.quantcast.com/hc/en-us/articles/13422322932379-Quantcast-Choice, accessed on 18 December 2023).

Users were asked to identify which products were available on the Quantcast website. They were given no information about the experiment, with the justification that this was a preliminary phase. In reality, the purpose of this request was to focus their attention on a specific task and observe their initial interaction with the banner upon entering an unknown website.

To ascertain if users read the banner before taking action, it is assumed that users can read at a rate of 238 words per minute in silent reading. Given that the Quantcast website's initial screen displays 152 words, it is considered that users did not read it if they performed the first action in less than 38 s. Once the users closed the banner, the purpose of the experiment was clarified, and the second phase of measuring the timing of expressing specific choices was initiated.



Quantcast

We value your privacy

We and our partners store and/or access information on a device, such as cookies and process personal data, such as unique identifiers and standard information sent by a device for personalised ads and content, ad and content measurement, and audience insights, as well as to develop and improve products. With your permission we and our partners may use precise geolocation data and identification through device scanning. You may click to consent to our and our partners' processing as described above. Alternatively you may access more detailed information and change your preferences before consenting or to refuse consenting.
Please note that some processing of your personal data may not require your consent, but you have a right to object to such processing. Your preferences will apply to this website only. You can change your preferences at any time by returning to this site or visit our privacy policy.

Figure 5. Cookie banner as it appears upon the first visit to the Quantcast website.

Users were then asked to apply the previously defined ACCEPT ALL, REJECT ALL, and ACCEPT P policies to the Quantcast website banner. In this specific case, ACCEPT P is equivalent to accepting cookies only for measuring content performance.

We evaluated the time required to implement the same policies using the proposed UPPMS. The UPPMS was developed as a browser extension that enforces the policy during a visit to a website. To assess UPPMS performance, we monitored the execution time of the web extension code and computed statistics. Page visits were conducted over a broadband internet connection, with a delay time of 100 ms between actions.

For generating rules outlined in Listing 2, we employed two GPTs (https://openai.com/blog/introducing-gpts, accessed on 18 December 2023)—custom versions of ChatGPT 4.0 designed to incorporate instructions, additional knowledge, and various skills. The first model specialized in identifying cookie banners, and locating the container ID that houses the banner on a given web page. The second model focused on interpreting banners to determine the subsequent rule for policy enforcement. Refer to Appendix B for the instructions used during model fine-tuning.

Listing 2. JSON file with the rules used in the experiments on the Quantcast website.

```
{
    "www.quantcast.com": {
        "ACCEPT_P": [
            ".css-1hy2vtq",
            "#Purposes-id\\:8 > button",
            "#Purposes-id\\:8 .qc-cmp2-toggle",
            ".qc-cmp2-footer .css-47sehv"
        ],
        "ACCEPT_ALL": [
            ".css-47sehv"
        ],
        "REJECT_ALL": [
            ".css-1hy2vtq",
            ".qc-cmp2-header-links > button.css-8rroe4:first-child",
            ".css-47sehv"
        ]
    }
}
```

6. Results

The experimentation phase involved a sample of 16 individuals aged between 23 and 58 years. All participants were experienced with computers, as they use them on a daily basis for work. Among the participants, 73% had a bachelor's or master's degree, while the remaining percentage had a lower degree.

The experiments were conducted by navigating a website using a laptop browser and a broadband connection. This section presents the results of the analysis of user interaction with the Quantcast cookie banner and a comparison with the application times of the policies implemented by UPPMS.

6.1. Users' Behavior

Figure 6 depicts the flow of user behavior during the experiment and their self-reported behavior in general. The first part distinguishes whether users read the initial information on the cookie banner and their first action on the banner. The subsequent section outlines the actions they reported taking and their privacy concerns.

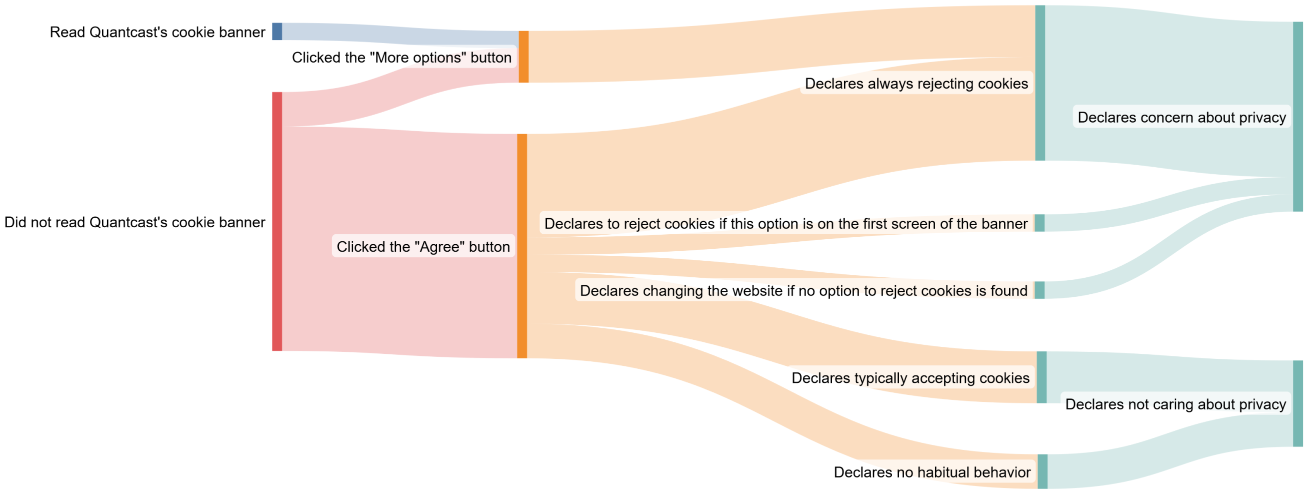


Figure 6. User behavior during the experiment with their self-reported behavior.

Some 94% of the participants took their first action on the banner in less than 4 s. Only one participant read the banner carefully, taking more than 38 s, before making the initial choice on the banner. In the observed behavior of participants regarding their first interaction with the cookie banner, it was found that the majority (81%) initially clicked the 'Agree' button. However, this action does not necessarily reflect their usual preferences. Only 23% of these individuals report regularly accepting cookies. Some 13% said that they have no fixed habits, but choose one of the fastest alternatives to remove the cookie banner, while a minority said that they would only reject cookies if the reject button was immediately visible. Another minority stated that they would leave the website if they could not locate the reject button. The majority of users who clicked the "Agree" button (46%) said that during the experiment they were focused on completing the task and therefore tried to close the banner as quickly as possible. Only 27% of the participants who said that they always refused all cookies really did so consistently over the course of the experiment.

6.2. Time Required to Apply Policies

The boxplot in Figure 7 illustrates the average time taken to perform the required actions for the three different policies under investigation in the experiment. The execution of the actions required for the "ACCEPT ALL" policy generally demonstrates quicker and less variable times. In contrast, the times required to perform actions for the "REJECT ALL" and "ACCEPT P" policies exhibit greater variability and higher average durations. Particularly, "ACCEPT P" highlights a case with a significantly longer duration than the others.

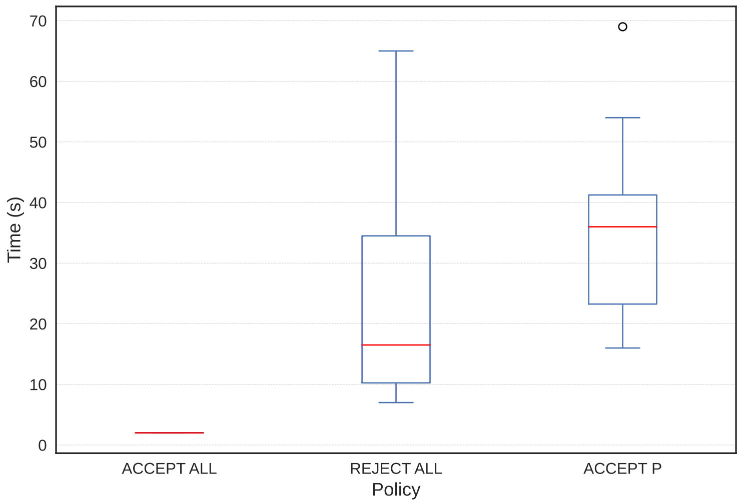


Figure 7. Time taken by users to execute policies.

Figure 8 shows the average time required to execute three policies, arranged in ascending order of complexity. Policy complexity is determined by the number of clicks needed from the user and the concentration required to understand the meaning of the clauses. On average, users spend several seconds performing a click, and this time increases with the number of required actions. In contrast, the UPPMS completes tasks within milliseconds. In detail, the UPPMS averages 0.011 s for the "ACCEPT ALL" policy, 0.203 s for the "REJECT ALL" policy, and 0.302 s for the "ACCEPT P" policy.

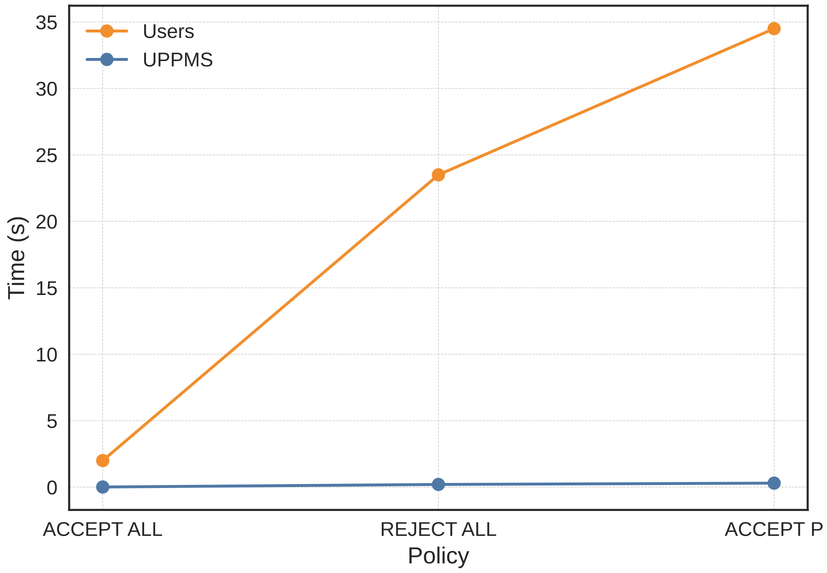


Figure 8. Comparison of average time taken by users and UPPMS.

7. Discussion

The results from the experiment show that half of the participants behaved inconsistently with their stated usual practices. The main reason given was that the need to complete a specific task led them to opt for the quickest way to remove the banner, which in this case was the "Agree" button that allowed for the acceptance of all cookies. This tendency suggests a predominance of operational efficiency over individual privacy concerns.

Most users did not read the privacy policy on the initial page of the banner despite valuing their own privacy. This can be attributed to the information overload present in these banners. For instance, it was estimated that reading all the information on the first page of Quantcast's cookie banner would take about 38 s.

Another factor contributing to this discrepancy between preference and actual choice is decision fatigue. Users find themselves overwhelmed by the abundance of choices, which pushes them towards the option perceived as less demanding.

Generally, we observed that users tended to adopt a binary strategy: they either accepted all cookies or rejected all. This behavior appears to be influenced by the design and usability of the banner itself, as often these inline options allow for quicker decision-making [29].

In the Quantcast cookie banner, the "Reject All" button was located in a secondary interface. Some participants reported that it was counterintuitive to use this "Reject All" button, as they expected the banner to close automatically after selecting it. However, an additional click on the "Save and Exit" button was required. This particular difficulty was not experienced by users who typically always reject all cookies. These latter users were also generally faster at implementing the ACCEPT P policy.

Our findings suggest that decision fatigue and information overload may contribute to the privacy paradox in user interaction with cookie banners. This paradox refers to users choosing the quickest option to close banners, usually by accepting all cookies, despite potentially having personal preferences for greater privacy. Although our experimental results align with the findings in the literature, our sample has limitations that prevent it from being representative of all user types. Our sample consisted entirely of users with demonstrated computer experience, which may have also provided an optimistic estimate of the time required for users to implement the policies. It is important to note that a user with limited computer experience may face additional challenges when carrying out the actions required by a policy.

The interview method may also have contributed to the optimistic estimates. This is because users were asked to implement the REJECT ALL policy first and then the ACCEPT P policy. Some users went through all the clauses thoroughly to make sure they had rejected everything, thus familiarizing themselves with the banner. This behavior may have given them a temporary advantage when performing actions for the following policy. However, the average time to execute the ACCEPT P policy was still the longest overall. The adoption of a tool like UPPMS eliminates the need for users to interact directly with cookie banners. This avoids the complexity and deceptive patterns commonly found in cookie banners and has the potential to address the privacy paradox in this context.

We highlight, however, that the decision-making process associated with cookie banners can be burdensome, even without deceptive patterns. The introduction of structured processes for standardizing cookie policies can increase user awareness and automate interactions with consent requests. In this regard, we have introduced standardized policies and a process for acquiring user preferences that feature a few categories with descriptions comprehensible to the average user.

8. Conclusions

In the absence of clear cookie banner implementation guidelines, users go through a time-consuming and intricate preference expression process that ultimately directs them toward accepting cookies. Our experiment highlighted the privacy paradox in which most users, due to decision making fatigue and information overload, choose the quickest option to dismiss cookie banners (the "Reject All" button), contradicting their privacy preferences. The proposed system guides users in formulating their cookie policy and generates rules to enforce user preferences on all visited websites. As a result, users can enjoy a smooth browsing experience as intricate banners with deceptive patterns disappear automatically in less than a second.

The UPPMS is currently a simple browser extension. However, in the future, it is intended to become a cross-platform service usable on both mobile and desktop applications. The UPPMS can suggest products and services that align with the user's privacy needs. For example, if a website's terms do not align with user privacy concerns, the UPPMS may suggest alternative websites that better meet their privacy needs. A tailored version of the UPPMS could also be used in other contexts, such as user experience, to identify deceptive patterns within interfaces or assess the level of difficulty for users in expressing their preferences.

Author Contributions: Conceptualization, L.P. and M.M.; methodology, L.P.; software, L.P.; validation, L.P.; formal analysis, L.P.; investigation, L.P.; writing—original draft preparation, L.P.; writing—review and editing, L.P., M.M., M.F. and F.P.; supervision, M.M., M.F. and F.P.; project administration, F.P.; funding acquisition, F.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by project SERICS (PE00000014) under the NRRP MUR program funded by the EU-NGEU.

Data Availability Statement: Data supporting reported results can be found at the following link: https://bit.ly/48yh86N.

Acknowledgments: This paper is part of the research activity carried out within the project PON "Ricerca e Innovazione" 2014–2020, action IV.6 "Contratti di ricerca su tematiche Green", issued by the Italian Ministry of University and Research.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Policy Definition Wizard

Below are steps to define your own privacy policy regarding the use of cookies. There are two distinct streams of users: one for experienced individuals (E) and one for beginners (B).

- **START: User Knowledge Checkpoint**
  - **Question:** Are you aware of the types of cookies and how they are used?
  - **Answers:**
    - **R1:** Yes, I know about cookies and how they are used.
    - **R2:** No, I need more information about cookies.
  - **Next step:** If the answer is R1, go to Step E1. If the answer is R2, go to Step B1.
- **Step E1: Choosing cookies from the list**
  - **Question:** Check the cookies you wish to accept:
  - **Answers:**

    ☒ Strictly necessary *(always allowed)*
    ☐ Analytics
    ☐ Functional
    ☐ Performance
    ☐ Targeting
    ☐ Other types of cookies
  - **Next step:** Go to END.
- **Step B1: Consent to functional cookies**
  - **Question:** Do you wish websites to use cookies to store your preferences, such as your selected language or your location for weather forecasts?
  - **Answers:**
    - **R1:** Yes, I allow to store functional cookies.
    - **R2:** No, I do not want to store functional cookies.
  - **Next step:** Go to Step B2.
- **Step B2: Consent to performance cookies**
  - **Question:** Do you wish websites to collect aggregated statistics about navigation, such as counting the pages you visit and the page load times?
  - **Answers:**

        - **R1:** Yes, I allow to store performance cookies.
        - **R2:** No, I do not want to store performance cookies.
    - **Next step:** Go to Step B3.
- **Step B3: Consent to analytical cookies**
    - **Question:** Do you wish websites to store your interactions with the website, such as which pages you visit, how long you stay, and what actions you perform on the website?
    - **Answers:**
        - **R1:** Yes, I allow to store analytical cookies.
        - **R2:** No, I do not want to store analytical cookies.
    - **Next step:** Go to Step B4.
- **Step B4: Consent to targeting cookies**
    - **Question:** Do you wish websites to display personalized advertisements based on your interests? For example, if you frequently visit travel sites, you might see ads related to travel offers on other websites.
    - **Answers:**
        - **R1:** Yes, I allow to store targeting cookies.
        - **R2:** No, I do not want to store targeting cookies.
    - **Next step:** Go to Step B5.
- **Step B5: Consent to other types of cookies**
    - **Question:** Do you wish websites to store other types of cookies that do not fall into any of the above categories and may be used for further unclassified purposes?
    - **Answers:**
        - **R1:** Yes, I allow to store other types of cookies.
        - **R2:** No, I do not want to store other types of cookies.
    - **Next step:** Go to END.
- **END: Summary**
  Summary of the chosen policy. For example, if the user wishes to store only functional and performance cookies, the following summary will be displayed.
    - **Policy ID**: ACCEPT F-P
    - **Description:**
      You agree to the storage of
        - (F): Functional cookies
        - (P): Performance cookies

      This means that websites can use cookies to store your preferences and collect aggregated statistics on navigation.

Appendix B. GPTs Fine-Tuning

Below are the instructions provided to the custom LLM models. Both models do not use additional capabilities such as Web Browsing, DALL-E Image Generation, and Code Interpreter.

Appendix B.1. Cookie Banner Identification

- **Task Description:** Your task involves analyzing HTML code for a cookie banner and scripting the necessary JavaScript actions to accomplish one of the following goals. Each goal is identified by a keyword followed by a description:
    - **ACCEPT_ALL:** Determine the JavaScript script sequence to click buttons for accepting all cookies and closing the banner.

  - **REJECT_ALL:** Determine the JavaScript script sequence to click buttons for rejecting all cookies and closing the banner.
  - **ACCEPT_P:** Determine the JavaScript script sequence to click buttons for accepting only performance cookies and reject all the rest. Performance cookies provide quantitative measures of website visitors. Information collected through these cookies is used to measure KPIs of the website or software, such as performance. For example, these cookies count visits and traffic sources.

- **Input Format:**
  - A keyword indicating the goal (ACCEPT_ALL, REJECT_ALL, ACCEPT_P).
  - The HTML code of the cookie banner.
- **Output Format:**
  - A JSON object indicating the task status and the sequence of JavaScript actions.
- **Detailed Instructions:**
  - Analyze the HTML code for interactive elements (es., buttons, toggles, checkboxes, etc. . .) relevant to cookie settings.
  - Provide concise responses depending on the task status:
    * If actions are partially identified, return a JSON object with “ONGOING” status and the actions taken so far, with additional required actions listed. Example:

      ```
      {
          "status": "ONGOING",
          "actions": [JavaScript actions list]
      }
      ```

    * If the goal is achieved, return a JSON object with “COMPLETE” status and the list of actions. Example:

      ```
      {
          "status": "COMPLETE",
          "actions": [JavaScript actions list]
      }
      ```

  - After an action that changes the banner (like expanding a section), request updated HTML code from the user.
  - Examine cookie labels in the HTML code to accurately identify the type of cookies.
  - Identify cookie categories based on their semantic description in the banner.
  - Before suggesting a JavaScript action, ensure all relevant sections are expanded and have actionable toggles.
  - Consider the need for confirmation buttons like “Save Preferences”, ”Save and Exit“, ”Confirm Settings”, etc. . . to finalize settings.
  - Continuously request updated HTML code from the user after actions that may change the banner’s content.
  - Focus on sections relevant to the specified cookie type. Expand only these sections.
  - In-depth examination of HTML code details (such as ‘id’, ‘class’, ‘aria-label’) is crucial for understanding each interactive element.
  - Check for classes like “expanded” in the HTML code to verify if a section is expanded.
  - Replace every \ with \\ for escaping purposes.
- **Additional Notes:**
  - Understand keywords based on their descriptions, not by their names.
  - Each step may require scripting multiple buttons. Analyze all potential buttons and their functions for JavaScript scripting.
  - Evaluate all user-actionable elements like toggles and checkboxes. Their descriptions might be in adjacent elements and could be crucial for making choices.

- Request updated HTML code from the user after each action that changes the banner′s content.
- Identify and expand only the sections relevant to the desired cookie type.
- Cookie banners may have multiple screens, which you need to navigate through to complete the task.

Appendix B.2. Cookie Banner Interpretation

- **Task Description:** Your objective is to meticulously analyze the provided HTML document to derive a unique selector for extracting the HTML code of the cookie banner. Focus on recognizing typical phrases and buttons characteristic of cookie banners.
  - Key Elements in Cookie Banners:
    - **Characteristic Phrases**: Look for phrases like "We use cookies to improve your experience.", "This website uses cookies to provide you with a great user experience.", "By using our website, you accept our use of cookies.", "Click accept to give your consent to accept cookies.", "We use cookies for analytics, personalization, and ads.", "Manage your cookie preferences.", "Adjust your cookie settings.", "Do you agree to our use of cookies?", "Manage cookies" or "Cookie preferences.", "Accept all cookies" or "Reject all cookies.", "To learn more about cookies and how we use them, view our cookie policy.", "This website stores cookies on your device.", "Our website uses cookies to remember your preferences.", "Cookies are used to track website usage and browsing habits.", "Your privacy is important to us, learn more about how we use cookies.", "By continuing to browse the website, you are agreeing to our use of cookies.", and "By continuing to browse the website, you agree to our use of cookies", and similar expressions.
    - **Common Button Texts**: "Accept", "Agree", "Consent", "I Agree", "OK", "Yes", "Allow Cookies", "Accept all", "Continue", "Got It", "Decline", "No Thanks", "Disagree", "Reject", "Reject All", "Opt Out", "No Cookies", "Refuse Cookies", "Settings", "Preferences", "Manage Cookies", "Cookie Settings", "Adjust Settings", "Customize", "Choose Cookies", "More Info", "Close", "X", "Later", "Not Now", "Learn More", "More Info", "Read More", "Policy", "Privacy Policy", etc.
- **Input Format:**
  - A comprehensive HTML document containing elements such as HTML, JavaScript, and CSS, inclusive of a cookie banner.
- **Output Format:**
  - A JSON file containing a selector for use with the JavaScript function document.querySelector() to obtain the complete HTML structure of the identified cookie banner, from its highest-level container.
- **Detailed Instructions**
  - Conduct a detailed review of the HTML document to pinpoint the cookie banner, using the provided key phrases and button texts as indicators.
  - If the cookie banner is not immediately visible, consider translating the page content into English to help identify banner-related keywords.
  - Focus on both the structural layout and the semantic content within the HTML to ensure accurate identification of the cookie banner.
  - Provide concise responses:
    - If no cookie banner is found, return a JSON with the status COOKIE_BANNER_NOT_FOUND. For example,

```
{
    "status": "COOKIE_BANNER_NOT_FOUND",
    "cookie_banner_selector": null
}
```

    * If a cookie banner is found, return a JSON with the status COOKIE_BANNER_FOUND and the selector code of the banner. For example,

```
{
    "status": "COOKIE_BANNER_FOUND",
    "cookie_banner_selector": <selector code>
}
```

- **Additional Notes:**
  - The task requires a careful examination of HTML elements, text content, and interactive features (buttons) to locate and retrieve the cookie banner code.
  - Ensure the extracted cookie banner code matches the original cookie banner code as loaded on the page.
  - Streamline the output to focus on essential information, thereby enhancing operation efficiency.

References

1. Skiera, B.; Miller, K.; Jin, Y.; Kraft, L.; Laub, R.; Schmitt, J. The Impact of the General Data Protection Regulation (GDPR) on the Online Advertising Market. Available online: https://www.gdpr-impact.com/ (accessed on 18 December 2023).
2. Degeling, M.; Utz, C.; Lentzsch, C.; Hosseini, H.; Schaub, F.; Holz, T. We value your privacy... now take some cookies: Measuring the GDPR's impact on web privacy. *arXiv* **2018**, arXiv:1808.05096.
3. Athey, S.; Catalini, C.; Tucker, C. The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. Available online: http://www.nber.org/papers/w23488 (accessed on 18 December 2023).
4. Aguirre, E.; Mahr, D.; Grewal, D.; De Ruyter, K.; Wetzels, M. Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *J. Retail.* **2015**, *91*, 34–49. [CrossRef]
5. Gerber, N.; Gerber, P.; Volkamer, M. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Comput. Secur.* **2018**, *77*, 226–261. [CrossRef]
6. Porcelli, L.; Ficco, M.; Palmieri, F. Mitigating User Exposure to Dark Patterns in Cookie Banners Through Automated Consent. In Proceedings of the International Conference on Computational Science and Its Applications, Athens, Greece, 3–6 July 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 145–159. [CrossRef]
7. Zhao, W.X.; Zhou, K.; Li, J.; Tang, T.; Wang, X.; Hou, Y.; Min, Y.; Zhang, B.; Zhang, J.; Dong, Z.; et al. A survey of large language models. *arXiv* **2023**, arXiv:2303.18223.
8. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762.
9. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving Language Understanding by Generative Pre-Training. Available online: https://openai.com/blog/language-unsupervised/ (accessed on 18 December 2023).
10. Gur, I.; Nachum, O.; Miao, Y.; Safdari, M.; Huang, A.; Chowdhery, A.; Narang, S.; Fiedel, N.; Faust, A. Understanding html with large language models. *arXiv* **2022**, arXiv:2210.03945.
11. Matte, C.; Bielova, N.; Santos, C. Do Cookie Banners Respect my Choice?: Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. In Proceedings of the 2020 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 18–21 May 2020; pp. 791–809. [CrossRef]
12. Utz, C.; Degeling, M.; Fahl, S.; Schaub, F.; Holz, T. (Un)informed consent: Studying GDPR consent notices in the field. In Proceedings of the ACM Conference On Computer And Communications Security, London, UK, 11–15 November 2019; pp. 973–990.
13. Hils, M.; Woods, D.W.; Böhme, R. Measuring the Emergence of Consent Management on the Web. In Proceedings of the ACM Internet Measurement Conference, Virtual Event, USA, 27–29 October 2020; pp. 317–332. [CrossRef]
14. Nouwens, M.; Liccardi, I.; Veale, M.; Karger, D.; Kagal, L. Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13. [CrossRef]
15. Thaler, R.H. Nudge, not sludge. *Science* **2018**, *361*, 431. [CrossRef] [PubMed]
16. Machuletz, D.; Böhme, R. Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR. *Proc. Priv. Enhancing Technol.* **2020**, *2*, 481–498. [CrossRef]

17. Mehrnezhad, M.; Coopamootoo, K.; Toreini, E. How Can and Would People Protect From Online Tracking? *Proc. Priv. Enhancing Technol.* **2022**, *1*, 105–125. [CrossRef]
18. Øverby, H. The Privacy Paradox. In *Encyclopedia of Cryptography, Security and Privacy*; Jajodia, S., Samarati, P., Yung, M., Eds.; Springer: Berlin/ Heidelberg, Germany, 2019; pp. 1–2. [CrossRef]
19. Iacono, M.; Mastroianni, M. Evaluating the Effectiveness of Privacy and Security Promotion Strategies. In Proceedings of the International Conference on Computational Science and Its Applications, Athens, Greece, 3–6 July 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 134–148. [CrossRef]
20. Fernandes, T.; Costa, M. Privacy concerns with COVID-19 tracking apps: A privacy calculus approach. *J. Consum. Mark.* **2023**, *40*, 181–192. [CrossRef]
21. Solove, D.J. The myth of the privacy paradox. *Geo. Wash. L. Rev.* **2021**, *89*, 1. [CrossRef]
22. Acquisti, A.; Brandimarte, L.; Loewenstein, G. Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *J. Consum. Psychol.* **2020**, *30*, 736–758. [CrossRef]
23. Hils, M.; Woods, D.W.; Böhme, R. Privacy Preference Signals: Past, Present and Future. *Proc. Priv. Enhancing Technol.* **2021**, *4*, 249–269. [CrossRef]
24. Sánchez, D.; Viejo, A.; Batet, M. Automatic assessment of privacy policies under the GDPR. *Appl. Sci.* **2021**, *11*, 1762. [CrossRef]
25. Zaeem, R.N.; German, R.L.; Barber, K.S. Privacycheck: Automatic summarization of privacy policies using data mining. *Acm Trans. Internet Technol. (Toit)* **2018**, *18*, 1–18. [CrossRef]
26. Belcheva, V.; Ermakova, T.; Fabian, B. Understanding Website Privacy Policies—A Longitudinal Analysis Using Natural Language Processing. *Information* **2023**, *14*, 622. [CrossRef]
27. Ravichander, A.; Black, A.; Norton, T.; Wilson, S.; Sadeh, N. Breaking down walls of text: How can nlp benefit consumer privacy? In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Online, 1–6 August 2021; ACL: Kerrville, TX, USA, 2021. [CrossRef]
28. Amaral, O.; Abualhaija, S.; Briand, L. ML-Based Compliance Verification of Data Processing Agreements against GDPR. In Proceedings of the 2023 IEEE 31st International Requirements Engineering Conference (RE), Hannover, Germany, 4–8 September 2023; pp. 53–64. [CrossRef]
29. Habib, H.; Li, M.; Young, E.; Cranor, L. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans LA USA, 29 April–5 May 2022; pp. 1–27.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.