# Framework for Ethical Acquisition of User-Data to Improve Recommendation Models' Accuracy in Digital Systems

Shaheer Hussain Qazi[1], M. Batumalay[2], Asheer Hussain[3], Ali Abbas[4]
Faculty of Data Science and Information Technology, INTI International University, Nilai, Malaysia[1, 2]
Faculty of Engineering and Technology, HITEC University, Taxila, Pakistan[3, 4]

***Abstract*—The modern digital ecosystem has evolved into a pervasive, opaque system where platforms collect and infer personal data through nearly every online action, search queries, email content, browsing history, and app usage, without transparency. Justified as a means to deliver "relevant" content and ads, this approach undermines user privacy, introduces bias, and normalizes surveillance. Through a comprehensive literature review, this study sought to: critically analyze the current landscape of user tracking, profiling, and privacy violations in online platforms and to evaluate the impact of existing legal, technical, and platform-driven mechanisms like GDPR, CCPA, ATT, and Privacy Sandbox in protecting user autonomy. It was learned that the current frameworks fall short by mostly being policy-based and having hard-to-access user controls. And a major flaw in existing systems is the assumption that all digital behavior reflects actual user preference, overlooking shared devices, accidental clicks, and non-user actions. To validate these insights, a survey of 572 privacy-aware participants was conducted, where nearly 71% preferring a proactive solution over passive regulatory frameworks and hard-to-navigate privacy menus/dashboards. Building on these findings, this study proposes a framework: a digital platform where individuals actively create and manage modular preference profiles, categorized by app type or content domain, which can be selectively and consensually shared with platforms in a standardized format. This concept facilitates high-quality, context-rich datasets for algorithms, enhancing personalization and recommendation models' accuracy and performance. By shifting from forced surveillance to invited participation, this approach advances ethical data-sourcing, enhances algorithmic accuracy, and aligns with SDG 9 and SDG 16 by prioritizing responsible digital solutions, process innovation, and safeguarding user autonomy.**

***Keywords*—Data handling; data privacy; online tracking; data collection; user profiling; model accuracy; ethical AdTech; open standards; user autonomy; SDG 9; SDG 16; process innovation***

## I. Introduction

Today's online environment is mostly about surveillance advertising. You can wish for something, watch a video, or just get news somewhere on the internet, and in either scenario, there is a greater likelihood that you will have your activity monitored, not only on that website, but on the entire internet environment. This mass cyber surveillance is traditionally described as a means to provide users with "more relevant Content". This justification is no longer accepted by privacy experts, Data Protection Regulators, and increasingly by everyday users.

Regulations like the EU's General Data Protection Regulation (GDPR), America's California Consumer Privacy Act (CCPA), Apple's App Tracking Transparency (ATT), and Google's Privacy Sandbox have taken steps to curb invasive data collection. While such efforts have made some progress on the transparency and control fronts, they largely shy away from ending the underlying machinery of surveillance capitalism. Instead, they've merely shifted how tracking is accomplished rather than whether it is accomplished. For example, ATT prevents third-party tracking on Apple devices but continues to allow Apple to build behavioral profiles in-house. In a similar manner, Privacy Sandbox prevents cross-site tracking while still allowing ad targeting systems based on aggregated user information.

The contemporary online advertising economy is based on users consenting, at least implicitly, to being monitored by virtue of accessing digital services. This consent is typically buried in unreadable terms of use or covered by manipulative interface design (a "dark pattern"). The result is that users are progressively profiled on the basis of search history, device use, app use, online and offline interpersonal interaction, geolocation, and even email content with little or no awareness or control. User's personal data is no more personal due to this excessive nonconsensual data collection.

These have effects greater than nuisance. Incorrect or misattributed profiling will warp the user experience, e.g., a Beach Holiday ad having inappropriate images displayed in a game application that is being played on a kid's device, or a shared device's search history leading to irrelevant or even inappropriate ad suggestions. Moreover, users' interests are usually deduced from noisy data, misplaced clicks, one-time browse for an alternative, or passing interest, thus constructing a tilted digital profile of the user different from their actual interest.

Despite data protection laws and consent mechanisms, most tracking is technically sophisticated and legally evasive. Companies get to play by the rule book without playing by its spirit. Fingerprinting, real-time bidding (RTB), and server-side tracking tend to get round even very stringent consent mechanisms. Law is either too slow to keep up, or too vague to effectively implement. And most importantly, they rarely ever give users a real choice: opting out typically means denying a service or receiving generic content and advertising. This is an extremely worrying scenario, not only in terms of privacy, but

also on a system design level. It is based on nonconsensual data inference and on imbalanced power dynamics between users and platforms.

It is important to clarify the scope and intent of this work. This study does not propose a new recommendation algorithm, personalization pipeline, or machine learning architecture, nor does it seek to benchmark predictive performance against existing models. Instead, the contribution lies at an earlier and often overlooked stage of the personalization lifecycle: the ethical and technical acquisition of preference data itself. While existing consent dashboards, preference managers, and data portability tools focus primarily on regulatory compliance or post-hoc control, it continues to rely on behaviorally inferred, noisy, and platform-owned data. In contrast, the framework proposed here introduces a user-authored, platform-agnostic preference declaration mechanism designed to replace inference with explicit intent. The central claim is not algorithmic novelty, but that recommendation and analytical models, regardless of their internal design, benefit structurally from cleaner, context-aware, and voluntarily provided accurate preference data.

This study pursues five core objectives:

- To critically analyze the current landscape of user tracking, profiling, and privacy violations in online platforms and advertising ecosystems.
- To evaluate the strengths and limitations of existing legal, technical, and platform-driven mechanisms like GDPR, CCPA, ATT, and Privacy Sandbox in protecting user autonomy through literature review.
- To conduct a survey to get real responses from people that have knowledge of the different Data Protection Acts, Online Data Privacy, and modern Advertising industry.
- To propose a user-centric conceptual framework for preference data sharing system that is consensual, transparent, secure, and technically feasible. This approach not only respects control by users, but it also has the potential to yield more accurate data and truly personalized relevance.
- To design a simple Technical Demo of the proposed solution. The solution is to be kept simple and easy to implement. So, it can be created by any organization or developer. Standardized downloadable format for easy integration.

Building on these objectives, this study contributes a conceptual and technical framework for ethical, user-governed preference data acquisition that addresses both privacy and system-performance limitations of contemporary personalization infrastructures. In place of being passively tracked for data with an effort to guess what they would like, users would explicitly state what their interests are in a structured and shareable form. For building a social profile, users may have multiple interest profiles, customized for specific contexts like shopping, social, video content, or games, and apply them selectively to apps or sites. By enabling individuals to explicitly construct, manage, and selectively share structured preference profiles, the framework reduces reliance on inferred behavioral signals that are frequently noisy, contextually misleading, or generated without meaningful consent. The proposed approach has implications beyond privacy protection alone: it offers a practical pathway toward higher-quality datasets for recommendation and data analysis models, reduces preprocessing overhead associated with behavioral noise, and supports interoperability through standardized, platform-independent data formats. From a broader perspective, the framework advances responsible digital system design by aligning personalization mechanisms with user autonomy, informed consent, and ethical data sourcing principles.

The remainder of the study is organized as follows: Section II reviews existing literature on behavioral tracking, privacy regulation, and personalization systems, highlighting conceptual and technical gaps. Section III describes the research methodology and survey design used to gather real users' views. Section IV presents results and discussion of empirical findings. Section V presents the proposed user-governed preference framework, its design principles, and integration patterns. It also presents a simple website for demonstrating the framework's workings. Section VI discusses the challenges and limitations of such a framework. Section VII suggests implementation roadmap and recommendations for future work. Finally, Section VIII concludes the study.

## II. Literature Review

There has been considerable effort on how modern digital advertising platforms use machine learning and behavioral data for targeting, with inherent trade-offs to systemic inefficiency and with privacy. In [1], the authors built a twin-framework consisting of an auction mechanism and machine learning to examine in-app advertising on mobile devices in Asia. Surprisingly, while improving click-through rates by over 66.80% when compared to a baseline, [1] discovered that optimal ad-network revenues may not always go with optimal behavioral targeting regime. This is a phenomenon reflecting a market misalignment where privacy-preserving strategies may, paradoxically, better match with greater interest of users as well as long-term platform stability than with high-dimensional personalization systems.

On the social media front, [2] examined how strong consumer–brand relationship can overcome subjective risks in personalized advertising. Such an experiment confirmed high levels of affiliation countering concerns over privacy, while poor connections drive resistance, with indiscriminate targeting likely to backfire on fragmented audiences. In support, [3] also provided evidence that personalized advertising has a positive effect on value perceived and purchase intention, but only with low concerns over privacy. The risk is that intrusive targeting will destroy campaigns, and in particular when communicated covertly or in an unsuitable manner. Political advertising based on personality was examined by [4], with matching users' psychographic profiles showing a measurable effect on persuasion. These results also attest to escalating sophistication in targeting conduct, as well as raising ethical concerns, when conducted covertly and without users' consent.

Even as platforms try in the direction of transparency, tools prove to be inadequate. Apple's Privacy labels, intended to make data more transparent, have been found by [5] to be poorly readable by users. Participants in that research were baffled by jargon and structural inconsistencies between app permissions and data disclosures, cutting off its day-to-day usefulness. This is representative of larger concerns addressed in [6], where post-GDPR web ecosystem continues to reveal glaring shortcomings. Although formal transparency was increased by that regulation, real usability and accessibility of data rights remain absent, at times even actively disenabled by "dark patterns".

More structural analysis is offered by [7], with personal information introduced not as a commodity in and by itself, but as a source for valuation regimes for Big Tech. As explained, "Techcraft" is how these companies force engagement toward predictive values on an economic level, eliminating concerns with consent or ownership altogether. This deeper structural development is behind this argument for redesigning value extraction from users' data, in aid of alternative systems such as those offered here. Attempts at syntheses and coordination on research in online behavioral advertising's distributed landscape are offered by [8]. These synthesize nearly two decades' worth of research and suggest a shortfall of integrative direction, and issue a call for integrative strategies and enhanced theory-based foundations. In contrast, [9] offers a behemoth, curated database of over a million privacy policies and reveals how these have progressively become more impenetrable over recent years. Despite policy measures like the GDPR, underlying disclosures continue to remain absent or obfuscated, suggesting a gap between practice and policy.

In [10], Protection Motivation Theory (PMT) was employed to study privacy practice and concluded that even when users perceive threats as serious, they perceive themselves to be powerless when addressing them. The powerlessness, combined with low actual-world sensitive data protection effort (e.g., cookie deleting) exhibits one deficit in empowerment as much as system design. The framework here in this study addresses this very question posed by [10]. Strategically, [11] analyzed how firms diverge with data strategy, showing how competition involving regulatory constraint and consumer mistrust produces different organizational responses, anything from aggressive monetization to defensive data sharing. Work here notes that concern with the privacy-performance trade-off is not one-size-fits-all and that company strategy has to take both ethical design as much as technical innovation into consideration. Computational research on predicting users' responses was summarized by [12], where the authors reported on pervasiveness in the environment by ML models. Technically, as impressive as this research is, it omits explorations on transparency or on augmenting user control, and thus exhibits that while algorithmic efficiency improves, that on trust, as much as compliance with regulations, lags behind.

The state of tracking in the post-cookie era was investigated in detail by [13], with an outcome such that tracker methods have migrated towards stateless ones like fingerprinting, ID leakage, and ID synchronization. Most disconcertingly, over 75% is taking place before users have even had a choice to use a cookie, or while opting-out actively. This helps to firmly substantiate this research's first hypothesis, that consent mechanisms are ineffective, and tracking has migrated towards more long-term and invisible methods. This study provided key empirical evidence for this theory that consent banners, as presently implemented, are symbolic rather than real. On the other hand, [14] took a human-focused approach, analyzing how users use privacy-enhancing technologies (PETs) and finding a broad disconnect between what sites are offering and what users use. Based on the survey by [14], it was concluded that most users are becoming educated on privacy features through unofficial sources (friends, relatives, etc.), suggesting a systemic failure with current systems to empower users. The study provides an added insightful perspective for this research in that it helps to illustrate the futility of current PET integrations, once again supporting a shift towards more open and consensual data sharing systems.

The behavioral impacts of personalized advertising are examined by [15] and show persuasive possibilities as well as the dangers of online behavioral advertising (OBA). Using SEM methodology allows systematic observation involving the double effect: personalization maximizes engagement, but ad avoidance is strongly implicated with concern about privacy. Despite its consumer-behavior orientation, it helps to identify how data exploitation strategies influence user trust as well as behavior, again affirming ethical redesigning preference data-sharing systems as a necessity.

On the development front, [16] they consider how ad networks target app developers with UI framing and nudging. Developers preferred less intrusive types of adverts when subjected to more transparent effects on privacy. This introduces an influential stakeholder, the developer, to the world of privacy. It is germane because current systems do not even provide for how backend decisions by developers shift users' data exposure. This discovery highlights how much ethical nudging by developers at a holistic level is critical. The authors concluded their findings by saying that "the developers need transparent options". And that is exactly what this proposed framework is. It does gives control in the users' hands and provides the developers and all other entities involved here with accurate data directly from the users themselves.

On the other hand, [17] takes a macroeconomic view of GDPR and argues that it has de facto concentrated monopolists like Google and Facebook. The paper attributes regulatory efforts to competition concerns, hypothesizing that current data protection laws, for all they benefit privacy, have enabled data power to shift towards accumulating in its hands. This is a sober reminder: unless novel models disrupting data collection monopolies are introduced, regulation alone may solidify the very mechanisms it aims to regulate. By contrast, [18] conduct a technical analysis of Apple's ATT and Privacy Nutrition Labels and report partial success. While tracking by IDFA has diminished, fingerprinting and cohort-based tracking remain unabated. Some apps even use backend logic or server-side code to circumvent limitations. These developments highlight the need for more transparent, enforceable, and attestable privacy mechanisms. Furthermore, the unclear nature of some disclosures also presents a case for a system-enforced, user-transparent privacy mechanism.

Adding more evidence to these claims, [19] also illustrates that Android apps largely engage in third-party tracking without consent, in violation of GDPR and UK privacy laws. Tracking libraries offload the burden on developers with little support or compliance frameworks. This is a systemic fault, everyone is guilty, no one is in charge, further supporting this study's approach of compliance by design frameworks. An online survey was conducted by [20] to examine how users perceive algorithmic profiling in Facebook. The findings illustrate users often do not appreciate or mistrust categories inferred to profile them, and there is a lack of transparency and an emotional disaffection with profiling systems. The description "algorithmic disillusionment" provided by [20] is particularly insightful; without users' comprehension or trust in algorithmic systems, consent is invalid. The overall message from [20] is the problem addressed by this study, which is that the users are tracked through their usage of the platform, in order to profile them, which inevitably results to inaccurate profiles. Then, [21] provide an exploration into the legality behind Real-Time Bidding (RTB), finding it by its very nature inconsistent with GDPR because of its inability to secure a legal basis for data processing, inability to meet transparency requirements, and inherent insecurity. A direct source of validation for a fundamental argument made by this research, by their legal analysis, is that AdTech infrastructure is, by its nature, incompatible with legal and ethical standards. Therefore, the regulations are not enough to secure a user's data.

On analyzing mobile and online environments' privacy dynamics, several studies have offered critical observations regarding digital platforms' management of users' information and its impact on the protection of privacy.

The research in [22] leverages Android and iOS apps' comparative privacy capabilities and reveals unsettling behaviors on both platforms. Third-party tracking, as well as sharing unique identifiers, were both prevalent, even among apps designed for children. While there was less advertising-based tracking on iOS apps, they were discovered to have more frequent access to kids' devices' locations than Android apps. This renders simplistic assumptions about the superiority of one platform's way over that of its competitor, and addresses the challenges that both platforms have compliance with privacy legislation such as GDPR. The evidence and analysis provided by [22] align with this study's underlying concerns, as pervasive disrespect for privacy legislation addresses the pressing necessity for a better privacy solution based on transparency and consent.

Likewise, [23] investigation on how users' opt-out of tracking cookies is impacted by legal and technical competence in terms of data collection highlights how enabling legal competence has encouraged consumers to actively opt out from data collection actions. Especially in recent years with such legislation as GDPR aimed at enabling consumers to exert greater control over personal data, the studies highlight that consumers are more likely to act (for their own online safety) when they have information about what legal ramifications follow data collection, and this serves to further highlight how better education for users is needed in conjunction with implementing privacy legislation. This aligns with the goal of creating systems that not only adhere to legal requirements but also empower users with greater choice and understanding. On a technical note, [24] demonstrates how conclusions are drawn differently when popular automatic crawlers are utilized compared to more realistic configurations for making measurements on the web in terms of privacy. The researchers argue that existing measuring tools are not able to adequately mimic actual users' experiences, and partial or incomplete conclusions form as a result. The research highlights that unless tools are adapted to actual users' action, seminal threats to privacy can be overlooked. This highlights a glaring shortfall in how research on privacy is conducted and aligns with the need for better, more accurate, and people-oriented methodologies for measuring privacy, an area on which this research aims to contribute forward-looking responses.

The use by online advertising of cookies is analyzed by [25] in considerable detail, by laying out how cookies form an extensive digital market infrastructure. The employment as a building element for online marketing and tracking by cookies is challenged here, with doubts over such systems' viability, particularly in this day when increased regulatory scrutiny exists. This consideration on "datafication" on human action is an important one in balancing trade-offs over commercialization of data and privacy. This research lends weight to concerns over having to rebuild ethics and operating procedures for tracking data and calls for innovative paradigms to move forward beyond an established cookie paradigm. Another critical analysis on how adoption has been impacted by GDPR is provided by [26], with e-commerce revenues and pageviews down by 12% when implemented. This economic analysis demonstrates that privacy regulations, while crucial for consumer protection, also carry real consequences for website performance and revenue. The findings here are key as these reflect more subtle consideration for trade-offs over protection for people's financial performance as well as people's privacy, an issue at the core of this research. By looking at techniques that minimize intrusions on people's privacy without sacrificing functionality, this research tries better to balance people's needs for protection of information with usability.

For further evidence supporting the whole regulatory framework, [27] study the impact of GDPR's opt-in principle, and conclude that it reduced consumers by 12.5%, with the remaining consumers becoming more traceable and therefore more valued by advertisers. The impact is a demonstration of the perverse effect of privacy legislation, whereby users of privacy influence others' actions, to advertisers' benefit. It is representative of the need for more sophisticated models for users' consent, that are able to exclude such privacy spillovers while maintaining control by users over information. One such design area for privacy features was proposed by [28], with particular consideration for the point that users struggle with getting around privacy features on the Internet of Things (IoT). This theoretical framework [28] proposed on building effective privacy features can also apply to other platforms, as an inspiration to building systems offering users more control over information regarding them. The idea behind this user-centered design is at the heart of the proposed research in that it is trying to build a system that not only complies with protocols on privacy but also enhances the user experience with open and

informed choice-making that users themselves take and communicate to platforms on voluntary grounds.

The dynamics behind cookie consent under GDPR compliance have been studied intensively by [29] and [30]. In [29], the researchers were interested in how cookie banners fail to provide users with adequate and accurate information to ensure consent, with an enormous 89% of banners in breach of legislation. This research frankly challenges the validity of current consent mechanisms on grounds of a presumed deficit in transparency and linguistic abuse as key barriers to users' comprehension. Conversely, [30] provides an engineering-based answer in the form of the CookieBlock add-on, grouping cookies automatically and imposing GDPR compliance. The research presents an implementable answer to concerns raised by research such as that by [29], taking into consideration how much scope there exists for tools to automatically assist users' privacy regardless of website cooperation.

On a slightly different subject, [31] investigated online student and parent cyberbullying behavior with an extension of the Theory of Planned Behavior (TPB) with a new construct, cyber parenting, to better explain online conduct. Its findings indicate a growing need for healthy digital conduct and surveillance, particularly by and on behalf of younger users and parents. This behavioral aspect has special relevance to the modern debate over digital marketing conduct now, where users' conduct, all too often removed from context or misinterpreted, fuels sneaky preference profiles. Just like how research by [31] demands worth from precise, conduct-informed data-gathering with regard to productive insight, this current concept proposes an alternative paradigm where users explicitly articulate ad/content interest themselves. Digital literacy and clear policies are necessary in this era. The proposed framework here avoids such artifacts like device-shared misguided presumptions or misclicks and reinstates users' agency, in line with the overall goal of ethical digital practice.

The comprehensive 2023 survey of [32] on American views of data privacy reveals a widespread erosion of public trust and a deep sense of powerlessness among users in the face of large-scale data collection practices. With 81% of respondents expressing concern over how companies use their data and 73% admitting they feel little to no control over that data, the findings highlight a pressing contradiction between user expectations of autonomy and the opaque, algorithm-driven surveillance ecosystems currently in place. Even more striking is the nearly universal concern over data practices affecting children, 89% worry about social platforms holding personal information on minors, and 85% object to advertisers using behavioral data to target young users. As platforms increasingly integrate AI to personalize content, nearly half (48%) of Americans find such applications unacceptable, yet remain trapped in systems where opting out is either obscured or functionally impossible.

Lastly, [33] offers a better perspective into the current global perception and operational stance on digital privacy, reinforcing the urgency and credibility of user-centric data governance models. With over 2,600 respondents across 12 countries, the study identifies strong global support for privacy legislation, 86% of organizations now view such regulations positively, marking a noticeable year-over-year increase. This reinforces that companies are not just legally compelled but also strategically motivated to respect user privacy. At the same time, nearly all participants (96%) asserted that privacy investments yield greater benefits than costs, underscoring a growing recognition that responsible data practices are not a trade-off but a competitive advantage. The study's findings validate the broader shift in enterprise thinking, which gives strong credence to the need for a transparent, structured, and user-authored preference platform as both a privacy measure and an ethical data-sharing personalization solution.

A comparative analysis of existing personalization, privacy, and tracking mitigation approaches versus the proposed user-governed preference framework is summarized in Table I below. Collectively, the reviewed literature demonstrates that contemporary personalization and advertising systems are constrained by a shared structural dependency on inferred behavioral data. Across regulatory, technical, economic, and human-centered perspectives, existing approaches attempt to mitigate harm through transparency tools, consent mechanisms, anonymization techniques, or post-collection controls. However, these interventions operate downstream of the core design assumption that user preferences must be *observed*, *predicted*, or *approximated* from activity traces. This assumption persists even when tracking is partially restricted, anonymized, or made less identifiable, resulting in persistent noise, contextual misinterpretation, and diminished user trust. The literature, therefore, converges on a common limitation: current systems focus on controlling or obscuring data collection practices rather than rethinking the origin and authorship of preference data itself.

Moreover, while several tools and extensions aim to block, reduce, or anonymize tracking, they remain platform-specific, require continuous maintenance by developers or organizations, and frequently shift the burden of privacy protection onto users without offering a constructive alternative. Anonymization-based approaches, while reducing identifiability, often degrade content relevance and recommendation quality, particularly in contexts involving shared devices or younger users, where contextual sensitivity is critical. Notably absent from the literature is a system-level framework that replaces behavioral inference with explicit, user-authored preference declaration in a standardized and portable format. This gap motivates the present work, which reframes ethical personalization not as a problem of limiting surveillance, but as one of redesigning how preference data is generated, owned, and shared at the source.

The following sections, namely subsection “A” and subsection “B” are extended explanations for the conclusive data in Table I.

TABLE I. COMPARATIVE ANALYSIS OF EXISTING PERSONALIZATION, PRIVACY, AND TRACKING MITIGATION APPROACHES VERSUS THE PROPOSED USER-GOVERNED PREFERENCE FRAMEWORK

| Dimension | Behavioral Tracking–Based Personalization | Regulatory & Consent-Based Mechanisms | Privacy Tools & Tracking Blockers | Anonymization & Aggregation Approaches | Proposed User-Governed Preference Framework |
|---|---|---|---|---|---|
| **Primary Data Source** | Implicit behavioral signals (clicks, views, dwell time, device data) | Same as tracking systems, with legal constraints | Same as tracking systems, partially suppressed | Behavioral data after de-identification | **Explicit, user-authored preference declarations** |
| **Mode of Data Acquisition** | Passive, continuous surveillance | Passive collection with formal consent | Passive collection with selective blocking | Passive collection with identity masking | **Active, voluntary user input** |
| **User Awareness** | Low to none | Limited, often symbolic | Moderate, tool-dependent | Low | **High and intentional** |
| **User Control Granularity** | None | Coarse (opt-in / opt-out) | Binary (block / allow) | None | **Fine-grained, category- and context-level** |
| **Reliance on Behavioral Inference** | Core dependency | Core dependency | Core dependency | Core dependency | **None** |
| **Data Noise & Misattribution** | High (misclicks, shared devices, incidental actions) | High | High | High | **Low (self-declared intent)** |
| **Context Sensitivity** | Poor | Poor | Poor | Poor | **Explicitly contextual via multiple profiles** |
| **Impact on Recommendation Accuracy** | Variable, often inflated by noise | Unchanged | Often degraded | Often degraded | **Improved via cleaner input data** |
| **Effect on User Experience** | Intrusive, opaque | Fatiguing, confusing | Disruptive or incomplete | Often less relevant | **Transparent and intention-aligned** |
| **Suitability for Children / Shared Devices** | Poor | Poor | Poor | Poor | **High (explicit profiles avoid misattribution)** |
| **Platform Dependency** | High (platform-owned data) | High | High | High | **Platform-agnostic** |
| **Interoperability Across Platforms** | None | None | None | None | **Designed for cross-platform use** |
| **Maintenance Burden** | Platform-side | Platform-side | User and developer dependent | Platform-side | **Minimal once profile is created** |
| **Scalability Across Ecosystems** | High (centralized) | High | Limited | High | **High via standardization** |
| **Compliance with Privacy-by-Design** | Weak | Partial | Partial | Partial | **Intrinsic** |
| **Ethical Alignment** | Low | Moderate | Moderate | Moderate | **High** |
| **Transparency of Data Semantics** | Opaque | Opaque | Opaque | Opaque | **Explicit and inspectable** |
| **Data Ownership** | Platform-controlled | Platform-controlled | Platform-controlled | Platform-controlled | **User-owned** |
| **Long-Term Sustainability** | Increasingly contested | Regulatory-dependent | Fragile | Fragile | **Systemically sustainable** |
| **Primary Limitation** | Systemically sustainable | Symbolic control | Fragmentation and relevance loss | Accuracy degradation | **Adoption and standardization effort** |

### A. Current System and Efforts Made to Overcome These Issues

User tracking is an omnipresent, albeit contentious, element of online advertising and targeted content in modern online society. A variety of sophisticated methods are used by websites, apps, and online services to collect great volumes of information on what users do, typically without users' conscious or active consent. The tracking is executed by a variety of tracking technologies, such as cookies, pixels, device fingerprinting, and device identifiers, by means of which companies are able to build very rich profiles on users based on how they use a range of different sites and services.

#### *1) How the Users are Tracked (for Ads and Content Preferences)?*

Tracking is also a multi-faceted practice that is carried out at virtually every phase of online activity. These businesses want to create a profile about users that will allow them to provide highly targeted, personalized advertising, optimizing digital promotion campaigns. Among the most critical means that are leveraged to track users are:

- Cookies and Tracking Pixels: These are little pieces of information stored on users' devices that track what users do on websites. Cookies allow advertisers to track users' activity data on different websites and connect it with individual users' profiles.
- Device and advertising identifiers (AdID, IDFA): Google's AdID and Apple's IDFA are two device identifiers marketers use to identify users across apps and services. These identifiers are associated with unique profiles belonging to users, making it possible for marketers to target users with personalized advertising based on their app-based activities.
- Browser Fingerprinting: It uses unique characteristics of the device of the user (like screen resolution, fonts, and operating system) to create a fingerprint for tracking users from website to website, even without cookies.
- Behavioral Analytics: Sites observe how long a consumer stays on a page, what they click on, and how they interact with a piece of content. From this, they deduce what a consumer prefers, even when a preference point has not been explicitly stated by them.

Despite its blatant encroachment, companies will most often justify this nonconsensual data collection or tracking by calling on the value in delivering users targeted advertising. The underlying idea is that, by tracking users, companies are better able to customize the experience and are thus better able to target advertising and content at users, making it more relevant and, as such, more valuable to users. This argument, however, is most often employed as a pretense to cover over a darker reality: invasion of privacy.

#### *2) Advertising Relevance as a Scapegoat for Privacy Infringement*

The term "relevant content/ads" is widely marketed as a quid pro quo for loss of privacy. Consumers believe that it is an acceptable trade-off to provide information to online sites in exchange for targeted advertising. The most frequent justification offered by the major tech organizations is that the fundamental motivation for tracking is to improve the consumer's user experience by showing them more targeted advertising and relevant content. However, this presupposes that users willingly agree to the tracking, a premise that is often questionable at best.

In reality, most users have no idea on what magnitude they are having data mining carried out on them, nor do they have an idea of long-term effects of such data becoming an input to algorithms for perpetual behavioral prediction. Additionally, ubiquity of tracking mechanisms leaves little or no choice for users to opt out, while "relevance" boasted by such mechanisms is overpowered by exploitation of users' personal data. The ad is shown not only on users' interests but also on inferred behavioral propensities without users' notice. This has created serious concerns regarding data ownership, consent, as well as manipulation of users' actions.

#### *3) Platform-Based Efforts*

There are several systems that are used to manage tracking and privacy in online marketing, and these systems have been heavily criticized as failing to keep users' information safe or offering adequate control over what is delivered. The most common are:

- AdID and IDFA: The two dominant identifiers for mobile device-based cross-app and cross-platform tracking are Google's AdID and Apple's IDFA. These two identifiers allow advertisers to monitor users' actions within applications and assign actions such as purchases or advertising interaction data to a specific user profile. While there have been efforts by Google and Apple to limit how such identifiers can be utilized (e.g., Apple's App Tracking Transparency (ATT) policy), the process remains to support tracking at mass levels based on inferred behavior rather than direct user consent.
- App Tracking Transparency (ATT): Apple introduced the App Tracking Transparency framework with its shift towards curtailing invasive tracking. The ATT framework requests apps to obtain consumers' permission before they can track consumers' actions on other apps and sites. While this was hailed as a consumer privacy victory, it has not solved the problem entirely. The majority of users, even when provided with an opportunity to opt out, are not aware of what this implies or do not have sufficient technical expertise to make an educated choice. The ATT mechanism also relies on an inferred user interests model, with users having no real control over where and how their data is used and shared.
- Privacy Sandbox: Google's Privacy Sandbox is an initiative to create alternatives to third-party cookies to preserve user privacy. The system is intended to provide advertisers with the tools to target users without making personal information directly available to advertisers. However, the system involves in-app activity-based profiling and attempts at deducing the desires or interests of users, an activity most detractors say still eliminates consumer control over personal information.

These efforts, as much as they are attempts at confronting privacy concerns, still don't provide users with full control over how they use their data. Despite efforts like ATT and Privacy Sandbox, users remain profiled outside of their direct, conscious consent, and companies still have considerable control over what data they collect and how they use it.

*4) Data Protection Acts and Policies in Place*

After increased online privacy concerns, several data protection acts have been enforced globally to safeguard users' personal information. Among them are two key laws:

- General Data Protection Regulation (GDPR): Implemented by the European Union in 2018, GDPR grants users robust data protection rights. It compels organizations to obtain users' consent first when gathering and processing their personal data. It also grants users the right to review, modify, and delete their information, among other rights. Enforcement has been spotty even with its strengthening, and businesses have kept finding loopholes in these regulations.
- California Consumer Privacy Act (CCPA): The CCPA, enacted by California in 2020, gives California users the right to know what information is gathered on them, gives them a right to access it, and gives them a right to have it removed. While more open than previously, it still does not address at its root the problem with tracking consent. It does not require companies to stop tracking at all, only provide an opportunity for users to opt out of having their personal information sold.

Although such legislation provides protection for users to some degree, it does not fully address consent to tracking, nor is its enforcement a success. Besides, they concentrate on collection and processing as opposed to tracking activities and users' lack of control over their online identities.

The current tracking system, with its target advertising and content tailoring, by its very nature disrespects users' trust. Even with measures such as Privacy Sandbox and ATT, they don't even come anywhere near making users have full control over their data, and most users aren't even aware of how much tracking is going on, as we learned through the literature. Regulations such as GDPR and CCPA are there for some protection, but they don't address the core issue of getting users' consent when it comes to tracking. The models currently in use, by inferred interest and data profiles, are by nature, defective and do not allow such ideals as transparency, autonomy, and privacy that users should have.

## B. Weaknesses and Improvement Gaps in the Current System and Efforts

Despite statutory data protection and transparency legislation, current online advertising and recommendation mechanisms are technically flawed. The technical flaws are not so much legal or ethical, but rather a design flaw intrinsic in how users' data is collected, inferred, and utilized. The mechanism based on pervasive surveillance and covert profiles cannot simultaneously grant real control to users and reduce advertising and content targeting quality.

*1) The Limited Role of GDPR, CCPA, and Other Data Regulations*

Regulations like the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) have been introduced to give users greater control over information about them. While these regimes provide legal protection, they are deficient in several key areas:

- Consent fatigue: Consumers are fatigued with cookie notices and consent banners that try to confuse or push users toward compliance. The majority click "Accept All" without reading and understanding what it means.
- Gross unenforceability: Despite GDPR mandating minimization and use limitation, its enforcement is aimed at pursuing large-scale blatant abuse and ignoring day-to-day misuse of inferred information and tracking loopholes.
- Vague "legitimate interest" clauses: Companies justify data collection in terms of "legitimate interests," with grey areas where profiling continues even when users try to opt out.

Lastly, laws such as these do not ensure ethical practice. The spirit of such legislation is always evaded by way of vagueness and exploitation by users.

*2) Inferred Interests: The Core Technical Fallacy*

At its heart, it is a risky presumption, one that machines are able to successfully infer users' interests from behavioral data.

- It uses passive data such as search history, clicks, scrolling, watch time, and hover actions to train algorithms to construct a profile of "what the user wants."
- Context is overlooked: The consumer may only want "baby strollers" due to an isolated reason, or on behalf of someone, or perhaps out of interest. The system is assuming this is an interest at the top and is showering them with baby advertising for months.
- Temporal interest fading is poorly implemented. Interest profiles are infrequently time-limited, and dated conclusions will continue long past when they should be applicable.

These systems are essentially flawed; they sacrifice accuracy for automation and are making assumptions where there should not even be any. This results in poor user experience as well as poor targeting.

*3) False Positives and Shared Device Challenges*

Dependence on device or browser-based behavior leads to an exceedingly high rate of false positives:

- Shared devices (e.g., schools, families, or cyber cafes) lead to different behaviors being correlated with a single profile.
  - One innocent action by a child will initiate travel high-ends advertising.

- A roommate's impulsive late-night shopping spree may also result in product suggestions that are not appealing to the original user.

- Disproportionate effect of single actions: A misclick, unintentional click, or experimental search may significantly skew ad/content suggestion for days or weeks.

This lack of disambiguation leads to data pollution, where extraneous interests get confused with real preferences, affecting advertiser ROI and user experience.

*4) Non-Consensual and Opaque Profiling*

Modern systems don't ask for preferences, they deduce them with no input from users:

- Users are rarely informed about how their data profiles are built.
- Inference models are not open-source, and most platforms don't allow users to review or verify what has been recorded about them.
- Profiling persists across devices and apps even when users erase identifiers or opt out, and it is most often carried out with browser fingerprinting, shadow profiles, and SDK data collection across apps.

Fundamentally, users are profiled by systems they don't understand, based on signals they didn't agree to provide, for purposes over which they have no control.

*5) Clunky and Deceptive Preference Management*

While there are some sites (Google, Facebook, Apple) that do have ad preference controls, they are plagued with usability problems:

- Hidden in menus: Choices are hidden behind several layers of menus with unclear labels.
- Auto-enabled interests: Platforms pre-define interest categories based on historical tracking and force users to opt-out.
- Inadequate granularity: Options are abstract (e.g., "Sports," "Travel") with no way to state sub-preferences or rule out presumed ones.
- No true opt-out: Even when interests have been opted out, there remains background behavioral tracking, and there is still targeting on other inferred information.

These are not intended for real control, rather to seem compliant while still being trackable.

*6) Defective Opt-Out Mechanisms*

Choosing not to participate in personalized advertising now is an illusion:

- Ad identifiers (e.g., Google's AdID or Apple's IDFA) may get cleared, but tracking is resumed shortly thereafter under a fresh identifier.
- Mechanisms such as opt-out cookies or Do Not Track (DNT) are not commonly honored or are erased when cookies are cleared.
- App-level opt-outs, such as ATT on iOS, only stop certain types of cross-app tracking while enabling server-side code tracking and profiling to still continue.

In essence, users are struggling with a hydra, severed from one tracking source, and two others from less obvious sources.

*7) Technical and Ethical Failures in Real-World Scenarios*

Consider these case studies and examples uncovering the serious shortcomings in current profiling systems:

- YouTube's feedback loop: Someone accidentally watches a video about a conspiracy theory. The site overloads them with related videos for weeks, propagating fringe viewpoints, a documented phenomenon called "algorithmic rabbit-holing".
- Amazon's Purchase Memory: One purchase of a vacuum cleaner leads to months of promotions for vacuum cleaners, something that the buyer won't need again for years.
- Political Ad Micro-Targeting: People are subjected to emotionally manipulative political messages during elections that are tailored to inferred personality attributes (e.g., fearfulness or extroversion) from past clicks and social media activity, and this is ethically objectionable when it impacts democracy and consent.
- Instagram Shopping Feeds: Your device is used by your friend to search for fashion goods, and your feed is filled with clothes from this domain, irrespective of it not aligning with your interests.

These instances depict how tracking often fails, and users are harmed because of poor inference.

*8) The Systematic Problem: Tracking is the Default*

We now have an online space where tracking is the rule and opting out is an exception. "If you're not paying, you're the product" has become a truism, though little is mentioned about how this business model is harming users and advertisers:

- Users lose control and confidence in the technology.
- Advertisers waste money targeting people with irrelevant ads/content based on noisy, inaccurate, or outdated preference profiles.
- Developers incorporate SDKs and third-party libraries that expose users' data without adequate transparency.
- Other platforms are also in the spotlight, as shown by recent events and consumer-led revolts (e.g., against Instagram and Facebook).

Not only is this model unviable, but it is also technically defective and ethically deficient.

## III. METHODOLOGY

After critical analysis of the literature on Data Integrity and Privacy issues, it seems necessary to obtain a broad knowledge base regarding actual users' views on digital tracking, data privacy measures, and alternative advertising models' perceived acceptability, a quantitative approach was utilized.

### A. Target Audience

A carefully designed questionnaire was shared among several online forums and communities focused on data privacy matters. These sites included boards and groups, where users typically have a more advanced understanding of digital rights, data handling practices, and implications of tracking technologies. This selection was intentional because individuals having a core understanding about data privacy provide richer answers, something indispensable to ideas like preference systems and user-consensual data sharing platforms. It was agreed that if this survey is distributed to normal public, then the results would be quite inaccurate, because, as it was learned through the literature review that the normal public usually doesn't have knowledge on technology being utilized to get their data and the Data Protection Acts in place that try to stop such Tracking.

### B. Survey Structure

The questionnaire included a set of Likert-scale, multiple-choice, and open-ended questions aimed at gauging participants' opinions about mandatory behavioral profiling, current data privacy regulations, the level of clarity regarding data use, and the feasibility of creating user-defined preference profiles. General themes examined included the perceived effectiveness of legislation like the GDPR and CCPA, the level of satisfaction with current opt-out features (such as App Tracking Transparency and the Privacy Sandbox), and attitudes toward the use of personal data to allow for better content personalization using artificial intelligence.

To protect the target audience's privacy expectations, the survey avoided collecting personally identifiable data, like age, gender, or geo-location. Rather, it focused on collecting relevant information pertaining to privacy expectations, trust in digital platforms, and potential participation in a new, permission-based model for advertising and content preferences data sharing platform. The tone of the questions was kept at a friendly level in order to let users select the best option, which reflects not only their opinion, but also their feelings. The survey was shared across privacy-oriented subreddits, encrypted messaging apps, and federated platforms, thus ensuring collected data was a true representation of users concerned with privacy. These responses were anonymized, and participants gave informed consent to participate in this study, along with explicit disclaimers provided informing participants about the non-commercial nature of the work. The design of this approach was not intended to provide generalizability to a wider populace, but merely to measure whether a well-informed audience would see value in this proposed framework in comparison to invasive tracking.

## IV. RESULTS AND DISCUSSION

A total of 572 valid responses were collected from privacy-aware online communities. The results provide empirical grounding for evaluating whether a user-authored, consent-driven preference framework is socially acceptable, practically viable, and theoretically aligned with existing privacy research. Microsoft Excel was used to create figures and charts for easy evaluation. Each subsection below discusses one thematic result and situates it within prior findings.

### A. User Sentiment on Tracking and Data Collection

As illustrated in Fig. 1, a dominant majority of respondents (62.59%) regarded behavioral tracking as inherently violating privacy and rejected it outright. An additional 30.77% indicated conditional acceptance of personalization, but only when preferences are explicitly provided by users themselves, rather than inferred through passive surveillance. By contrast, only 2.10% were comfortable with tracking in exchange for free services.

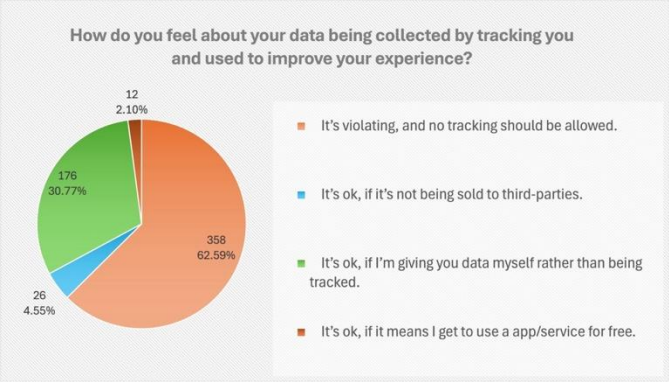


Fig. 1. User sentiment on tracking and data collection.

This finding directly reinforces prior work demonstrating that method of data acquisition, rather than personalization itself, is the primary driver of user discomfort [2], [3], [15]. It also empirically supports observations in [20] that users distrust inferred profiles and experience what has been described as algorithmic disillusionment. The strong preference for voluntary data provision further aligns with Protection Motivation Theory findings in [10], where perceived lack of agency suppresses protective behavior even among informed users. Collectively, these results validate the study's premise that user-declared intent is more socially acceptable than algorithmic inference, even when outcomes (e.g., personalization) appear similar.

### B. Attitudes Toward Existing Data Protection Mechanisms

Fig. 2 demonstrates widespread dissatisfaction with current regulatory approaches. Only 2.27% of respondents considered existing data protection laws sufficient, while 70.98% explicitly stated that a functional solution is needed rather than additional rules.

This sentiment closely mirrors empirical critiques of GDPR effectiveness found in [6], [13], and [29], where compliance mechanisms such as consent banners were shown to be largely symbolic. The results also echo macro-level analyses in [17] and [27], which argue that regulation alone can unintentionally consolidate data power instead of redistributing it. Respondents' desire for actionable tools aligns with findings in [14] and [23], where users expressed frustration at being formally empowered yet practically constrained. These results strengthen the

argument that compliance-by-design systems, rather than policy-only approaches, are required to operationalize privacy rights.

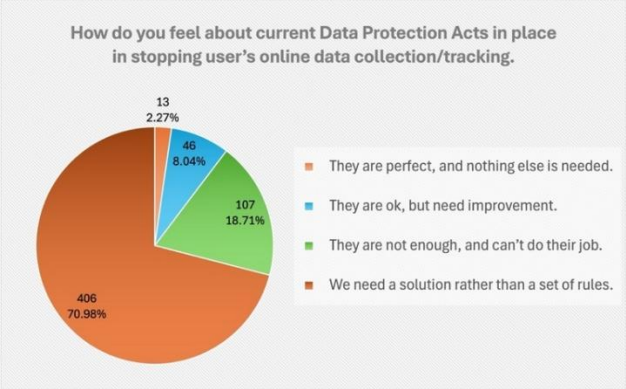


Fig. 2. Attitudes toward existing data protection mechanisms.

### C. Preferred Type of Solution

Responses in Fig. 3 reveal a nuanced, but decisive preference landscape. While 83.22% favored stopping data collection entirely, reflecting deep-rooted distrust, substantial support was also observed for voluntary preference sharing (77.80%), prohibiting third-party data sales (72.90%), and data anonymization (56.47%).

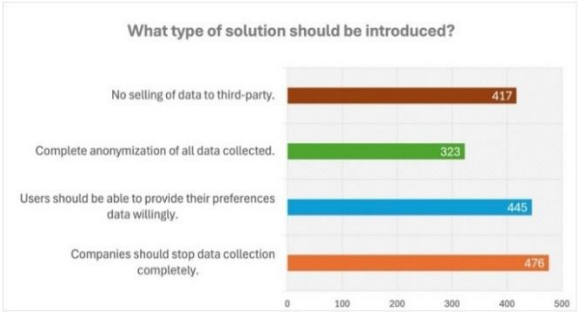


Fig. 3. Preferred type of solution.

This apparent tension between idealistic rejection and pragmatic compromise is consistent with prior findings in [15] and [26], which highlight trade-offs between privacy protection and system functionality. Importantly, the strong endorsement of voluntary preference sharing provides empirical validation for alternative paradigms proposed in [7], [8], and [28], where value extraction is decoupled from covert surveillance. Unlike anonymization-only approaches, which can degrade content relevance and raise concerns for vulnerable users such as children [22], [32], user-authored profiles preserve contextual accuracy while respecting consent. The findings thus support a middle-ground model that balances personalization utility with ethical data governance.

### D. Willingness to Use a User-Controlled Preference Platform

As shown in Fig. 4, willingness to adopt the proposed model was notably high. A combined 86.19% indicated they would use such a platform either unconditionally (9.97%) or provided it remained simple to manage (76.22%). Only 7.34% rejected the need entirely.

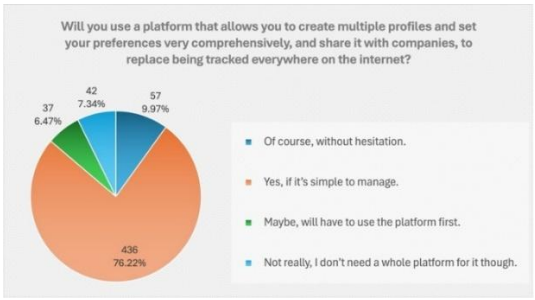


Fig. 4. Willingness to use a user-controlled preference platform.

These results reinforce usability-focused findings in [5], [14], and [28], where complexity and opacity were identified as primary barriers to meaningful privacy engagement. The emphasis on simplicity also corroborates observations in [10] that perceived effort reduces protective action. From a systems-design perspective, this underscores that technical feasibility alone is insufficient; adoption depends on intuitive interfaces and low cognitive overhead. The results further suggest that multi-profile support, allowing users to contextualize preferences, may address shortcomings of one-size-fits-all consent mechanisms.

### E. Trust in Who Should Build and Maintain the System

Trust attribution results in Fig. 5, reveal a pronounced institutional divide. Privacy-focused organizations (90.20%), open-source communities (73.60%), and non-profits (52.80%) were overwhelmingly preferred, while Big Tech (9.97%) and governments (28.50%) ranked lowest.

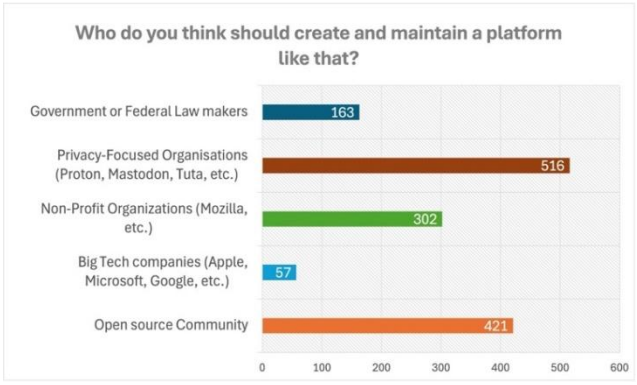


Fig. 5. Trust in who should build and maintain the system.

These findings are consistent with broader trust erosion documented in [32] and [33], as well as structural critiques of Big Tech's data governance in [7] and [17]. The strong preference for open and privacy-native stewardship aligns with research advocating transparency, auditability, and decentralization as trust-building mechanisms [18], [21]. This suggests that for the proposed framework to be credible, governance structure is as critical as technical design, favoring open standards, inspectable codebases, and non-commercial custodianship.

### F. Discussion and Implications

To conclude, the results demonstrate a clear mismatch between current data collection practices and user expectations. While prior research has extensively documented the harms and

inefficiencies of behavioral tracking [1], [13], [21], this study contributes empirical evidence that even privacy-literate users favor constructive alternatives rather than absolute disengagement. The findings support a paradigm shift from inferred behavioral profiling toward explicit, user-authored preference signaling, addressing long-standing concerns about consent validity, trust, and ethical personalization raised in [8] and [20].

Importantly, the results do not contradict existing literature, but rather operationalize its conclusions: users want personalization without surveillance, control without complexity, and privacy without sacrificing functionality. The proposed framework responds directly to these conditions by enabling consent to become an active, expressive process rather than a defensive reaction. In doing so, it aligns regulatory intent with technical implementation and user agency, offering a viable path beyond the limitations of both tracking-centric and regulation-only models.

## V. Proposed Model: A User-Governed Data Sharing Platform and Standard

### A. Overview of the Concept

The current ad environment is built on an interest-inference information system based on passive tracking. Companies track users through apps, sites, devices, and off-line activity and build hidden profiles by unknown means. Based on this inferred interest profile, supposedly "relevant" advertising and material is served. Its justification has always been: "you get relevant ads/content". Never once along the chain is anything asked of users: what they are interested in watching or how they want to get profiled. The system is by its nature defective.

The innovation here is a move from hidden tracking to open, user-consented preference sharing. Rather than getting tracked everywhere, users would have control over actively building and sharing interests on one platform. The platform described here can enable users to create their own Preference Profiles, to be linked to apps or sites in an easy, privacy-preserving way. Think of it as creating a social media account, but rather than uploading photos and updates, users select what they want to see: advertising, kinds of content, topics of interest, and more.

This is not an ad business model, but a technological solution to restoring control of data and trust by users in online environments. It is a simple solution, though one with pressing practical application as people's frustration over loss of control, over personal data tracking and its use in training AIs escalates.

### B. Core Features

User Accounts like Social Media: The initial point is a user account, identical to what you would have on a social media website. The difference here is that there is no messaging, posting, or networking aspect to this account. It is only used for storing preference and interest, only for passing on to apps or companies that people will be interacting with. Sharing is opt-in and under users' control.

User-Generated Interest Profiles with Individualized Preference Profile IDs: Any given user can create multiple Interest Profiles in their account. Each can have a unique alphanumeric Preference Profile ID as its "username" or API key. These IDs or direct downloads of JSON format of profiles may then be shared with sites that request advertising or content preferences from the user. This will enable users to use different profiles for different purposes without disclosing personal data or correlating all activity with one single identity.

Multi-Profile Support for ADs, Interests, Content Preferences: Users are not homogeneous, and their interests vary based on context. An individual might want new productivity applications' ads in a working environment and entertainment content during leisure time. Hence, this model supports multiple profiles per user account, which enables users to define various sets of interests for:

- Ads (e.g., job offers, tech products)
- Content (e.g., comedy, music genres, educational videos)
- General interests (e.g., travel, fitness, finance)

This system avoids having ad engines guess and instead relies on declarative data provided by users.

App-Specific Profile Preferences: There can be profiles created specifically for a particular platform. For example:

- A video application profile like a YouTube or TikTok profile, for instance, "ContentProfile_YT" with interests like "Tech Reviews, Music, No Politics, Comedy, Memes, Pranks".
- Another for professional networks like LinkedIn (e.g., "CareerAds_LI" with interests like "Project Management, Startups, Recruitment").

Apps only request a Preference Profile ID or a direct JSON file, and the user will provide them with one of their preference IDs or files. This avoids over-personalization with unwanted information and promotes context relevance. It is an ethical way of sharing your interest data for online content/ads, rather than being tracked everywhere online.

### C. Functionality

User Experience Design: The system must have optimal usability. Within each profile, the users should have the ability to select interests from pre-designed template options, search by categories, and build profiles through a guided, intuitive interface. The users should have a sense of empowerment rather than frustration with the process. Selecting pre-designed default template options (e.g., "Gaming Profile," "Work Profile," "Music Preferences") can provide optimal reduction in cognitive load without sacrificing fine-tuned personalization.

Linking and Sharing Mechanisms: The platforms can request Preference Profile IDs from users. This can take place through:

- Manual Entry: The users enter or paste their Preference Profile ID into an input field.
- Account Linking: As with OAuth, users can directly link their preference profile via an API handshake.
- File Upload: Directly upload the exported JSON file of that Preference Profile on the requesting platform.

- Username-based linking: Social apps already have a facility for linking via usernames, this method can emulate the same with Preference Profile IDs.

This will allow developers to have flexibility and ensure that consent is specific and accurate.

Integration with Applications and Sites: Developers can be offered an API or an SDK. This allows applications and sites to:

- Query a user's Preference Profile ID (read-only).
- Fetch updated preferences dynamically (optional).
- Respect content/advertising category exclusions (e.g., avoid political ads).

Access to all data would be governed by an open protocol without permanent identifiers or behavioral surveillance.

Open Standards, Open-Codebase: In order to foster trust and collaboration, the platform should be open-source under an open standard. This prevents vendor lock-in, and third-party auditing is feasible. Developers can propose categories, tailor template logic, and perform integrations. Open protocols also support modular implementations, and decentralized control or even self-hosting is feasible.

Decentralization and Identity Ownership: Decentralized Identity (DID) standards or principles of Web3 can expand on this model further. This gives users control over profiles with cryptographic ownership, makes the entire system tamper-resistant, and could support zero-knowledge proofs whereby users reveal preferences data without identity disclosure.

AI integration for Profile Creation: AI can easily be embedded within this platform. It can create a user's profile based on user's text input by harvesting the power of NLP techniques available in modern AI age.

Prompt style simple example:

*"I need a general profile. I like all music except jazz. I hate drama movies, but I love all other movie categories, especially comedy. I love everything related to technology, but I hate marketing and SEO stuff. I love the power of AI and what it can do, but I don't like AI generated content (as in music, video, photos, etc.) as it seems lazy. I'm an outdoor person, so I have many physical hobbies, like Sports and exercising. I enjoy traveling a lot, but I'm not a luxury freak. I love almost all animals, but I'm not quite fond of dogs or pigs. And when it comes to News, I absolutely hate politics. Please generate a good profile for me and try to understand my personality to get all the preferences correct."*

There're few ways it can be done, as listed below:

- Own Trained Model: This will probably be the best implementation, as the AI doesn't need to external knowledge base to be used. An AI Model can be trained on manually crafted examples and their outputs, which will allow for highly accurate User's Preference Profile creation.
- API Linking: An advanced AI Model can be used with the help of API Keys. This way when the users writes the text of what they want from the profile, the prompt can be sent to the AI Model's server using its API to get the categories list having users' likes and dislikes. This way the users have created their preference profile via an API handshake. In this, most of the Processing will be done by the outsourced server.
- Local AI model: This method is by far the toughest to get right, as the entire AI Model will have to run locally, taking up a lot of RAM and GPU power. It is quite inefficient in terms of what the overall platform looks to accomplish.

Overall, this brings us to an open, privacy-based internet with the actual user at its center. Some of these features are too advanced and out of scope to implement right now but are to be considered in future developments of the platform. In the survey conducted, the respondents emphasized simplicity of the platform, linking it to usability. The Demo web interface developed here ensures the basic requirements of the proposed system are met and showcases the simplicity of this solution's implementation.

### *D. Demo Website Showcasing the Workings of the Proposed Framework*

To illustrate the practical feasibility and interaction flow of the proposed user-governed preference data acquisition framework, a lightweight demonstration website was developed. The purpose of this prototype is not to present a production-ready system or a novel software contribution, but to concretely demonstrate how the conceptual framework described in the previous section can operate in practice from a user and platform perspective.

The demo implementation focuses on three core aspects of the framework: 1) explicit user-authored preference declaration, 2) multi-profile and context-aware preference management, and 3) standardized, portable export of preference data for third-party use.

Accordingly, the system was intentionally kept minimal in scope and complexity, prioritizing conceptual clarity over feature completeness or optimization. The prototype serves as a reference implementation to support understanding, reproducibility, and future extension by researchers, developers, or organizations.

#### *1) System Overview and Interaction Flow*

At a high level, the demonstration platform allows individuals to create and manage one or more preference profiles, each representing a distinct context of interest (e.g., content consumption, advertising, professional use). These profiles are explicitly authored by users through direct input rather than inferred from behavior. Once created, a profile can be selectively shared with external platforms via a unique identifier or exported in a standardized machine-readable format (JSON).

The interaction flow is intentionally simple: users register, define preferences through structured categories, optionally refine or update those preferences over time, and provide the resulting profile to platforms that request preference data. This flow reflects the framework's central design principle of

replacing passive behavioral inference with explicit, user-controlled preference declaration.

*2) Entry Point and Navigation*

The platform entry point is a minimal landing interface that introduces the purpose of the system and provides navigation to user-facing actions such as registration and login. As shown in Fig. 6, the landing page is deliberately sparse, reflecting the framework's goal of minimizing friction and cognitive overhead during onboarding.

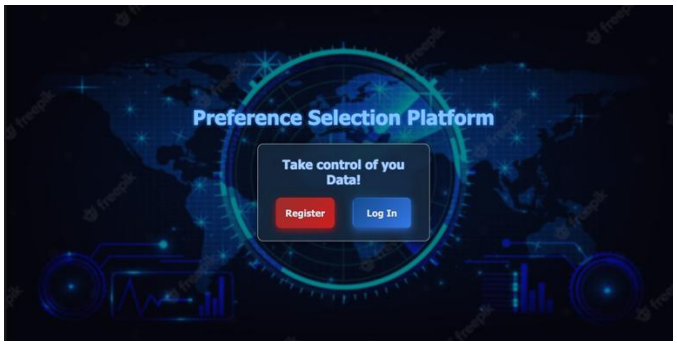


Fig. 6. Landing page.

*3) User Authentication and Account Access*

Users interact with the system through a basic account access that enables secure access to preference management functions. The registration and login interface, illustrated in Fig. 7, exists solely to associate multiple preference profiles with a single user.

Importantly, authentication is treated as an enabling mechanism rather than a profiling mechanism. No behavioral tracking or implicit data collection is performed during account creation or login.

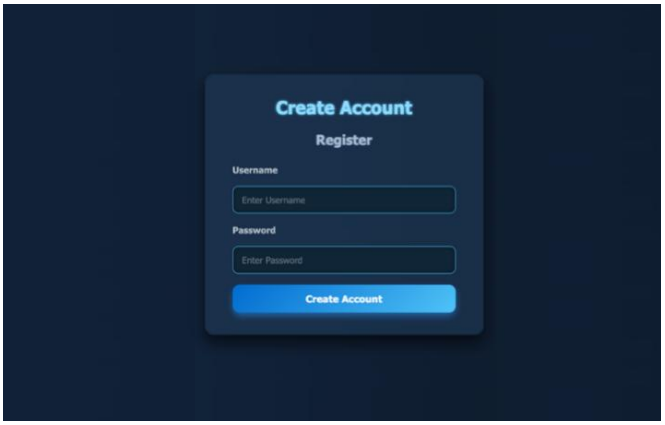


Fig. 7. Register page.

*4) Preference Profile Management Dashboard*

Once authenticated, users are presented with a centralized dashboard that lists all preference profiles associated with their account. As shown in Fig. 8, the dashboard functions as the primary control surface for managing preference contexts.

Each profile is treated as an independent, self-contained declaration of interests, allowing users to maintain multiple profiles for different use cases without cross-contamination. This design directly addresses issues identified in the literature related to shared devices, mixed contexts, and misattributed behavioral signals.

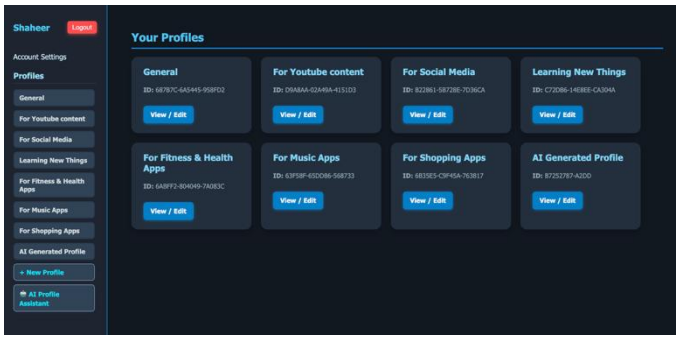


Fig. 8. Dashboard page.

*5) Account-Level Controls*

In addition to profile-level management, the demo includes a basic account settings interface, shown in Fig. 9, enabling users to update credentials or remove their account entirely. This functionality is included to demonstrate user autonomy and data lifecycle control, consistent with privacy-by-design principles.

Account deletion removes all associated profiles, reinforcing the notion that preference data remains user-owned and revocable rather than persistently retained by the system.

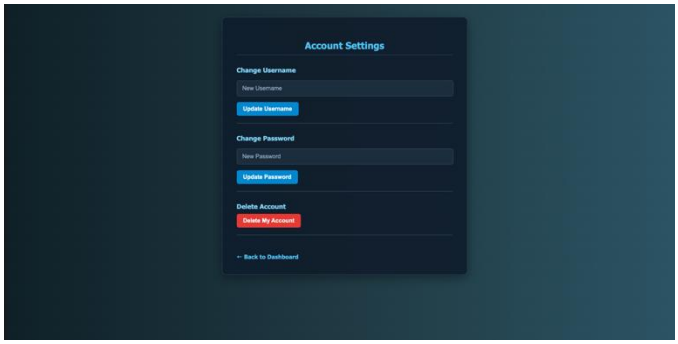


Fig. 9. Account settings page.

*6) Explicit Preference Profile Creation*

The creation of a new preference profile is the core interaction through which the framework's principles are operationalized. As illustrated in Fig. 10, users define preferences through a structured category hierarchy, explicitly indicating positive (Like), negative (Dislike), or neutral interest (Not Decided) states. Categories can be very comprehensive.

This explicit declaration mechanism replaces behavioral inference entirely. No assumptions are made based on past activity, accidental interactions, or temporal proximity. The hierarchical structure supports both coarse-grained and fine-grained preference expression while remaining interpretable to both users and consuming platforms.

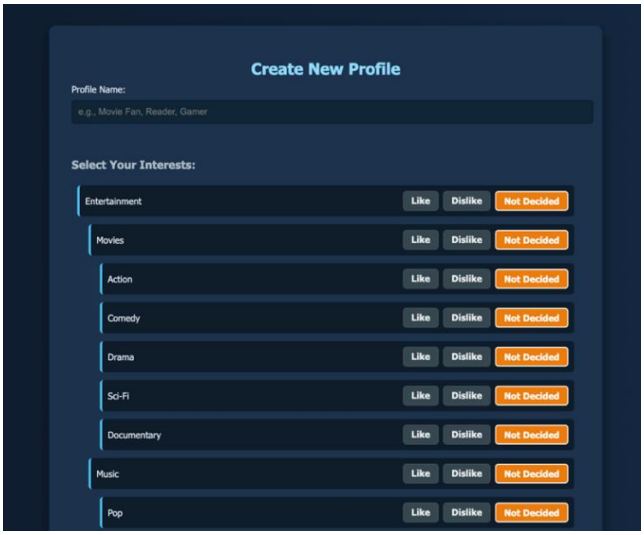


Fig. 10. New profile creation page.

*7) Profile Review, Modification, and Export*

Once created, preference profiles can be reviewed, modified, or deleted through a dedicated profile management interface, shown in Fig. 11. Preferences are not static; users may revise them as interests evolve or contexts change.

Crucially, profiles can be exported in a standardized JSON format that preserves the full preference structure and semantics. This export functionality demonstrates how platforms can consume high-quality preference data without engaging in surveillance, fingerprinting, or behavioral tracking. The JSON format is used as an example standard that can be established for this framework.

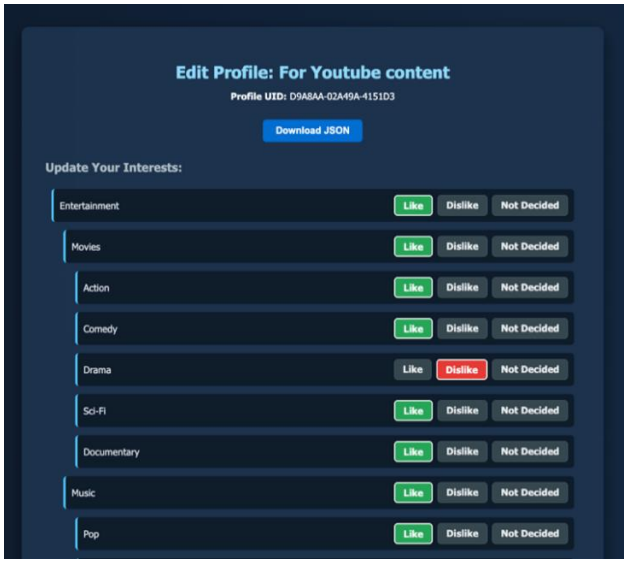


Fig. 11. Profile management page.

*8) Assisted Preference Authoring via Natural Language Input*

To reduce user effort and improve accessibility, the demo includes an optional AI-assisted preference authoring interface, illustrated in Fig. 12. This component allows users to describe their interests in natural language, which are then translated into structured preference categories. The performance of this depends on the AI model(s) used by either the individual or the organization, as explained in the above subsection B. Core Features.

This feature is intended as a usability enhancement rather than a replacement for user control. Users remain free to review, modify, or reject any generated preferences before saving the profile. The inclusion of this panel demonstrates how assistive technologies can support explicit preference declaration without reverting to opaque behavioral inference.

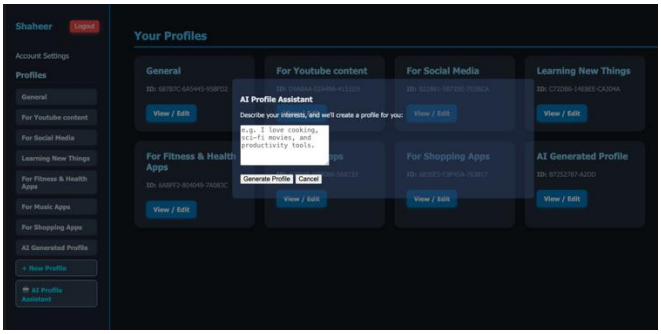


Fig. 12. AI profile creation panel.

*9) Preference Sharing and Third-Party Access*

Finally, the demo includes a minimal interface through which external entities can retrieve a preference profile using a user-provided identifier. As shown in Fig. 13, this interaction models how platforms may request and consume preference data without requiring user accounts, cookies, or tracking identifiers.

This mechanism highlights the framework's platform-agnostic and interoperable design, where preference data is shared deliberately and contextually rather than collected implicitly.

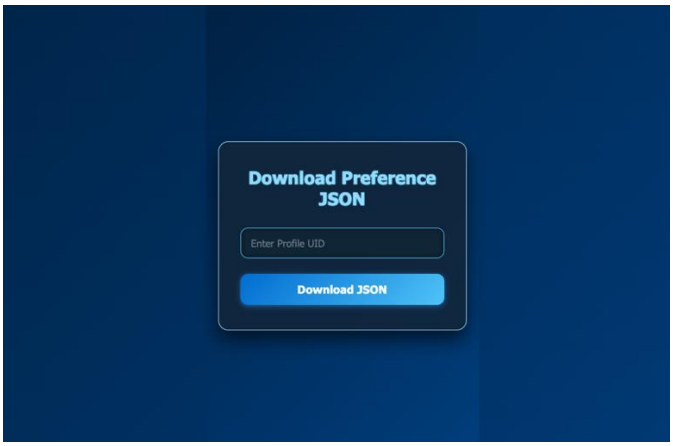


Fig. 13. Downloader page.

*10) Scope and Limitations of the Demonstration*

It is important to emphasize that this demo website is not intended as a complete system, performance benchmark, or commercial platform. Its role is illustrative: to demonstrate feasibility, interaction flow, and conceptual alignment with the proposed framework.

Advanced concerns such as large-scale deployment, governance models, standardization bodies, security hardening, and integration with production recommender systems are intentionally left out of scope and are identified as directions for future work.

### *E. Benefits of the Proposed Solution*

The system proposed here, an opt-in and user-driven Ad/Content Preference Profile (data sharing) platform, introduces paradigm-shifting benefits to the entire digital environment. This is different from invasive ad-supported models that are now in existence and is privacy-preserving without sacrificing relevance. It's a real win-win for all stakeholders in digital marketing.

#### *1) For Users: Autonomy, Privacy, and Relevance, Without Surveillance*

For most people on the internet, such a platform model truly gives freedom. Rather than having to be passively followed everywhere on the entire Internet, users would have control over what types of data, information, content and advertisements they want to have displayed. It is relevant content/ads delivery with permission rather than at the sacrifice of privacy protection.

Users can create personalized profiles for different sites, such as different job interests for LinkedIn, different genres of movies and television for YouTube/Streaming platforms, or type of products for Amazon/Shopping. It's less intrusive and on purpose. The days have gone by when users get shown an ad because they typed in one stray search term or because someone was on their computer. Users decide, Users select, Users provide. It's advertising at user's consent.

#### *2) For Businesses: Cleaner Data, More Accurate Targeting, Greater Brand Trust*

This model is revolutionary when looked at from a business perspective. Current ad tech is rooted mostly on inferred data compiled predictively and inaccurately. The result is a clutter of false positives and poorly targeted postings that users actively avoid.

Having openly stated user-generated interest profiles, brands receive high-quality, precise, and volunteered data. Not only is this improved targeting, but it also rebuilds trust within an industry where consumers are becoming cynical towards Big Tech. There is no risk in taking a leap of faith on shady third-party trackers or complicated behavioral inference engines. The user data is clean, relevant, and ethically provided.

On top of this, companies can use this paradigm as a marketing theme, "We value your Data and only show you something that you want to see". This is positive branding when concern for data privacy is on the rise.

#### *3) For Regulators: Compliance by Design*

There are laws on the books to protect users, like GDPR and CCPA, but with modern tech, compliance is more about damage control than damage prevention. The proposed conceptual platform reverses that. Ethics are built-in here. No loopholes, no dark patterns, no fine print. The users themselves create their Profiles and willingly share their Interest data with the Apps, companies, and organizations.

Every ad or content displayed through this system is backed by consent on behalf of the users, explicitly given through their preference profile. There is no need for regulators to chase companies for breaches as the consumer is in charge in the first instance. It's not damage control, it's damage prevention.

#### *4) For Developers: Simplicity, Flexibility, and Future-Proof Design*

Today developers must deal with complicated SDKs, notice banners for privacy, cookie banners, and third-party data providers. This Model answers it as one open standard. Create a preference profile, add preferred categories, share the ID of that profile, and finally get shown accurate content, no background tracking or hidden algorithms required.

This concept makes it easier to create ethically sound, lightweight, and privacy-oriented apps and still allows for monetization through advertising or personalization. As an open-source platform, developers can contribute, adapt, and expand on the platform as required by use cases.

#### *5) For Data Scientists: Cleaner Data, Minimum Preprocessing, and Consent-based*

The proposed framework offers data scientists a clean, structured, and consent-driven source of user preference data, removing much of the noise, ambiguity, and ethical burden associated with current tracking-based methods. Unlike opaque behavioral inference systems, where interests are derived from cluttered signals like accidental/incidental clicks or shared devices, this approach provides explicit, user-declared profiles that can significantly improve model accuracy and personalization outcomes. It eliminates guesswork in distinguishing actual intent from outliers, allowing data teams to work with high-integrity datasets that are legally and ethically sound. This not only enhances recommendation engine quality and ad targeting algorithm quality, but also facilitates better governance and auditing for data sets.

For machine learning pipelines, it means faster training cycles, reduced labeling overhead, and clearer feedback loops, while avoiding the risks associated with overfitting on biased or improperly collected data. By switching from invasive data mining to consensual data sharing, this framework creates a more sustainable and transparent data ecosystem where scientific rigor and user autonomy can coexist.

#### *6) Overall Novelty and Conceptual Leap*

The most significant aspect of this proposal is that none of the existing research or practice has attempted to address this issue from this angle. The most common proposed remedies and conversations are damage control, minimizing harm inflicted by existing tracking methods, or implementing opt-out functionality afterwards. This proposal is proactive rather than reactive.

Instead of trying to fix the broken system of surveillance capitalism, it offers an alternative: decentralized, consent-based, user-governed, and future-enabled system. It doesn't just

minimize harm; it renders tracking obsolete. This is the true innovation that can be enhanced through further research.

### F. Conclusive Comparison Between Current System and Proposed Solution

#### 1) Current System: Invasive by Default

Today's online economy rests on an idea that has become hardened: when you're getting it for no charge, you are the product. Sites and apps are scooping up everything, what you search, click, watch, stream, and even mouse-hover. The users didn't give any permission to get tracked everywhere online. Most don't even know what's being monitored or are too busy to stop it.

Even GDPR and CCPA, as excellent as they are, have been disadvantageous in certain situation as concluded through the literature review. The cookie notices are drafted to mislead, rather than reveal, privacy settings are buried behind a chain of multiple menus, and consent is implied rather than actual. Compliance is on study and not in spirit.

To make things worse, profiling is inaccurate. A parent using their child's iPad or a one-time Google search can skew their entire digital identity data. Ad relevance often feels like guesswork, and users feel more intruded upon than ever.

#### 2) Proposed System: Consent-Driven by Design

By contrast, the platform suggested here eliminates surveillance. People create interest profiles for themselves, directly and by conscious choice. Users decide what every app learns about them and can toggle or shift tastes at will. There is no guessing. There is no spying. There is no tracking across unrelated platforms.

One may have a variety of preferences for different activities such as work, entertainment, shopping, etc. and map them with the respective apps with a simple identifier like a username. This preference profile will only become shareable when desired by the user.

*3) No tracking. No third-party cookies. No behavioral inference.*

This approach solves many of the core issues with current user data exploitation systems:

- No more unwanted advertising or irrelevant ads.
- No more misattributed activity from shared devices or one-time search data.
- No more hidden profiling or silent consent.
- No more complex legal back-and-forth.

#### 4) Privacy Becomes the Default, Not the Afterthought

Where other systems ask for forgiveness or hide behind technobabble (incomprehensible technical jargon), the solution presented here asks for permission. It makes an online community where privacy is not just kept intact, it is engraved as an unbreakable ideal.

Companies still get data, but better data, and that data will voluntarily come from users who know what they are sharing with the Company or app. Ads and content are still personalized, at real user intent rather than usual tracking or guesswork. And users can still use services for free, without having that uneasy feeling of being watched.

This framework is not an extension to what already exists, but a rethink on how relevance and privacy might best be aligned.

## VI. Challenges and Limitations

For all its potential to revolutionize consumer and online advertising privacy, mandating a preference-governed, consumer-led Preference Data sharing system is not without its hurdles. As with every shift away from established systems, it must overcome a multi-faceted environment of opposition, logistics obstacles, as well as an education shortfall. The significant hurdles to mention are as follows:

Resistance by Big Tech Giants: The largest obstacle is entrenched interest by the Giants. They now monetize information accumulated with invasive surveillance methods. The entire business model is based on behind-the-scenes profiling and data collection instead of having users provide data on an opt-in arrangement. A move toward a consent-based framework would entail some loss of profit and control. So, Google, Meta, and Amazon businesses would not welcome this with open arms naturally. It is not just tech, they would need to rethink their data stack, ad framework, and platform dynamics. This is a big order.

Risk of Abuse - Spoof or Manipulated Profile: Granting users control over what they are interested in presents one loophole risk: malicious players will game the system. Companies or spammers may also create spoof Preference Profiles to poison ad targeting or reap benefits assigned to designated segments. In order for a system to successfully function, it would require verification tools or behavioral checks without reviving intrusive tracking, a tricky thing indeed.

Monetization and Incentivization Strategy: When advertising is served based on users defined criteria, advertisers will naturally ask: Where's the optimization? Currently, machine learning optimizes ad delivery in real-time on inferred behavior. With a system based on user-defined profiles, there is some loss of granularity. This is a bigger question, how do platforms monetize such a system, and how do advertisers measure performance? There will have to be new cost-per-interest models, new measures of efficiency of conversion, and perhaps even incentivization by users, paying users to opt-in to have preferences refreshed or to answer honestly.

The two decades and billions invested in the surveillance-based data collection and tracking economy have no click-of-a-switch expectation of a flash transition to an alternative system. Such systems are built on top of heavyweight APIs, SDKs, real-time bidding systems, and price-per-data-point selling by third-party data brokers, all delicately tuned to behavioral targeting. Transitioning from there to consent-based would have gradual, perhaps even agonizing transition. Hybrid models will have a temporary relief, but these can become backdoors for inserting covert tracking unless regulated hard.

Failure to Educate General Users: In the real-world situation, most users don't even have a clue about what's spying on them online, let alone how to manage Preferences Profiles. Assuming all users will develop rich preference profiles is an expectation of a sort of digital literacy that no one possesses. In order for this to take off as a success, there needs to exist a robust layer of user education, tutorials, onboarding tours, perhaps even gamified experiences to help users manage interests. If it's difficult or an enormous effort, users will revert back to "what works" and return control to tracking systems.

Interoperability and Standardization: In order for such a system to achieve size, it cannot possibly operate on a single app or platform only. It must be made into an open standard, something that any app, game, or website can integrate with seamlessly. To make that happen will require open protocols, interoperability arrangements, and likely regulatory promotion. Without standards, the idea risks becoming fragmented, one platform implementing it and one that doesn't, and users would have to navigate several systems. Standardization may also drive healthy competition and ad tech innovation that is respectful to users' privacy.

While there is immense potential in having a preference-based economy, it will not be a cakewalk. It goes against the way business is now carried out online, and it's not going to come with open arms. But that's precisely what makes it worth trying. Privacy shouldn't have to be the price you pay for convenience. The entire industry has an honest opportunity now to reimagine online relevant content delivery and advertising, not as a method by which to sell more, but as an operating framework based on respect, on consent, and on true relevance. Cooperation, transparency, and most fundamentally, an agreement to get off the road to least resistance will take it.

## VII. Implementation Roadmap, Recommendations, and Future Work

As learned from the literature review and the survey taken, making an open, opt-in ad economy is not a vision, it's a mandate. The people are actively looking for a solution rather than waiting for Big Tech to comply with Data Protection Acts and regulations. In the previous sections, this research's perspective of such a platform was presented which can successfully fulfil the users' requirements and not cause an issue for the companies. Stating again, it's a win-win solution for all digital entities involved. The users get the control over their data whereas the companies get the data by peacefully asking for it from the user. This section maps out a pragmatic way back to restoring control for users without shattering the "data tracking" supported web.

### *A. Cross-Industry Alliance for Ethical Advertising*

No single actor can compel this shift. What one needs is an inter-industry group comprising stakeholders with a genuine interest in building trust on the digital space, privacy-oriented developers, ethical start-ups, digital liberties groups, open-source groups, as well as progressive advertisers. These players, as a group, can:

- Set transparent ethical standards for consent-driven data sharing and personalization.
- Advocate for privacy-by-design principles.
- Promote adoption of open standards and APIs for sharing content/ads preferences data.

Such a platform should have no linkage with Big Tech and rather should be funded by neutral nonprofits, open-source devices, and local platforms committed to users' autonomy.

### *B. MVP Idea: Open-Source Website to Build Interest Profiles*

Phase one is an MVP: A fundamental open-source website on which users may create and manage ad and content preference profiles. The key features can include:

- Easy-to-use UI for selecting and categorizing interests.
- Multiple profile creation (e.g., Shopping, Music, Gaming, Social Media).
- Option to generate a unique alphanumeric Profile ID for each profile.
- Support to refresh interests at any time with real-time syncing features.

This MVP is an operational utility as well as a proof-of-concept. By open-sourcing the platform/app, there is transparency, as well as promotion of collaboration with the global developer and Privacy community.

### *C. API Standard for App Integration*

For this to scale, an open, standardized API is required on which apps, sites, and platforms can ask this profile for users' preferences instead of making their own in secret by tracking behind users' backs. The most fundamental functionality offered by the API must be:

- Requesting a specific profile by ID (with consent).
- Syncing updated interests securely for linked profile.
- Allowing apps to offer category-specific personalization (e.g., only show Music and Comedy categories from the user profile).
- Full revocation of access from the user's end at any time.

This provides people with control and encourages platforms to not collect data without reason.

### *D. Partnerships: Start Small, Grow Gradually*

Though Big Tech might not sign up on day one, smaller developers and privacy-oriented companies are great partners upon whom to test and prove out the system. They include:

- Indie game developers (who currently rely on ad SDKs that disrespect user privacy).
- Ethical data privacy friendly platforms such as Proton, Mastodon, or Lemmy.
- Music or podcast applications with interest in community-based personalization.
- Content filtering tools or browser extensions that want to build in clean advertising alternatives.

Such partners would provide aid in credibility building, interoperability verification, and proving real-world applicability of preference-based personalization without surveillance.

### E. Governance: User-Led, Not Company-Controlled

To avoid it becoming "just another platform," its governing framework should have collective accountability, transparency, and control at its core. One possibility is a user-led or consortium, cooperative where users, contributors, and even occasional users participate in voting on:

- New features and categories.
- API access rules.
- Third-party app partnerships.
- Security protocols and breach responses.

The idea is not to become locked into control by a few individuals, but to create a decentralized, organized system where the interest of the user is not an afterthought, but the priority.

## VIII. Conclusion: A Future Beyond Tracking

The conclusion drawn from literature review was that the current practices in the Data Integrity and Privacy space are trying their best to protect and equip Users with some control, yet they have certain definite flaws that cannot be solved by making more policies. As seen in the survey results, the users were willing to adapt a new solution rather than depending on policies. No doubt, Data Protection Acts are doing their best, but it's just not enough, as companies can find loopholes and workarounds every time. This creates privacy issues for users and their personal data as discussed extensively in this study.

Imagine a world where you only see what you truly value, with none of what you're doing online tracked and recorded. That's the vision here: Relevance without surveillance. This is not just another idea. It is a shift in mindset. Instead of having businesses follow users from app to website to collect their data in the name of providing "relevant" content, what if users had a choice and were able to communicate to companies what they truly want to see?

At its essence is a basic, yet powerful principle: have users develop ad and content preference profiles like they create social media profiles. Instead of shrouded by veil-of-uncertainty algorithms, users would tell different platforms (or share their data) explicitly and openly about what kind of ads or what kind of content interests them. Companies would not have to track anything. No more guessing, no more eavesdropping, no more false data, no more embarrassing advertising for something your friend checked out on your device. Just preferences, by choice.

This is not only an ethical principle, it's pragmatic. It's a win-win. Businesses still get data to sell advertising and keep services free. But they get clean data, direct feedback rather than piecing together conclusions through intrusive tracking. Consumers, on the other hand, have privacy and have more control over what they watch, write, buy, etc. Both consumers and businesses are better off.

Of course, this is just a simple framework. But it's a solution to such a problem that hasn't had any so far. A platform like this can easily be built. The technology is there. It can start as an open-source online platform/application. It can become an open protocol, an open standard. Just like demo website in this study, it is the simple website but has the ability to share users preference data with any organization. Developers, companies, and people can get behind such a thing. But it would require people, researchers, designers, coders, policymakers to mold it. The framework introduced in this research will certainly get data scientists, privacy researchers, and tech companies to one table. Although, it might look like a simple solution, but that is the purpose. It won't be beneficial to make an overly complex solution that users hesitate to use, as proven by the Survey Results. The simplicity itself will get everyone onboard.

Let's reimagine how we do relevance online. Let's create systems that honor users rather than take advantage of them. Instead of teaching algorithms to guess what users want, let users simply tell them. If we can get organized, do some additional research, and start prototyping this vision on a bigger scale than what this work presents in the previous sections, people may be able to guide the internet in a path that serves all of us ethically.

## References

[1] O. Rafieian and H. Yoganarasimhan, "Targeting and Privacy in Mobile Advertising," Marketing Science, vol. 40, no. 2, Dec. 2020, doi: https://doi.org/10.1287/mksc.2020.1235.

[2] J. L. Hayes, N. H. Brinson, G. J. Bott, and C. M. Moeller, "The Influence of Consumer–Brand Relationship on the Personalized Advertising Privacy Calculus in Social Media," Journal of Interactive Marketing, vol. 55, no. 1, pp. 16–30, Aug. 2021.

[3] L. F. Lina and A. Setiyanto, "Privacy Concerns in Personalized Advertising Effectiveness on Social Media," SRIWIJAYA INTERNATIONAL JOURNAL OF DYNAMIC ECONOMICS AND BUSINESS, vol. 5, no. 2, pp. 147–156, Aug. 2021, doi: https://doi.org/10.29259/sijdeb.v1i2.147-156.

[4] B. Zarouali, T. Dobber, G. De Pauw, and C. de Vreese, "Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media," Communication Research, vol. 49, no. 8, p. 009365022096196, Oct. 2020, doi: https://doi.org/10.1177/0093650220961965.

[5] S. Zhang, Y. Feng, Y. Yao, L. F. Cranor, and N. Sadeh, "How Usable Are iOS App Privacy Labels?," Proceedings on Privacy Enhancing Technologies, vol. 2022, no. 4, pp. 204–228, Oct. 2022, doi: https://doi.org/10.56553/popets-2022-0106.

[6] M. Kretschmer, J. Pennekamp, and K. Wehrle, "Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web," ACM Transactions on the Web, vol. 15, no. 4, pp. 1–42, Jul. 2021, doi: https://doi.org/10.1145/3466722.

[7] K. Birch, D. Cochrane, and C. Ward, "Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech," Big Data & Society, vol. 8, no. 1, p. 205395172110173, Jan. 2021, doi: https://doi.org/10.1177/20539517211017308.

[8] K. Varnali, "Online behavioral advertising: An integrative review," Journal of Marketing Communications, vol. 27, no. 1, pp. 1–22, Jun. 2019, doi: https://doi.org/10.1080/13527266.2019.1630664.

[9] R. Amos, G. Acar, E. Lucherini, M. Kshirsagar, A. Narayanan, and J. Mayer, "Privacy Policies over Time: Curation and Analysis of a Million-Document Dataset," Proceedings of the Web Conference 2021, pp. 2165–2176, Apr. 2021, doi: https://doi.org/10.1145/3442381.3450048.

[10] S. C. Boerman, S. Kruikemeier, and F. J. Zuiderveen Borgesius, "Exploring Motivations for Online Privacy Protection Behavior: Insights From Panel Data," Communication Research, vol. 48, no. 7, p.

009365021880091, Oct. 2018, doi: https://doi.org/10.1177/0093650218800915.

[11] S. Quach, P. Thaichon, K. D. Martin, S. Weaven, and R. W. Palmatier, "Digital technologies: Tensions in Privacy and Data," Journal of the Academy of Marketing Science, vol. 50, no. 1, pp. 1299–1323, 2022, doi: https://doi.org/10.1007/s11747-022-00845-y.

[12] Z. Gharibshah and X. Zhu, "User Response Prediction in Online Advertising," ACM Computing Surveys, vol. 54, no. 3, pp. 1–43, Jun. 2021, doi: https://doi.org/10.1145/3446662.

[13] E. Papadogiannakis, P. Papadopoulos, N. Kourtellis, and E. P. Markatos, "User Tracking in the Post-cookie Era: How Websites Bypass GDPR Consent to Track Users," arXiv:2102.08779 [cs], Feb. 2021, doi: https://doi.org/10.1145/3442381.3450056.

[14] M. Mehrnezhad, K. Coopamootoo, and E. Toreini, "How Can and Would People Protect From Online Tracking?," Proceedings on Privacy Enhancing Technologies, vol. 2022, no. 1, pp. 105–125, Nov. 2021, doi: https://doi.org/10.2478/popets-2022-0006.

[15] S. Aiolfi, S. Bellini, and D. Pellegrini, "Data-driven digital advertising: benefits and risks of online behavioral advertising," International Journal of Retail & Distribution Management, vol. 49, no. 7, pp. 1089–1110, May 2021, doi: https://doi.org/10.1108/ijrdm-10-2020-0410.

[16] M. Tahaei, A. Frik, and K. Vaniea, "Deciding on Personalized Ads: Nudging Developers About User Privacy," Symposium On Usable Privacy and Security, pp. 573–596, May 2021, doi: https://doi.org/10.7488/ds/3045.

[17] D. Geradin, T. Karanikioti, and D. Katsifis, "GDPR Myopia: how a well-intended regulation ended up favouring large online platforms - the case of ad tech," European Competition Journal, vol. 17, no. 1, pp. 1–46, Dec. 2020, doi: https://doi.org/10.1080/17441056.2020.1848059.

[18] K. Kollnig, A. Shuba, M. Van Kleek, R. Binns, and N. Shadbolt, "Goodbye Tracking? Impact of iOS App Tracking Transparency and Privacy Labels," arXiv:2204.03556 [cs], May 2022, doi: https://doi.org/10.1145/3531146.3533116.

[19] K. Kollnig et al., "A Fait Accompli? An Empirical Study into the Absence of Consent to Third-Party Tracking in Android Apps," arXiv (Cornell University), Jun. 2021, doi: https://doi.org/10.48550/arxiv.2106.09407.

[20] M. Büchi, E. Fosch-Villaronga, C. Lutz, A. Tamò-Larrieux, and S. Velidi, "Making sense of algorithmic profiling: user perceptions on Facebook," Information, Communication & Society, vol. 26, no. 4, pp. 1–17, Oct. 2021, doi: https://doi.org/10.1080/1369118x.2021.1989011.

[21] M. Veale and F. Zuiderveen Borgesius, "Adtech and Real-Time Bidding under European Data Protection Law," German Law Journal, vol. 23, no. 2, pp. 226–256, Mar. 2022, doi: https://doi.org/10.1017/glj.2022.18.

[22] Konrad Kollnig, A. Shuba, R. Binns, M. V. Kleek, and N. Shadbolt, "Are iPhones Really Better for Privacy? Comparative Study of iOS and Android Apps.," Privacy Enhancing Technologies, Sep. 2021.

[23] Dr. J. Strycharz, Dr. E. Smit, Dr. N. Helberger, and Dr. Guda van Noort, "No to cookies: Empowering impact of technical and legal knowledge on rejecting tracking cookies," Computers in Human Behavior, vol. 120, p. 106750, Feb. 2021, doi: https://doi.org/10.1016/j.chb.2021.106750.

[24] J. Jueckstock et al., "Towards Realistic and ReproducibleWeb Crawl Measurements," pp. 80–91, Apr. 2021, doi: https://doi.org/10.1145/3442381.3450050.

[25] K. Mellet and T. Beauvisage, "Cookie monsters. Anatomy of a digital market infrastructure," Routledge eBooks, pp. 14–33, Apr. 2021, doi: https://doi.org/10.4324/9781003130154-2.

[26] S. G. Goldberg, G. A. Johnson, and S. K. Shriver, "Regulating Privacy Online: An Economic Evaluation of the GDPR," American Economic Journal: Economic Policy, vol. 16, no. 1, pp. 325–358, Feb. 2024, doi: https://doi.org/10.1257/pol.20210309.

[27] G. Aridor, Y.-K. Che, and T. Salz, "The Effect of Privacy Regulation on the Data Industry," Proceedings of the 22nd ACM Conference on Economics and Computation, Jul. 2021, doi: https://doi.org/10.1145/3465456.3467543.

[28] Y. Feng, Y. Yao, and N. Sadeh, "A Design Space for Privacy Choices: Towards Meaningful Privacy Control in the Internet of Things," Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, May 2021, doi: https://doi.org/10.1145/3411764.3445148.

[29] C. Santos, A. Rossi, L. S. Chamorro, K. Bongard-Blanchy, and Ruba Abu-Salma, "Cookie Banners, What's the Purpose?: Analyzing Cookie Banner Text Through a Legal Lens," Workshop on Privacy in the Electronic Society, pp. 187–194, Nov. 2021.

[30] Bollinger, D., Kubicek, K., Cotrini, C., & Basin, D. (2022). Automating cookie consent and {GDPR} violation detection. In 31st USENIX Security Symposium (USENIX Security 22) (pp. 2893-2910).

[31] Ke Xin Ch'ng, A. Y. Lee, Wou Onn Choo, and Pei Boon Ooi, "Students' and Parents' Online Behavior and Exposure towards Cyberbullying," Oct. 2021, doi: https://doi.org/10.1109/icriis53035.2021.9617112.

[32] C. Mcclain, M. Faverio, M. Anderson, and E. Park, "How Americans View Data Privacy The role of technology companies, AI and regulation -plus personal experiences with data breaches, passwords, cybersecurity and privacy policies FOR MEDIA OR OTHER INQUIRIES," vol. 18, p. 2023, Available: https://www.pewresearch.org/wp-content/uploads/sites/20/2023/10/PI_2023.10.18_Data-Privacy_FINAL.pdf

[33] "The Privacy Advantage: Building Trust in a Digital World." Available: https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/cisco-privacy-benchmark-study-2025.pdf