From Regulation to Relevance: Integrating User Values into Privacy Policy Scoring

Brian Kim and Suzanne Barber
The Center for Identity
The University of Texas at Austin
Austin, Texas, United States of America (USA)
e-mail: briankim31415@gmail.com, sbarber@identity.utexas.edu

***Abstract*—Despite increasing regulatory efforts to protect user data, such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), privacy policies—the main way users encounter these protections—remain difficult to understand and largely ineffective in guiding informed consent. Existing automated tools for evaluating these policies prioritize legal compliance but often overlook what users personally value in data privacy. This paper proposes a user-centered approach that integrates individual privacy values into policy scoring systems. A survey capturing user comfort levels with various types of Personally Identifiable Information (PII) reveals wide variability in privacy concerns. These insights are embedded into the PrivacyCheck™ tool to enable personalized policy evaluations. The results show that incorporating user values can significantly shift policy scores, especially in areas related to user control, highlighting a gap between regulatory standards and real user priorities. This work moves toward privacy tools that better reflect what users actually care about, supporting more meaningful and transparent data governance.**

Keywords–Privacy policy evaluation; User privacy values; Personalized scoring; Personally Identifiable Information.

I. Introduction

Digital platforms increasingly collect, store, and process personal data, making the protection of PII a critical public concern. Governments and regulatory bodies have responded with major privacy regulations, such as the European Union's GDPR, the CCPA, and its amendment, the California Privacy Rights Act (CPRA).

While such regulations shape data protection norms, users experience them primarily through privacy policies—the official documents describing how websites and applications collect, use, and share personal data. Ideally, policies should empower informed decisions, but studies show they are long, written at a college reading level, and filled with legal or technical jargon, making them inaccessible to the general public [1]. As a result, most users skip them or misunderstand their content, leading to uninformed consent and reduced transparency in data practices [2][3].

To address this shortcoming, automated tools interpret and evaluate privacy policies using machine learning [4]–[6]. PrivacyCheck™ [7] rates policies against regulatory frameworks, such as the GDPR and Fair Information Practice Principles (FIPPs), but these tools often apply uniform criteria and overlook variation in what users value. As a result, two users with different sensitivities can receive the same policy score, even when the policy aligns well with one user's priorities and poorly with another's.

We argue that privacy policy evaluation should go beyond regulatory compliance to reflect personal values. We propose a user-centered approach that integrates individual privacy preferences into automated policy scoring systems. We conducted a user study to assess comfort with sharing different types of PII and privacy-related practices. Results show general caution—especially for possession-based information—but substantial variation across individuals, highlighting the need for personalized evaluation frameworks. This variation suggests that a single "good" score can mask trade-offs across factors, such as data retention and third-party sharing, that users weigh differently.

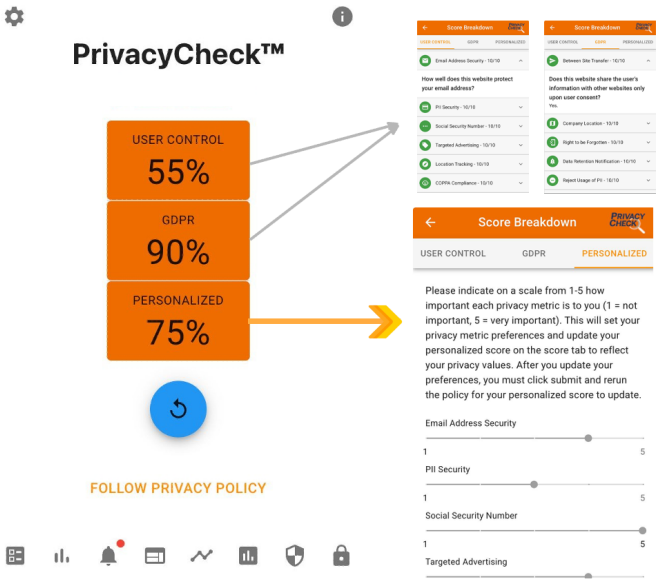


Figure 1. An example of PrivacyCheck™ scoring with the personalized scoring mechanism.

We map the survey insights onto the PrivacyCheck™ scoring model, enabling personalization based on aggregated or individual values (see Figure 1). Incorporating user values shifts policy scores—particularly for user control—revealing gaps between regulatory benchmarks and user priorities. By embedding user values into policy assessments, this work bridges legal compliance and meaningful user understanding.

The remainder of this paper is organized as follows. Section 2 reviews related work, Section 3 describes the survey, Section 4 details the value-aligned scoring and evaluation, and Section 5 concludes.

II. Related Work

A. Tools for Interpreting Privacy Policies

Automated tools analyze privacy policies using Natural Language Processing (NLP) and machine learning to improve transparency, regulatory compliance, and user understanding. Early systems, such as Privee [5], extract key practices from policy text. Polisis [4] uses deep learning with hierarchical attention to classify policy content and present it through an interactive user interface.

More recent tools include PrivacyGuide [6] and PolicyChecker [8]. They assess GDPR compliance using NLP and supervised learning

but often overlook individual user values. The Value-Centered Privacy Assistant (VcPA) [9][10] addresses this gap by helping users make app decisions aligned with personal value profiles derived from survey data.

B. User Values Regarding Privacy Policies

Surveys and experiments have examined user attitudes toward privacy policies and privacy-enhancing technologies. Ibdah et al. [11] show that despite concern about data privacy, participants often avoid policies due to complexity, length, and helplessness, prioritizing convenience over informed consent. Choi et al. [12] and Ebbers et al. [13] report similar themes in smart speakers and digital assistants, emphasizing transparency and control with features, such as customizable data collection and meaningful consent. These studies collectively suggest that usability and perceived control, not just formal compliance, shape whether privacy tools influence real behavior.

Wang & Li [14] propose a machine learning framework to infer privacy preferences from demographic and behavioral data, enabling personalized recommendations with limited user input. Building on this foundation, we directly integrate user-reported comfort levels with different types of PII into an automated policy evaluation system, aligning assessments with user values.

C. Privacy Policy Rating Systems PrivacyCheck™

We focus on PrivacyCheck™ [7], a machine learning–based privacy policy rating system. Implemented as a Chrome browser extension, PrivacyCheck™ identifies and scores key privacy practices from policy webpages, providing structured assessments aligned with frameworks, such as the GDPR and FIPPs.

Originally launched in 2015, the initial version of PrivacyCheck™ detected policy content across ten core dimensions, including PII collection (e.g., email, Social Security Number (SSN)), data aggregation, and law enforcement sharing. It used keyword detection and parsing to assign categorical risk scores reflecting GDPR and Federal Trade Commission (FTC) standards. Later iterations added scalability, version tracking, and user engagement.

In its current implementation, PrivacyCheck™ rates privacy policies across 20 factors derived from regulatory guidelines [15], as shown in Table I. Factors span disclosures, such as data usage, retention, third-party transfers, user access rights, and consent mechanisms. Each factor is scored and aggregated into two overall scores:

- GDPR Score: Reflects the presence or absence of disclosures relevant to specific GDPR articles
- User Control Score: Indicates the extent to which a policy outlines mechanisms for user agency and data governance based on FIPPs

In this work, PrivacyCheck™ provides the baseline for comparing regulatory compliance to user privacy value coverage, supporting an empirical investigation of transparency and data governance in online privacy policies.

III. User Attitudes Toward the Collection and Processing of PII

A. Survey Design

To evaluate users' comfort levels with digital privacy practices, we conduct a survey to capture individual attitudes toward the collection and processing of various types of PII by websites and mobile applications. This study was approved by the Institutional Review Board (IRB). The survey is structured around six thematic categories as listed below.

The first two categories address data policies and processing practices, assessing participants' perceptions of broader organizational behaviors, such as data retention, third-party data sharing, and transparency in privacy policies. The other four categories focus on participants' attitudes toward specific types of PII. These categories are informed by foundational principles in information security and user authentication, specifically the types of data individuals have, are, do, and know. Together, these six categories were designed to comprehensively reflect real-world contexts of digital data collection and usage.

1) Data Policies (17 questions)
2) Data Processing (8 questions)
3) What You HAVE (16 questions): possession-based identifiers (e.g., device IDs)
4) What You ARE (10 questions): biometric information (e.g., fingerprints)
5) What You DO (10 questions): behavioral data (e.g., browsing history)
6) What You KNOW (20 questions): knowledge-based credentials (e.g., passwords)

The final survey included a total of 81 questions. Each question in the survey is phrased in a standardized format to ensure clarity and consistency across categories. The wording generally follows the structure: "How comfortable are you with websites or mobile apps... [question topic]." The question topic typically begins with an action verb, such as "collecting" or "sharing," depending on whether the item relates to data acquisition or dissemination, and is followed by a specific type of PII; for example, "How comfortable are you with websites or mobile apps collecting your online shopping patterns?" or "How comfortable are you with websites or mobile apps sharing your data for academic or research purposes?" This consistent phrasing helps participants easily interpret the intent of each question while maintaining focus on the specific privacy-related scenario being assessed.

Participants respond to each item using a five-point Likert scale: (1) very comfortable, (2) slightly comfortable, (3) no difference, (4) slightly uncomfortable, and (5) very uncomfortable. The selected scale enables nuanced expression and fine-grained analysis of user comfort levels, revealing both the types of PII users are most sensitive about and the intensity of that sensitivity.

To ensure the quality of the responses and detect inattentive participation, one question from each of the six main categories is deliberately duplicated in another category where the context still made the question relevant. These duplicated questions serve as internal consistency checks. A response is flagged as inconsistent if a participant answered a duplicated question pair in a way that suggested a reversal in sentiment—for example, if an answer shifted from either very uncomfortable or slightly uncomfortable to either very comfortable or slightly comfortable, or vice versa. Additionally, if a participant changed their response from an extreme stance—defined as very uncomfortable or very comfortable—to no difference, or from no difference to an extreme stance, this is also treated as an inconsistency. Submissions containing more than one of these response reversals are considered unreliable and excluded from the dataset. Using this criterion, nine participant submissions are removed.

Demographics. In the survey, we also collect participants' demographic information, including gender, age, ethnicity, occupation, and education level. Additionally, we ask five questions about participants' general attitudes toward privacy practices and their daily use of websites and applications, as listed below. This information is used to explore potential correlations between user profiles and their

TABLE I. PRIVACYCHECK™ SCORING FACTORS.

User Control Factors
How well does this website protect your email address?
How well does this website protect your credit card information and address?
How well does this website handle your Social Security Number?
Does this website use or share your PII for marketing purposes?
Does this website track or share your location?
Does this website collect PII from children under 13?
Does this website share your information with law enforcement?
Does this website notify or allow you to opt-out after changing their privacy policy?
Does this website allow you to edit or delete your information from its records?
Does this website collect or share aggregated data related to your identity or behavior?
GDPR Compliance Factors
Does this website share the user's information with other websites only upon user consent?
Does this website disclose where the company is based or where the user's PII will be processed and transferred?
Does this website support the right to be forgotten?
If they retain PII for legal purposes after a user request to be forgotten, will they inform the user?
Does this website allow the user to reject the use of their PII?
Does this website restrict the use of PII of children under the age of 16?
Does this website advise the user that their data is encrypted even while at rest?
Does this website ask for the user's informed consent before processing data?
Does this website implement all principles of data protection by design and by default?
Does this website notify the user of security breaches without undue delay?

comfort levels with digital privacy, enabling analysis of trends across different demographic groups.

1) On average, how many hours per day do you spend online or on mobile apps?
2) How often do you read each website's or mobile application's privacy policy?
3) How comfortable are you with technology?
4) Do you agree with the statement that "I feel that I get more accomplished because of technology?"
5) List 3 to 5 of your most used websites and mobile apps.

Survey participants are recruited from the undergraduate and graduate student populations in the Electrical and Computer Engineering (ECE) and Computer Science (CS) departments at the University of Texas at Austin. Participation is entirely voluntary. In total, we collect 99 valid responses. Of the participants, 70.7% identify as male, 23.2% as female, and 6.1% as other gender identity. In terms of ethnicity, 59.6% identify as Asian, 17.2% as White, and 16.2% as Hispanic or Latino. The average age of participants is 21 years.

A third of participants (33.3%) report that they never read privacy policies, while 42.4% say they rarely do. Only 24.2% indicate that they read them sometimes or often. Regarding technological proficiency, 43.4% of participants consider themselves experts, and 46.5% report an advanced level of comfort with technology. Participants also report their average daily internet usage, with responses spanning a broad range, as illustrated in Figure 2. Finally, participants list their top 3 to 5 most frequently used websites and applications. Aggregated results reveal the 20 most-used platforms, shown in Figure 2. Social media platforms dominate usage patterns, with Instagram emerging as the most frequently mentioned.

B. Survey Results

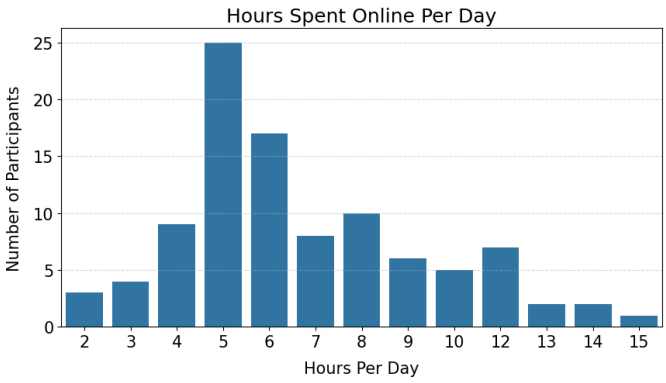
The overall mean score across all survey questions is 3.85 on a five-point Likert scale. As shown in Figure 3, questions in the "Data Policies" and "What You HAVE" categories receive higher scores, suggesting that users are generally cautious about websites and apps collecting their PII, particularly possession-based information. In contrast, scores are lower in the "What You KNOW" and "Data Processing" categories, indicating more users are comfortable with sharing knowledge-based PII and are less concerned about how their data is processed. The questions with the highest and lowest scores are shown in Table II.

We also calculate the standard deviation for each survey question to assess variability in participants' responses, as shown in Figure 4. The average standard deviation across all questions is 1.06. We observe that questions with higher mean scores—indicating lower comfort levels—tend to have lower variance, suggesting stronger consensus when participants feel uncomfortable about specific data practices or sharing certain types of information. In contrast, questions with lower mean scores—reflecting higher comfort—exhibit greater variance, indicating a wider range of opinions. This pattern reveals a general agreement on what makes users uncomfortable, while comfort tends to be more subjective and varies significantly across individuals.

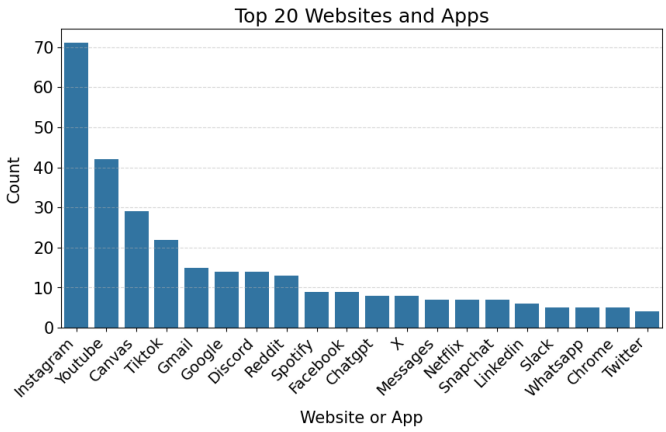
In Table III, we list questions with the highest and lowest variance. Notably, all five questions with the highest variance belong to the "What You KNOW" category, whereas the questions with the lowest standard deviations are from the "Data Policies" and "What You HAVE" categories.

Additionally, to assess the relationships between users' responses across questions, we compute the Pearson correlation coefficient (r). The results, shown in Figure 5, summarize the number of strongly correlated question pairs ($r > 0.6$) across category combinations. Few questions have negative correlations, and none are strongly negative. Most strong correlations occur within the same category, likely reflecting thematic consistency. In particular, questions in the "What You HAVE" and "What You KNOW" categories exhibit high internal correlations. Notably, the "What You KNOW" category is the only group to show strong correlations with questions from other categories, suggesting it may capture broader user attitudes that extend beyond its specific theme.

To ensure that correlation patterns are not merely the result of universally agreeable questions, a subset of questions is selected based on higher distributional variability (standard deviation greater than 1). Pearson correlation results for the refined question set are shown in Figure 6. Even with this stricter filtering, the "What You KNOW" category remains prominent, exhibiting a high number of strong correlations, especially within its own group.



(a) Distribution of hours spent online per day.



(b) Top 20 most frequently used websites and apps.

Figure 2. Demographic Information.

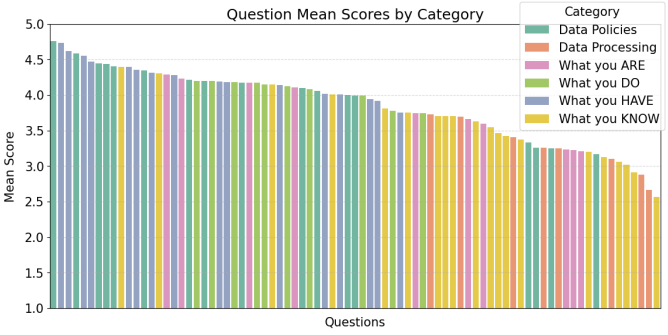


Figure 3. Distribution of question mean scores color-coded by category.

Finally, we calculate correlations between participants' average scores on high-distribution questions and their demographic information. For continuous demographic variables, we used Pearson correlation; for categorical variables, we used analysis of variance (ANOVA). None of the correlations are strong, suggesting no meaningful relationship between demographic factors and their responses to more polarizing or nuanced questions.

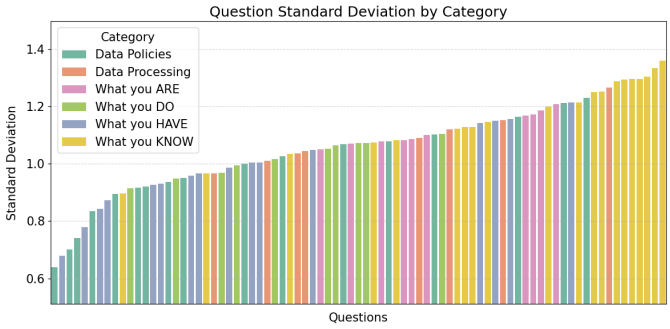


Figure 4. Distribution of question score standard deviations color-coded by category.

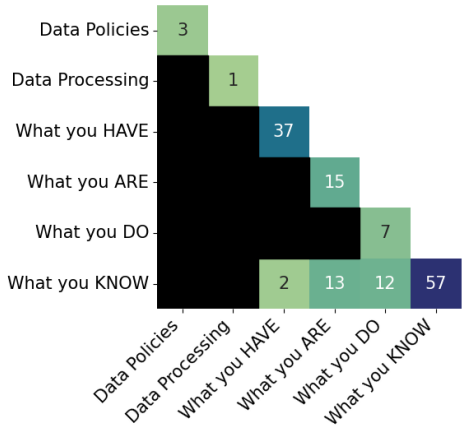


Figure 5. Count of Strongly Correlated Question Pairs across Categories.

Figure 6. Counts of Strongly Correlated High-Variance Question Pairs across Categories.

TABLE II. QUESTIONS WITH THE HIGHEST AND LOWEST MEAN SCORES.

Question	Category	Mean
... not notifying you if a data breach has occurred?	Data Policies	4.76
... having access to and/or collecting copies of your SSN?	What You HAVE	4.74
... having access to and/or collecting copies of your birth certificate?	What You HAVE	4.62
... collecting information from children under the age of 13?	Data Policies	4.59
... having access to and/or collecting copies of your passport?	What You HAVE	4.56
... collecting your email?	What You KNOW	3.02
... collecting your gender?	What You KNOW	2.91
... collecting your data for curated recommendations?	Data Processing	2.88
... collecting your data for developer bug detection and analytics?	Data Processing	2.67
... collecting your preferred language?	What You KNOW	2.57

TABLE III. QUESTIONS WITH THE HIGHEST AND LOWEST STANDARD DEVIATIONS IN SCORE DISTRIBUTIONS.

Question	Category	St.D	Mean
... collecting your ethnicity?	What You KNOW	1.36	3.13
... collecting your name?	What You KNOW	1.33	3.20
... collecting your phone number?	What You KNOW	1.30	3.46
... collecting your preferred language?	What You KNOW	1.30	2.57
... collecting your gender?	What You KNOW	1.30	2.91
... having access to and/or collecting copies of your birth certificate?	What You HAVE	0.78	4.62
... sharing your data with business partners that are not explicitly listed?	Data Policies	0.74	4.40
... not notifying you if a data breach has occurred and your information may be compromised?	Data Policies	0.70	4.76
... having access to and/or collecting copies of your SSN?	What You HAVE	0.68	4.74
... collecting information from children under the age of 13?	Data Policies	0.64	4.59

IV. PERSONALIZED PRIVACYCHECK™ WITH VALUE-ALIGNED SCORING

Using the collected data on users' attitudes toward privacy practices, we incorporate values into the privacy policy rating system—PrivacyCheck™. We map each survey question to corresponding factors and re-weight these factors based on aggregated responses to better align PrivacyCheck™ scores with users' values. We also implement a personalization feature that allows individual users to adjust factor weights and compare the resulting scores with the original regulatory scores.

A. Aligning PrivacyCheck™ Scoring with User Value

As described in subsection 2.3, PrivacyCheck™ evaluates privacy policies using two scoring dimensions: GDPR Compliance and User Control. Each dimension comprises a set of factors. For each factor, the model assigns 0/1/2 for not addressed, unclear, or clear. These raw scores are normalized to a 0–10 scale per factor and summed to produce two overall scores out of 100—one reflecting GDPR Compliance and the other User Control. The list of corresponding questions is provided in Table I.

At a high level, if $x_f \in \{0, 1, 2\}$ is the raw score for factor f, then the base score for dimension d (with factors F_d) is computed as S_d. To incorporate user values, we map survey responses to scoring factors; each survey question links to one or more criteria used in PrivacyCheck™. We compute average responses per question and aggregate them by factor to assign a user value weight, then normalize these weights within each dimension to sum to 1. For population-level analysis, v_f is the average response across participants; for deployment, v_f can be computed per user. We normalize v_f to survey weights w_f, which initialize slider values. Users adjust slider values s_f, normalized via a softmax function to user weights $\mathbf{u}$ (with u_f as the fth element). The base score sums normalized factor scores across F_d, scaling each x_f to a 0–10 range. The weighted score is

$$w_f = \frac{v_f}{\sum_{k \in F_d} v_k}, \quad S_d^{(u)} = 10 \cdot \sum_{f \in F_d} u_f \cdot \frac{x_f}{2}$$

We integrate this feature into the PrivacyCheck™ browser extension via a new "Personalization" tab, as shown in Figure 1. The interface presents adjustable sliders for each factor, initialized with survey-derived weights. Users can modify these sliders to reflect their preferences, and the values are normalized to sum to 1, allowing users to increase the influence of specific factors and receive customized scoring.

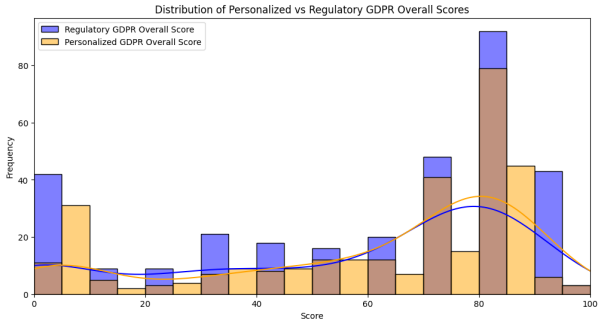


Figure 7. Distribution of personalized vs regulatory GDPR scores.

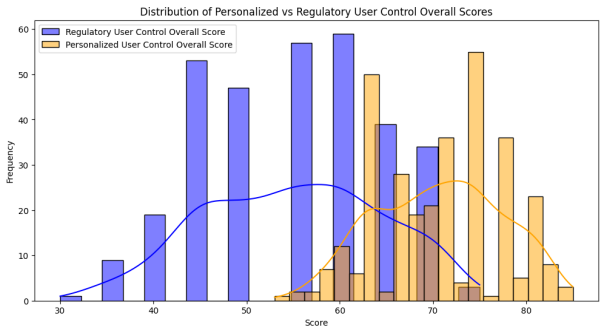


Figure 8. Distribution of personalized vs regulatory User Control scores.

B. Evaluation of the Alignment

To evaluate the value-aligned scoring system, we run the aligned PrivacyCheck™ on 321 privacy policies from the original 392-policy corpus [7], excluding invalid URLs or scoring errors. We compare value-aligned scores—generated using survey-based weights (without user-adjusted sliders)—against the original regulatory scores. Figure 7 and Figure 8 present the results for GDPR Compliance and User Control, respectively.

GDPR Compliance scores remain relatively similar under value-aligned weighting, whereas User Control scores increase. This suggests users emphasize control over how their data is collected and used, especially for sensitive "What You HAVE" information, such as Social Security Numbers and location data. This shift indicates regulatory assessments may underrepresent user priorities, and value-aligned scoring can surface concerns de-emphasized in traditional evaluations.

V. Conclusion and Discussion

This work addresses whether privacy policies reflect the values and comfort levels of the users they aim to protect. Through a user-centered survey and a personalized policy evaluation mechanism in PrivacyCheck™, we provide an empirical and technical foundation for answering that question.

Our findings show that strong, consistent discomfort with opaque data practices, particularly those involving sensitive identifiers, such as government-issued documents and breach notifications. Responses to less invasive practices, such as email or language collection, are more tolerant and varied. This pattern suggests contextual and value-driven interpretations of privacy. While regulations, such as the GDPR, provide a foundation for protecting user rights, they do not fully capture individual privacy concerns.

By integrating user preferences into PrivacyCheck™, we show that personalized scoring shifts evaluations, particularly in user control. Conventional regulatory assessments may underestimate areas that users deem critical. At the individual level, differences in priorities can shift factor emphasis and change policy rankings, supporting user-specific decision-making, such as comparing apps against stated privacy priorities. This opens the door to interfaces that let users tune scores to their values without requiring them to parse long policy text.

A. Limitations

The survey sample is limited to 99 respondents recruited from one university's ECE and CS populations, which may not represent broader user populations. Survey responses are self-reported and may not fully capture how users behave when making real privacy decisions. The mapping from 81 survey items to the 20 PrivacyCheck™ factors also introduces modeling assumptions, and aggregated factor weights may mask nuanced preferences that are not yet expressed in the survey.

The policy evaluation uses a subset of the original 392-policy corpus after excluding invalid URLs and scoring errors, so results may differ on other corpora or newly updated policy text. Finally, the current evaluation focuses on score shifts rather than downstream outcomes such as whether personalized scoring changes users' consent behavior or long-term trust. Broader demographic sampling, additional policy datasets, and behavioral validation would strengthen external validity.

To translate these findings into practice, organizations should enhance transparency by explaining, in clear language, how personal data is collected, used, and shared—especially for sensitive identifiers. It may also be helpful to provide more intuitive and granular options for users to manage privacy preferences around consent and data sharing. While legal compliance remains essential, aligning practices with user expectations can foster trust and engagement. Tools like the personalized version of PrivacyCheck™ show how user feedback can bridge regulatory intent and lived experience.

In closing, we aim to narrow the gap between what privacy policies claim and what users value. Centering user preferences supports tools and frameworks that are compliant yet intuitive and respectful. As digital ecosystems expand, integrating user values into privacy design can help build more transparent and trustworthy data governance.

References

[1] J. Tang, H. Shoemaker, A. Lerner, and E. Birrell, "Defining privacy: How users interpret technical terms in privacy policies," *Proceedings on Privacy Enhancing Technologies*, vol. 2021, no. 3, pp. 70–94, 2021. DOI: 10.2478/popets-2021-0038.

[2] C. McClain, M. Faverio, M. Anderson, and E. Park, *How Americans view data privacy*, Pew Research Center, Report 18 [retrieved: January, 2026], 2023. [Online]. Available: https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/.

[3] B. Auxier et al., *Americans' attitudes and experiences with privacy policies and laws*, Pew Research Center: Internet, Science & Tech, [retrieved: January, 2026], 2019. [Online]. Available: https://www.pewresearch.org/internet/2019/11/15/americans-attitudes-and-experiences-with-privacy-policies-and-laws/.

[4] H. Harkous et al., "Polisis: Automated analysis and presentation of privacy policies using deep learning," in *27th USENIX Security Symposium (USENIX Security 18)*, 2018, pp. 531–548.

[5] S. Zimmeck and S. M. Bellovin, "Privee: An architecture for automatically analyzing web privacy policies," in *23rd USENIX Security Symposium (USENIX Security 14)*, 2014, pp. 1–16.

[6] W. B. Tesfay, P. Hofmann, T. Nakamura, S. Kiyomoto, and J. Serna, "Privacyguide: Towards an implementation of the EU GDPR on internet privacy policy evaluation," in *Proceedings of the fourth ACM international workshop on security and privacy analytics*, 2018, pp. 15–21.

[7] R. N. Zaeem, R. L. German, and K. S. Barber, "Privacycheck: Automatic summarization of privacy policies using data mining," *ACM Transactions on Internet Technology (TOIT)*, vol. 18, no. 4, pp. 1–18, 2018.

[8] A. Xiang, W. Pei, and C. Yue, "Policychecker: Analyzing the GDPR completeness of mobile apps' privacy policies," in *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, 2023, pp. 3373–3387.

[9] S. E. Carter, "A value-centered exploration of data privacy and personalized privacy assistants," *Digital Society*, vol. 1, no. 3, p. 27, 2022.

[10] S. E. Carter et al., *In pursuit of privacy: The value-centered privacy assistant*, [retrieved: January, 2026], 2023. [Online]. Available: https : / / api . semanticscholar . org / CorpusId : 260775437.

[11] D. Ibdah, N. Lachtar, S. M. Raparthi, and A. Bacha, ""Why should I read the privacy policy, I just need the service": A study on attitudes and perceptions toward privacy policies," *IEEE Access*, vol. 9, pp. 166 465–166 487, 2021.

[12] H. Choi, J. Park, Y. R. Choi, and Y. Jung, "User preferences of privacy-enhancing attributes of a smart speaker," *International Journal of Human–Computer Interaction*, vol. 39, no. 18, pp. 3649–3662, 2023.

[13] F. Ebbers, J. Zibuschka, C. Zimmermann, and O. Hinz, "User preferences for privacy features in digital assistants," *Electronic Markets*, vol. 31, pp. 411–426, 2021.

[14] W. Wang and B. Li, "Learning personalized privacy preference from public data," *Information Systems Research*, pp. 1–20, 2024, Articles in Advance. DOI: 10.1287/isre.2023.0318.

[15] R. N. Zaeem et al., "Privacycheck v3: Empowering users with higher-level understanding of privacy policies," in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 2022, pp. 1593–1596.