ToneCheck: Unveiling the Impact of Dialects in Privacy Policy

Jay Barot
Vanderbilt University
Nashville, TN, USA
jay.d.barot@vanderbilt.edu

Ali Allami
Vanderbilt University
Nashville, TN, USA
ali.allami@vanderbilt.edu

Ming Yin
Purdue University
West Lafayette, IN, USA
mingyin@purdue.edu

Dan Lin
Vanderbilt University
Nashville, TN, USA
dan.lin@vanderbilt.edu

ABSTRACT

Users frequently struggle to decipher privacy policies, facing challenges due to the legalese often present in privacy policies, leaving trust and comprehension shrouded in ambiguity. This study dives into the transformative power of language, exploring how different linguistic tones can bridge the gap between legal, technical jargon, and genuine user engagement—through a comparative analysis involving diverse focus groups, immersing them in three distinct policy variations: *legalistic*, *casual*, and *empathetic*. We explored how these tones reshape the user experience and bridge the gap between legal discourse and comprehension. Analysis of the data revealed significant associations between linguistic tone and user trust and comprehension. The adoption of an empathetic tone significantly enhanced user trust, as evidenced by a 40.4% increase compared to alternative language styles. This preference highlights the human desire for genuine connection, even in the intricate domain of data privacy. Furthermore, comprehension indices arise for both empathetic and casual tones, leaving legalistic language lagging far behind. This suggests a clear path towards user-friendly policies, where clarity exceeds complexity. Our exploration goes beyond mere compliance. We illustrate the complex gap between subtle linguistic shifts and user perception. By deciphering the language that resonates with trust and understanding, We plant the seeds for the development of privacy policies that not only meet legal requirements but also enhance user trust and comprehension.

CCS CONCEPTS

• **Social and professional topics** → **Privacy policies**; • **Security and privacy** → *Privacy protections.*

KEYWORDS

Privacy Policy, Linguistic Tones, User Trust, Comparative Analysis

SACMAT 2024, May 15–17, 2024, San Antonio, TX, USA

ACM ISBN 979-8-4007-0491-8/24/05
https://doi.org/10.1145/3649158.3657035

ACM Reference Format:
Jay Barot, Ali Allami, Ming Yin, and Dan Lin. 2024. ToneCheck: Unveiling the Impact of Dialects in Privacy Policy. In *Proceedings of the 29th ACM Symposium on Access Control Models and Technologies (SACMAT 2024), May 15–17, 2024, San Antonio, TX, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3649158.3657035

1 INTRODUCTION

In our rapidly advancing digital age, the safeguarding of user privacy has emerged as a critical concern. As individuals navigate the intricate landscape of online platforms, privacy policies emerge as pivotal documents, shedding light on the intricate practices surrounding collecting, storing, and utilizing personal data. Essentially, these policies serve as indispensable beacons, designed to furnish users with vital insights into the ways their data is handled by the platforms they engage with [24].

The prevailing approach adopted by numerous websites, tools, and online services in crafting their privacy policies involves the utilization of automated generation tools. Notable examples include Termly, Termsfeed, and GetTerms, which have become trending privacy policy generators. These tools deploy templates based on policy standards such as the California Consumer Privacy Act (CCPA) and General Data Protection Regulation (GDPR) [38]. While small organizations and websites often rely on auto-generated templates, medium to large organizations exhibit a more tailored approach, creatively presenting their privacy policies while adhering to CCPA and GDPR norms. Additionally, specialized tools have been developed to facilitate automated compliance checking and enhance control over privacy policies, ensuring alignment with CCPA and GDPR regulations [2, 6].

Extensive research has explored trust dynamics, cultural influences, linguistic subtleties, content analyses, and the role of design and readability in these documents. While previous studies established a strong link between policy content and user trust, this research stands out by actively proposing solutions, testing diverse language tones, and suggesting strategies for trust and comprehension improvement. It complements existing work on policy analysis, content evolution, and readability by introducing innovative approaches to enhance these dimensions. The study focuses on the complex relationship between trust and loyalty among users experiencing software, websites, or web tools for the first time, particularly during the critical phase of establishing initial trust. The overarching problem is the challenge of articulating privacy

policies in a clear and user-friendly manner, prompting the need for a paradigm shift in content, language, and presentation formats.

Beyond merely meeting legal requirements and refining analysis and presentation methods, the significance of privacy policies transcends into crucial areas such as transparency, building trust, and empowering users. They go beyond being mere regulatory obligations, playing a crucial role in creating an atmosphere where users can make informed decisions about their online interactions. Privacy policies, when strategically crafted, can become powerful communicative tools, fostering an environment where users not only comprehend the intricacies of data handling but also trust the platforms they engage with. The current dilemma lies in the fact that, despite recognizing their potential impact on user experience, the majority of privacy policies tend to adhere to a standardized language dictated by legal frameworks.

Our study presents a compelling opportunity for the exploration of a novel path. By diving into the detailed dialects employed in privacy policies, we aim to perform a comparative analysis of how variations in tones ranging from formal *legalistic"* to *"casual"* conversational and user-centric *"empathetic"* can reshape the user experience. Beyond compliance, our focus is on understanding how these subtle linguistic shifts can bridge the gap between legal discourse and user comprehension, ultimately contributing to the establishment of privacy trust and the empowerment of individuals.

To conduct this study, we carefully selected diverse focus groups spanning various disciplines. Each focus group was exposed to three distinct versions of privacy policies, each crafted to embody one of the three tones under examination. These policies were adapted from a prominent web source or tool commonly utilized in their respective domains. The presentation took place in a controlled environment, ensuring a standardized experience for all participants. Following the exposure to the privacy policies, participants were administered a survey designed to evaluate their trust, comprehension, and tone preference. This structured approach allowed us to systematically analyze the impact of different tones on user perception within the context of privacy policies.

Key findings from our research reveal distinct patterns in user responses to different privacy policy tones. Firstly, trust levels are notably higher among users when privacy policies adopt an *empathetic* tone. Secondly, the comprehension index for *empathetic* and *casual* tones remains consistently high, surpassing the *legalistic* tone. Furthermore, in terms of user preference, our findings indicate a clear inclination towards favoring *empathetic* language, with 40.4% expressing a preference for this tone. In contrast, 33.3% of users favored the *casual* tone, while a mere 26.3% expressed a preference for the *legalistic* tone.

1.1 Our Contributions

This paper explores the link between language, security, and privacy. It goes beyond legal compliance to examine how subtle wording choices impact user understanding of privacy policies. Current practices often use confusing legalese, leaving users vulnerable. This research proposes clear, user-friendly language to bridge this gap. The findings reveal a fascinating connection between language and trust. The research suggests that employing an empathetic tone in privacy policies leads to higher trust levels among users. While acknowledging the potential for misuse, the paper emphasizes the power of responsible language in fostering user agency and informed consent. Ultimately, these insights offer invaluable considerations for building privacy policies that go beyond legal compliance, fostering genuine understanding and trust, and empowering users to make informed decisions about their personal information. This user-centric approach paves the way for a more secure and privacy-conscious future for all.

To summarize our contributions come in several folds as follows:

(1) We aim to explore how different linguistic tones can affect user trust and comprehension of privacy policies across various domains.
(2) To the best of our knowledge, we propose the first insight to enhance the readability and user-friendliness of privacy policies by testing and comparing three tones: legalistic, casual, and empathetic. The paper also aims to identify the optimal tone-domain combination that maximizes user trust and comprehension.
(3) To assess the impact of linguistic tone on privacy perception, we recruited 99 participants across 8 industries and presented them with 3 versions of privacy policies (legalistic, casual, empathetic) followed by a survey measuring trust, comprehension, preference, and user demographics/privacy experience.
(4) We analyze the data using descriptive and inferential statistics, and reveal several key findings. For example, an empathetic tone significantly boosts user trust, with a 40.4% increase compared to others, reflecting a desire for connection and transparency in data privacy. Comprehension is consistently high for empathetic and casual tones, surpassing legalistic ones, suggesting a path towards user-friendly policies prioritizing clarity. While no single tone-domain combo maximizes both trust and comprehension, a trade-off exists. The optimal choice depends on context, organizational goals, and user needs.

The following sections explore related work (Section 2), define the problem statement (Section 3), present the proposed methodology (Section 4), detail the experimental study (Section 5), analyze the results (Section 6), and conclude with insights (Section 7).

2 RELATED WORK

In the ever-evolving digital landscape, privacy policies play a pivotal role in shaping user experiences on websites, impacting trust, comprehension, and overall engagement. Extensive research has delved into diverse facets, such as trust dynamics, cultural influences, linguistic subtleties, comprehensive content analyses, policy coverage, and the significance of design and readability in these documents.

Trust and Cultural Nuances: Previous research has established a clear link between privacy policy content and the willingness of individuals to share personal information online [50], emphasizing the importance of cultural factors in shaping trust and privacy concerns [5, 10, 42, 50]. These studies suggest that cultural backgrounds significantly influence perceptions of trust in privacy policies, through both content and linguistic nuances. Our research

builds on this foundation by specifically exploring how different linguistic tones within privacy policies affect user trust, highlighting the critical role of cultural contexts in digital interactions and trust formation.

Dialect Impact: Research on the linguistic aspects of privacy policies has primarily focused on their legalistic language, which often seems irrelevant to users. Early investigations, like the one by Pollach and Irene [40], began mapping out the linguistic terrain of privacy policies, but it's the identification of language as a barrier [43] to user understanding that has sparked interest in innovative solutions. Studies advocating for clearer alternatives, such as the Compact Privacy Policy Language (CPPL), aim to make privacy policies more accessible [29]. Our work distinguishes itself by not only addressing these identified challenges but by going further to propose and test creative solutions across various linguistic tones and domains, enhancing user engagement and comprehension.

Tools and Models for Analysis: Other studies have directed their attention toward identifying problems within privacy policies, employing tools and models to accurately detect questionable policies across various domains, including Android applications and websites [28, 39, 51, 52]. In particular, some noteworthy studies examine privacy policy changes over time, utilizing a web crawler as a 'time machine' to collect and aggregate all modifications [3]. However, the primary emphasis of these works lies in analyzing policies and policy content evolution rather than proposing strategies for trust and comprehension improvement.

Content and Coverage: Research on privacy policies has extensively explored their content, focusing on identifying key information such as data collection practices, disclosure recipients, and security measures [24]. These studies, spanning multiple disciplines [35, 36, 41], have established a foundation for advanced policy analysis, emphasizing the development of systematic approaches for assessing policy coverage [36]. Innovations in this field have introduced automated and manual techniques to detail and categorize policy content [45], leading to the creation of models aimed at improving analysis accuracy [47]. Our research builds on these efforts by examining the impact of linguistic tones on user perception, suggesting that as linguistic tones evolve, analytical tools may need updates to remain effective. Ultimately, our goal is to contribute to the development of privacy policies that are not only comprehensive but also accessible and user-friendly, moving away from complex legal jargon.

Comprehension and Readability: Recognizing the pivotal role of privacy policies in fostering effective user engagement on websites, scholars on privacy policies has consistently underscored their critical role in user engagement, advocating for their clear and understandable presentation [33]. Research efforts have extensively evaluated the readability and comprehension of privacy policies, with comparative analyses across major organizations aiming to identify communication patterns [17]. Further studies have probed the relationship between the clarity of privacy policies and users' trust and perceptions of website credibility [7]. While these investigations have significantly advanced our understanding of privacy policy readability and its impact on user trust [15, 22], our research introduces innovative methods to further enhance policy readability and comprehension. By building on previous findings, our study proposes new approaches to making privacy policies more accessible to users, marking a distinct advancement in the quest for clearer and more engaging privacy communications.

User Interface: Pilot studies have established a foundational relationship between design and trust [48], leading to subsequent investigations in the domain of how the format and visual presentation of privacy policies impact trust and comprehension. Experimental comparative studies have explored how alterations in the appearance of privacy policies can enhance the perceived trustworthiness of online services [1, 37]. These studies not only evaluate the impacts of various formats but also offer recommendations for what they consider as best practices [11]. In recent years, the findings from these studies have been actively adopted by numerous organizations, resulting in observable changes in the presentation of privacy policies. Furthermore, there have been studies aimed at improving the visualization of policies that do not adopt impactful design patterns [12, 26]. However, it is essential to note that our study takes a distinctive approach by not focusing on the effects of the user interface on policy directly. Instead, we treat the user interface as an independent variable in a moderated environment within our study design, as we will elaborate on in later sections.

Loyalty and Trust: While existing studies have delved into the intricate relationship between loyalty and trust established through privacy policies within organizations [21], others have explored how loyalty can influence perceived privacy [47]. Moreover, certain studies have proposed loyalty program protocols designed to stimulate a more privacy-aware user base [30]. Notably, these investigations predominantly concentrate on the established user base. In contrast, our study adopts a distinctive approach by placing a central focus on first-time users during the pivotal phase of initial trust establishment

3 PROBLEM STATEMENT

Privacy policies function as vital instruments for companies to communicate with users, elucidating the intricacies of data handling and answering the questions of how, why, and for what purposes private data is collected [33]. However, a significant challenge surfaces as these policies often encompass intricate social, legal, and technical facets, demanding a clear and user-friendly articulation. Notably, the readability of privacy policies poses a considerable hurdle, with most presented in ways that require college-level reading, thereby presenting users with the formidable task of deciphering legalistic, confusing, or jargon-laden language [37].

Consider, for instance, the average user encountering a privacy policy dense with legalese, technical terminology, and convoluted clauses. Such complexity invariably obstructs their understanding, fostering a sense of opacity rather than transparency. Consequently, as shown by many studies, users do not trust online services concerning the use of their private data [1]. The research underscores that, instead of instilling trust, privacy policies tend to heighten user concerns regarding data privacy. This conundrum prompts some of the research to propose a crucial need for a strategic redesign of the entire content, language, and presentation formats [40].

Take, for example, the challenge organizations face in conveying specific legal and technical details, often necessitating expert

language. Striking a balance between conveying accurate information and ensuring user comprehension becomes a formidable task [33]. Beyond the mere fulfillment of legal requirements, the significance of privacy policies transcends into pivotal areas such as transparency, trust-building, and user understanding.

Consider the broader implications, of a user navigating a website, attempting to discern how their data is handled. Privacy policies, if strategically crafted, possess the potential to be transformative communicative tools. Yet, despite their potential impact on user experience, the majority of privacy policies remain entrenched in a standardized language dictated by legal frameworks. This creates a dilemma, underscoring the imperative for a paradigm shift towards user-centric, comprehensible approaches that empower users to make informed decisions about their online interactions.

In response to a critical juncture in the field, our research introduces a novel innovative approach centered on linguistic experimentation. Our proposed solution involves a comprehensive exploration of diverse dialects. The objective is to transform the user experience with privacy policies. Through empirical investigation, our study aims to study the impact of different linguistic tones on user trust, comprehension, and preference. By strategically testing and comparing these varied dialects in the context of privacy policies, we seek to uncover insights that address the current limitations associated with standardized dialects commonly employed in such documents. Furthermore, our research contributes valuable considerations to the ongoing discourse on privacy practices. These insights are intended to inform the development of policies, guiding how organizations can enhance the quality of their privacy-related communications.

3.1 Problem definition

Given the above problem statement, this section paves the path for defining the problem of evaluating privacy policy dialects for user trust and comprehension.

Definition 1. Given dialect (D) a specific stylistic and lexical register is used within a privacy policy. User Trust (T): A measure of a user's confidence in an organization regarding its data handling practices. User Comprehension (C): A measure of a user's understanding of the information conveyed in a privacy policy. Linguistic Tone (L): A specific set of stylistic and lexical features characterizing a dialect.

The objective: Given a set of privacy policies (P) expressed in different dialects (D) and linguistic tones (L), we aim to identify the dialect-tone combination (DL) that maximizes both user trust (T) and comprehension (C).

The Constraints: Dialects must comply with all relevant legal and regulatory requirements. Dialects should be grammatically correct and stylistically consistent.

Trust-Comprehension Score (TCS): It seeks the dialect and tone that simultaneously maximize both user trust and comprehension, measured by their product as follows:

$$TCS = Maximize(T(DL) * C(DL))$$

4 PROPOSED METHODOLOGY

The objective of our methodology is to analyze the impact of linguistic tones on user trust (T) and comprehension (C) of privacy policies across varied domains (d). We start by defining the following sets and datasets:

4.1 Tone Set (Λ)

Without loss of generality, the tone set can be represented as follows:

$$\Lambda = \{Legalistic(l), Casual(c), Empathetic(e)\}$$

In our exploration of the impact of linguistic dialects on privacy policies, the selection of three foundational tones forms the cornerstone of our study design. We begin with the ubiquitous formal legalistic tone, which has long been the standard in generating privacy policies. This tone serves as a vital baseline, allowing us to draw insights into the enduring dynamics associated with formal language, even after years of its prevalent use.

Considering the contemporary landscape influenced by transformative language models, such as generative transformers, it becomes crucial to acknowledge the growing prevalence of conversational tones [9, 20, 25]. These models, trained to understand and generate natural language, have become integral to various platforms, enabling users to engage in conversations seamlessly [31]. Given their widespread adoption and the shift towards conversational interfaces [46], we include a casual conversational tone as a parameter under study. This choice aligns with the current trend of integrating generative transformer models, especially in applications like chatbots, where their conversational nature enhances user interactions.

Furthermore, recognizing the inherent link between empathy and trust in human communication [9, 18], we strategically incorporate an empathetic tone as the third parameter under study. Understanding that empathy is a powerful catalyst for building trust, especially in sensitive matters like data privacy [34], this choice aims to explore the detailed impact of a user-centric and empathetic approach within privacy policies. By contrasting these three tones—formal legalistic, casual conversational, and empathetic we aim to unravel the intricate dynamics that language tones bring to the user's perception of privacy policies.

4.2 Focus Groups (G), Domain Set (Ω), and Privacy Policy Dataset (Φ)

After finalizing our linguistic tones, our focus shifts to the strategic selection of focus groups (G), domain set (Ω), and privacy policy dataset (Φ) which are defined as follows:

$$G = \{G_1, ..., G_8\}$$

where G_i represents participants from domain D_i.

$$\Omega = \{Education\&Training, Software\&IT, Banking\&Financial, Healthcare, Engineering, RetailWholesale\&Distribution, JuridicalSciences(Law), NonProfit\}$$

$$\Phi = \{(D_i, t_j, P_{ij}) : D_i \in \Omega, t_j \in \Lambda\}$$

. Where P_{ij} is a curated paragraph from policy in D_i with tone t_j, $i = \{1, \ldots, 8\}$, and $j = \{1, \ldots, 3\}$.

As illuminated in the related works section, privacy policy research spans a myriad of fields due to the pervasiveness of digital information across diverse domains. The prevalence of digital data, from sensitive health information to financial data in banking [19] and the everyday data collected by websites, positions privacy policies as a topic of paramount importance across all sectors.

In recognition of the varied landscape of data usage practices in major domains, we purposefully select participants representative of eight major disciplines: ***Education & Training***, ***Software & IT***, ***Banking & Financial***, ***Healthcare***, ***Engineering***, ***Retail Wholesale & Distribution***, ***Juridical Sciences (Law)***, and ***Non-Profit***. By incorporating a broad spectrum of disciplines, we ensure a comprehensive understanding of how linguistic tones within privacy policies resonate across varied industries. This approach enriches the depth of our study and also enhances the applicability of our findings, making them relevant and insightful for a diverse range of applications.

In constructing our dataset of privacy policies representative of the eight focus groups, a meticulous process was undertaken to curate a list of websites for each domain. Our objective was to ensure that the selected website encompassed information and tools pertinent to the respective domains, effectively capturing the diversity of data handling practices within each focus group.

To gauge the popularity and relevance of websites within a domain, we adopted a procedure rooted in standard Search Engine Optimization (SEO) policies commonly used for competitor analysis [16, 32]. Leveraging SEO tools such as SemRush, Ahrefs, and readily available indexes like SimilarWeb, we delved into key metrics including Average Bounce Rate, Conversion Rate, Average Session Duration, Average Time on Site, and Unique Visitors. These metrics provided valuable insights into the popularity and engagement levels of websites within a particular domain [14].

As an illustrative example of our procedure, consider the domain of Software & IT. OpenAI.com stood out as a prominent website for this domain based on its impressive metrics. The website demonstrated a low Average Bounce Rate, indicating high user engagement, a noteworthy Conversion Rate suggesting effective user interaction, and a substantial number of Unique Visitors. The site's comprehensive coverage of topics ranging from artificial intelligence research to cutting-edge technologies positioned it as a representative choice for the Information Technology focus group. This meticulous approach was consistently applied across all eight domains, ensuring the selection of websites that not only reflected the popularity within their respective domains but also encompassed diverse facets of data handling practices. Table 1 provides detailed information about each of the selected websites, offering comprehensive insights into the diverse domains represented in our study.

After curating our list of websites representative of each industry domain, the next step involves extracting their privacy policies. Notably, we observed diversity in the structure of these policies, with some adhering to a standard one-page format (e.g., OpenAI), while others featured a multi-page layout with interactive sections (e.g., CDC). To systematically analyze the content, we manually compiled a list of headings extracted from these privacy policies.

Industry Domain	Website
Education & Training	coursera.org
Software & IT	openai.com
Banking & Financial	economist.com
Healthcare	cdc.gov
Engineering	construction.autodesk.com
Retail, Wholesale & Distribution	waze.com
Juridical Sciences (Law)	justice.gov
Non-Profit	jstor.org

Table 1: Industry Domains and Websites

To ensure a focused examination, we compared these headings against the guidelines outlined by CCPA and GDPR [4]. Guided by these regulations and identifying commonalities among headings across all privacy policies, we refined our list to include four key headings consistently present in the policies. These headings, namely *Introduction And Who We Are*, *Information We Collect*, *Age and Children*, and *Changing Our Privacy Notice*, were identified as central components that encapsulate crucial information within the privacy policies under study. It is worth noting that during the manual extraction of headings from the privacy policies, we encountered a few other common titles such as "Contact" and "Complaints." However, due to their relatively brief length and frequent placement within the footer sections of the policies, we decided to exclude these sections from our narrowed list.

Upon identifying our four common sections, we created three versions of the industry-specific privacy policy, each exclusively consisting of the aforementioned four sections. Importantly, to enhance readability, each section was limited to a single paragraph [17, 27]. The creation of these three versions for each policy involved careful curation, with the original text under each section modified to represent one of the three linguistic tones: *legalistic*, *casual*, and *empathetic*. The text curation process underwent moderation and validation using *Grammarly* and *Sapling* tools to ensure the authenticity and representativeness of the chosen linguistic tones [44, 49]. This meticulous procedure was replicated for the privacy policies of all eight industry domains, resulting in the curation of a total of $8x3 = \mathbf{24}$ privacy policies, each offering an in-depth representation of the different tones across diverse domains.

Recognizing the absence of explicit tone categories within popular online language tools such as Grammarly and Sapling, we devised a strategy to synthesize their extensive list of tone markers into our targeted tone set (Λ). It is important to acknowledge that this process is inherently subjective, rooted in the conventional implications and usage of these tones in everyday communication. Our methodology hinged on selectively aggregating tones that aligned with our focused categorizations, intentionally omitting tones that, while potentially relevant, did not directly contribute to our primary analysis. For the legalistic tone, we combined markers such as "*Neutral*," reflecting objectivity; "*Approving*," for contexts of agreement; and "*Disapproving*," indicative of legal dissent. The empathetic tone was shaped by combining "*Sympathetic*," to convey understanding; "*Grateful*," for expressions of appreciation; and "*Loving*," to evoke a sense of affectionate communication. Lastly, the

casual tone encompassed markers like "*Amused*," embodying humor; "*Curious*," for informal inquiry; "*Eager*," representing enthusiasm; and "Excited," to capture spirited discourse.

To demonstrate the transformation of the "*Introduction and Who We Are*" section from OpenAI's privacy policy into three distinct tones—legalistic, casual, and empathetic—we start with the original text as our baseline. This initial version is formal, serving as the foundation for our adaptations, characterized by its clarity and compliance, which is typical of a legalistic tone. In the legalistic version, we retain this formal tone, emphasizing precision and authority to ensure the text remains clear and direct, as seen in phrases like "*We, at OpenAI OpCo, LLC, together with our affiliates, respect your privacy...*". Conversely, the casual version takes a markedly different approach by transforming the content into a more relaxed and engaging narrative. It employs everyday language and expressions to demystify the legal jargon, aiming to make the policy more accessible and engaging for the general public, evident in the friendly introduction "*We're the crew at OpenAI OpCo, LLC, and we really care about your privacy...*". Lastly, the empathetic version is specifically crafted to resonate on a personal level with the reader, using language that conveys understanding, care, and concern about their privacy, highlighted by the welcoming "*Welcome to OpenAI OpCo, LLC, where we truly value your privacy...*".

Having successfully curated our dataset, we proceed to execute the experimental study. In the subsequent stages of the methodology design, we present the three curated industry-specific privacy policies to the respective focus groups representing each domain. It is crucial to note that members of the focus group are kept uninformed about the specific goals and motives of the study. Instead, they are informed that the overarching objective of the project is to enhance overall privacy and security. This intentional lack of detailed information aims to eliminate bias formation among participants [8, 13, 23].

Additionally, participants are queried about their prior experience with reading privacy policies as part of data collection and demographic profiling. This information serves to enhance our understanding of the participant's background and may reveal any correlations in the data post-experiment. Moving forward, the three privacy policies are presented to each focus group member in a randomized sequence to ensure that the order of presentation does not become a confounding factor; or independent variable in the data results. Importantly, participants are not compelled to thoroughly read each policy, mimicking the natural setting where individuals may casually browse through such documents. Once participants have perused all three privacy policies, they are then invited to complete a post-survey. This survey comprises questions diving into their perceptions of trust, comprehension, and preference regarding the presented privacy policies.

4.3 Survey Instruments

Next, to understand how different ways of writing privacy policies affect people, we designed special tests. After reading each policy version, participants answered questions to see how much they trusted the company (T_{ij}) and how well they understood the information (C_{ij}). These tests helped us uncover how different writing styles influence people's feelings and understanding of privacy policies. For this, we define our survey instruments as follows:

- **Trust Measure (T_{ij}):** Measured after participant in G_i reads P_{ij}
- **Comprehension Measure (C_{ij}):** Measured after participant in G_i reads P_ij

Our research rests on three key hypotheses as follows:

- **H1:** T_{ij} is dependent on t_j and D_i. We believe the way a privacy policy is written (its tone) and the context in which it is applied (the domain) will influence how much people trust it.
- **H2:** C_{ij} is dependent on t_j and D_i. Similarly, different tones and domains are likely to affect how well people understand the information presented.
- **H3:** Exist D'_i and t'_j such that $max(T'_{ij}C'_{ij}) > max(T_{ik}C_{ik})$ for all $(i,k) \in [(1,8),(j,k)in(1,3)]$. However, our ultimate goal is to identify a "sweet spot" – a specific tone and domain combination that maximizes trust and comprehension.

In other words, we're searching for the magic words and settings that unlock the highest level of user empowerment when it comes to understanding and trusting how their data is handled.

To uncover the magic formula for effective privacy policies, we launched a carefully controlled experiment. We recruited N participants for each domain (G_i), ensuring a diverse pool of perspectives. Each participant encountered three versions of a domain-specific privacy policy, each crafted in a different linguistic tone ($P_{ijt}, P_{ikj}, P_{ilm}$). To avoid bias, the order of presentation was shuffled, and after each encounter, participants faced challenges to gauge their trust (T_{ij}) and comprehension (C_{ij}) of the information. Beyond the policy content, we also gathered demographic data and their prior experience with privacy policies, giving us a richer understanding of how individual characteristics might influence their responses. This detailed procedure ensured a comprehensive and unbiased evaluation of the impact of linguistic variations on privacy policy effectiveness.

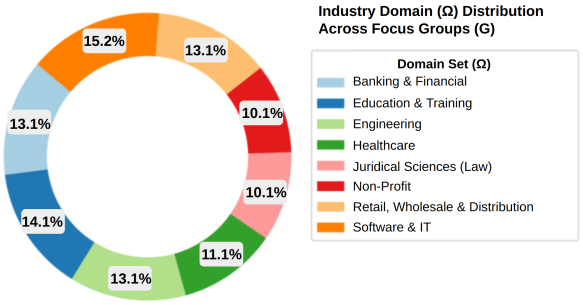

Figure 1: Participants Demographic - A

5 DATA COLLECTION

In the data collection phase of our study, participant recruitment was essential for ensuring diversity and representativeness across eight domains. The process unfolded in two phases, starting with a pilot study to obtain preliminary results for identifying trends and

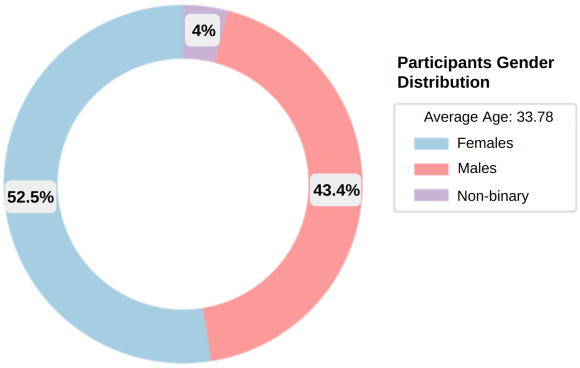

Figure 2: Participants Demographic - B

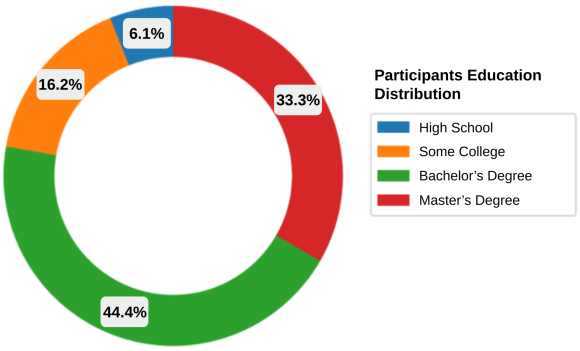

Figure 3: Participants Demographic - C

assessing data quality. A multifaceted recruitment strategy was utilized, involving email, social media, and academic announcements, coupled with a $5 monetary incentive to value participants' time and encourage their involvement. This approach resulted in the active participation of 19 individuals in the pilot study, laying a solid foundation for the subsequent comprehensive analysis.

Building on the foundation laid by the pilot study, our research expanded to utilize the Clickworker crowdsourcing platform for further data collection. This subsequent phase was meticulously organized into five iterative campaigns on Clickworker, enabling thorough verification of the data gathered. The combination of Clickworker and our initial pilot study facilitated the accumulation of a total of 99 data points across diverse domains. Specifically, data distribution was as follows: 13 data points from the Banking & Financial sector, 14 from Education & Training, 13 related to Engineering, 11 from Healthcare, 10 pertaining to Juridical Science (Law), 10 from Non-Profit organizations, 13 associated with Retail, Wholesale, and Distribution, and 15 from the Software & IT domain.

To ensure a well-rounded representation, our study carefully balanced the demographic distribution, as depicted in Figure 1, 2. This deliberate effort through rigorous prescreening aimed to capture diverse perspectives across industry domains. The resulting participant pool enriches the study's findings, contributing to a more comprehensive understanding.

To ensure ease of access and enhance participant engagement, a user-friendly interface was developed utilizing the Angular TypeScript framework. This interface, referred to by the header "*Understanding User Perceptions of Privacy Policies*", remains accessible for public review and future reference, designed to streamline the participant experience.

The participant journey initiates with a pre-screening process designed to assess eligibility according to alignment with a predefined set of industries, domains, majors, or job functions. Participants not meeting the criteria are promptly redirected to a notification page, and measures are implemented to block their Clickworker ID and IP address, thereby preventing any subsequent participation attempts. After this initial screening, the study unfolds in three distinct, sequential steps, where progression to the next phase is contingent upon the successful completion of the preceding one. This structured approach aims to clarify the process, streamline participant engagement, and promote active, informed participation.

Step 1: Consent and Demographic Information: The participants were directed to the consent and demographic information section. Here, the process was designed to be straightforward and transparent. Participants were presented with a clear overview of the study's general objectives, without examining specific details. However, they were provided with comprehensive information about the steps involved, the overall process, and an estimate of the time required to complete their participation, which was to be around 15 minutes on average.

Following the consent phase, participants proceeded to provide their demographic information, a critical component for tailoring the study to individual characteristics. This information included their Clickworker ID serving as a unique identifier throughout the study. It was particularly useful in later stages, especially during the post-survey, where it helped match participants with their responses.

Participants provided crucial demographic details, including gender, age, and their affiliated industry domain. This information was instrumental in accurately presenting privacy policies tailored to each participant's specific domain. Education level referenced as Figure 3 was also collected to assess how it might influence participants' understanding of privacy policies. The survey included a question about participants' prior experience with reading privacy policies and the frequency with which they update their privacy settings on online accounts, aiming to explore potential correlations. Responses to this question are depicted in a pie chart (referenced as Figure 5, 6).

Step 2: Privacy Policies: In the subsequent phase of the study, participants engage in a crucial task: reviewing privacy policies specific to their industry. This step is vital for maintaining the integrity of the research. To counter potential biases and ensure objective participation, each participant is assigned a randomized sequence of shapes—circle, square, or triangle. The design of the web application incorporates a unique logic, where each shape symbolizes a different linguistic tone: circles are associated with a legalistic tone, squares with a casual tone, and triangles with an empathetic tone. This methodological strategy aims to provide participants with a clear, visual cue regarding the linguistic tone conveyed by each shape, facilitating an intuitive and unbiased interaction with the

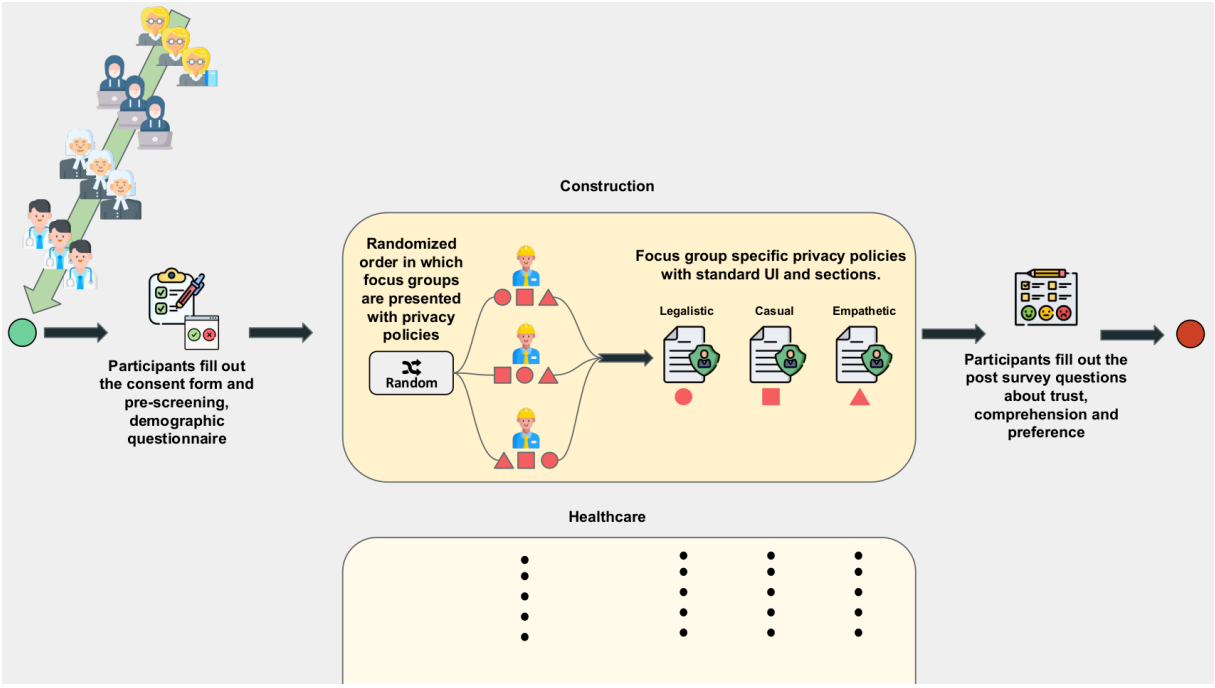

Figure 4: Experiment Flow Chart

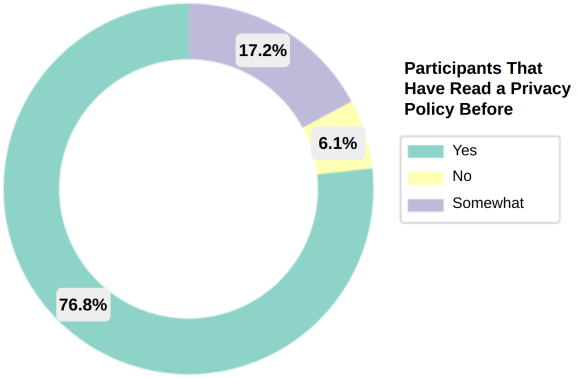

Figure 5: Prior Experience with Privacy Policies - A

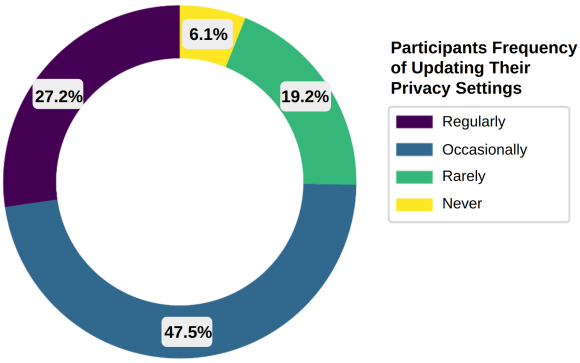

Figure 6: Prior Experience with Privacy Policies - B

privacy policy content. Table 2 illustrates the association between shapes and linguistic tones for clarity.

Shape	Linguistic Tone
Circle	Legalistic
Square	Casual
Triangle	Empathetic

Table 2: Shapes and Linguistic Tones

The strategic development of the front-end design, coupled with the integration of a curated dataset of 24 privacy policies, presented multiple benefits within the experimental study's framework. Upon selecting their industry domain, participants navigated the privacy policies following a specific shape sequence assigned to them, creating a highly controlled environment for the experiment. This setup was instrumental in enforcing a minimum engagement time of one minute per privacy policy, ensuring participants at least skimmed through the content before proceeding.

Furthermore, the design of the user interface (UI) played a pivotal role as the independent variable in the study. By presenting the privacy policies within this controlled UI environment, the study effectively mitigated potential biases associated with differing visual elements that might influence participant perception. Importantly, the UIs of the original websites or platforms from which the privacy policies were derived were not used. This decision was critical to maintain uniformity and control over the experimental conditions. Elements such as fonts, headings, and paragraph styles were standardized across all privacy policies shown to participants, with the linguistic tone being the sole variable. This approach ensured a

consistent visual and interactive experience, isolating the impact of linguistic tone on participant responses.

Step 3: Post Survey: In the concluding phase of the experimental study, participants proceeded to the post-survey, a pivotal component designed to elicit their views on trust, comprehension, and preference regarding privacy policies. The objective of this phase was to collect detailed feedback on participants' responses to the varied linguistic tones encountered within the policies. To ensure a thorough exploration of their experiences, the question sequence was carefully organized. Data collection utilized Likert scale indices, allowing participants to express the degree of trust and comprehensibility of each privacy policy symbolized by the circle, square, and triangle shapes framed within different linguistic tones. This rating scale spanned from 1 (Low Trust) to 5 (Very High Trust), facilitating a comprehensive assessment of trust levels. Similar scale was incorporated for comprehension levels.

For example, participants used this scale to answer the question "How easy was it for you to understand the privacy policy represented by the legal tone?" from the comprehensibility survey. Our approach to framing these questions was designed to capture the immediate subjective insights on trust and comprehension level of each policy version as directly experienced by the participants, without incorporating indirect tests or inferential questions. The survey's final segment asked participants to express their preference for a specific privacy policy shape within their industry domain. This open-ended question was crucial for understanding participants' subjective inclinations. Moreover, participants were encouraged to articulate the reasons behind their preferences, adding depth and qualitative richness to their responses. This qualitative aspect played a pivotal role in uncovering the nuanced factors influencing participants' choices, providing critical insights into the interplay between linguistic tone, preference, and industry context.

6 EXPERIMENT RESULTS

To ensure the accuracy and integrity of our analysis, we implemented rigorous data-cleaning procedures. This process involved manually correcting spelling and grammatical errors to improve dataset consistency and clarity. We also removed irrelevant responses from crowdsourcing platforms and excluded incomplete data from participants who abandoned the survey after prescreening. These measures significantly enhanced the dataset's quality, reliability, and interpretability, providing a robust foundation for our experiment's detailed analysis.

6.1 First Finding: The Power of Empathy

The analysis of Likert scale responses from participants across diverse industry domains underscores the significant role of empathetic linguistic tones in fostering trust. A consistent pattern emerges, revealing a marked preference for privacy policies articulated in an empathetic tone, which participants perceive as more trustworthy than those framed in legalistic or casual tones. This inclination is statistically substantiated by the mean trust scores, with policies employing an empathetic tone achieving the highest average score of μ=3.82. The standard deviation for trust in empathetic tones, at σ=1.043, indicates a relatively tight clustering of responses around this high mean, suggesting that participants broadly agree

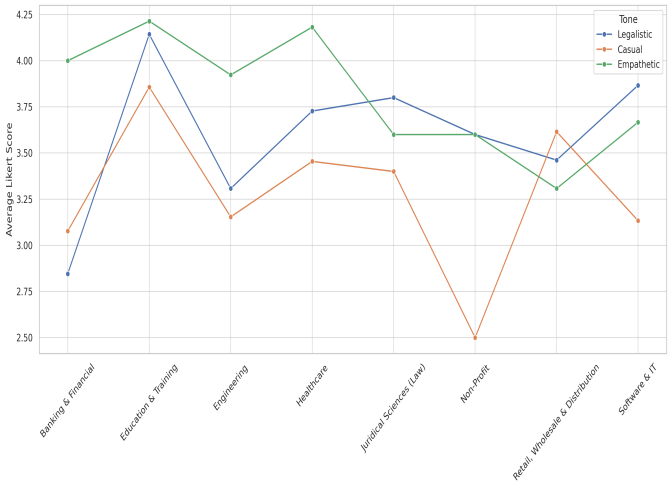

Figure 7: Trust Assessment

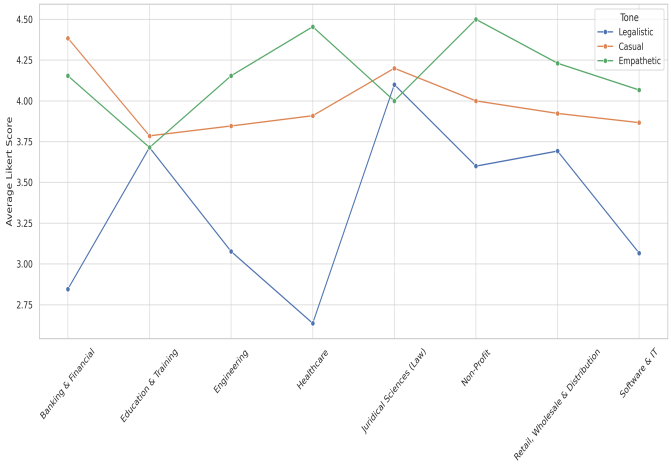

Figure 8: Comprehension Assessment

on the trustworthiness of empathetically toned policies. This trend is visually encapsulated in the Trust Assessment Figure shown in Figure 7, highlighting the nuanced differences in trust perceptions across tones.

The reasons underlying users' preference for an empathetic privacy policy in their respective fields become apparent when examining the post-survey responses. Participants frequently cited elements such as *"The design is like talking to a friend"* and policy phrases like *"We want you to know"*. These responses indicate that users appreciate the conversational and personalized nature of empathetic language in privacy policies. The use of language that fosters a sense of connection and transparency contributes to users' comfort and trust.

This inclination towards empathetic communication is further reinforced by industry-specific feedback. A participant from the Banking & Financial domain highlighted the language as "More assuring and straight to the point. It felt like it was more tailored to my concerns and understanding," emphasizing the value of a

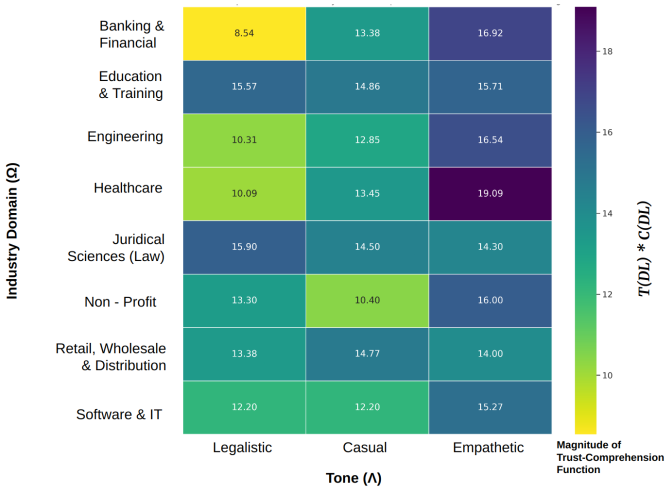

Figure 9: Trust-Comprehension Score

policy that appears customized to address specific user concerns with clarity and assurance. The empathetic tone's importance is particularly pronounced, with a Healthcare domain participant noting, "I went for this policy style because it just feels more personal and caring. It's important in healthcare to feel that your data is treated with care and empathy." This reflects the critical need for privacy policies in healthcare to convey a sense of empathy and personal care, given the sensitive nature of the data involved.

In contrast, legalistic tones, with a mean trust score of μ=3.60 and a standard deviation of σ=1.029, demonstrate the second-highest level of trust among participants. This indicates a moderate level of consensus on the trustworthiness of legalistic tones, albeit less so than empathetic tones. Casual tones, with the lowest average trust score of μ=3.29 and the highest standard deviation of σ=1.311, reveal the greatest variability in trust assessments, suggesting divergent views on their trustworthiness.

The data in Figure 9 demonstrates variations in the Trust-Comprehension Score (TCS) across different tones and domains. For instance, the Healthcare focus group exhibited a high Empathetic TCS score of 19.09, significantly higher than its Legalistic TCS score of 10.09 and Casual TCS score of 13.45. This variance supports the hypothesis **H1** that both the tone of privacy policies and the domain in which they are applied influence the level of trust.

Hence, although the majority of respondents may not have opted for legalistic tones as their preferred choice in industry-specific privacy policies, a qualitative inquiry sheds light on the underlying reasons behind this discernible pattern. Few respondents underscored the significance of upholding formal policy standards, especially within the framework of dispute resolution, drawing from their extensive industry expertise. This insightful comment highlights a pragmatic viewpoint where more formal tones are perceived to imbue policies with an aura of authenticity. This nuanced discovery implies that, within specific professional domains, striking a balance between formal and empathetic tones may prove pivotal in fostering trust.

6.2 Second Finding: Casual and Empathetic Tones Outperform Legalistic

The analysis of Likert scale data across various focus groups illustrates the remarkable consistency in comprehension levels when comparing empathetic and casual tones, both of which notably outperform the comprehension index for legalistic tones. An in-depth examination of the dataset reveals that the mean comprehension scores for casual and empathetic tones stand at approximately μ=3.98 and μ=4.14, respectively. This notable parity is further evident in Figure 8's comprehension chart, highlighting consistent trends across diverse disciplines. This subtle difference, while indicating a slight edge for the empathetic tone, showcases a notable parity in the ease of understanding privacy policies framed in either a casual or empathetic linguistic style.

This parity in comprehension is further evidenced in the standard deviation values for these tones, which are σ=0.93 for the Casual tone and σ=0.78 for the Empathetic tone. These values highlight a consistent trend across diverse disciplines, affirming that, irrespective of the industry domain, users exhibit a notably higher comprehension index when presented with privacy policies employing casual or empathetic tones. This observation is underscored by the findings that consistently show lower comprehension levels for legalistic tones, with a mean score of approximately μ=3.32 and a higher standard deviation of σ=1.13. The higher variability and lower mean score associated with the Legalistic tone underscore a potential challenge in user understanding when confronted with formal and legalistic language.

The analysis of the Trust-Comprehension Score (TCS) from Figure 9 reveals that comprehension varies with tone and domain. For example, the Education & Training focus group showed a higher Empathetic TCS score (15.71) compared to its Legalistic TCS score (15.57), indicating that the tone in which information is presented and the specific domain context affect comprehension levels. These numerical findings suggest that different tones have a distinct impact on how well privacy policies are understood across various domains. Hence, we accept **H2**.

The consistent trend of lower comprehension levels in legalistic tones emphasizes the need for a more user-centric and accessible approach to language in privacy policies. This observation highlights a crucial gap in comprehension that needs to be bridged to ensure that users across diverse industry domains can fully grasp the content and implications of these policies. The findings suggest that employing casual or empathetic tones not only facilitates a better understanding of privacy policies but also represents a strategic approach to enhancing user engagement by making these documents more accessible and relatable. By prioritizing comprehension through the choice of tone, policymakers, and drafters can significantly improve the efficacy of privacy policies, thereby fostering a deeper understanding and engagement among users.

6.3 Third Finding: User Preference for Empathetic Language Across Diverse Disciplines

A substantial majority, **40.4%** of participants, express a preference for an empathetic linguistic tone in privacy policies within their respective domains. This overwhelming preference is reflected

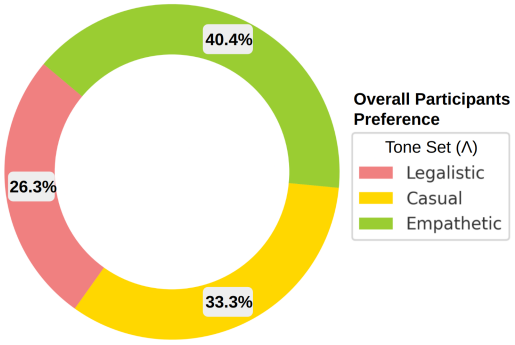

Figure 10: Overall Preference

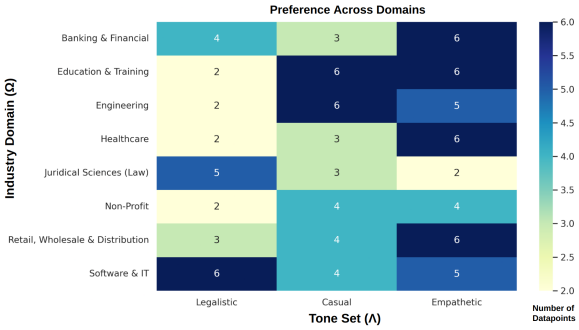

Figure 11: Preference Across Domains

in participant responses, such as "I experienced a stronger sense of connection and trust; it made me sense they truly cared about my privacy." These responses provide valuable depth to users' preferences, emphasizing the importance of building a connection and trust through empathetic language in privacy policies. Another participant noted a preference for "not throwing around legal jargon just for the sake of it," highlighting a desire for a more user-friendly and approachable communication style.

In contrast to the significant preference for empathetic language, our analysis reveals that **33.3%** of users favored the casual tone, indicating a noteworthy but lesser inclination towards a more relaxed and informal communication style in privacy policies. Additionally, a mere **26.3%** expressed a preference for the legalistic tone.

Incorporating individual domain trends into our analysis, particularly through the lens of overall preference indices, reveals nuanced preferences across various fields as captured in the Heat-map Figure 11. This heat map, representing the majority percentages of preferences for the three tones, shows a gradient of preference intensity, with darker shades indicating a stronger preference for the Empathetic tone and progressively lighter shades as preferences shift towards the Legalistic tone.

A significant observation from the heat map is the distinct inclination toward Legalistic tones within the Software & IT, Banking & Financial, and Juridical Sciences fields. This trend, also mirrored in their Trust Assessment in Figure 7, suggests a nuanced industry-specific approach to privacy policy preferences. The Information Technology sector, pivotal in the creation and implementation of online privacy policies, displays a notable trend of preferring Legalistic tones, despite lower Comprehension levels associated with these tones. This preference could be attributed to the sector's emphasis on precision, clarity, and the legal robustness of privacy statements, which might be critical in mitigating risks and ensuring regulatory compliance. Similarly, the Banking & Financial, and Juridical Sciences sectors, known for their bureaucratic nature and stringent regulatory environments, exhibit a higher trust in Legalistic tones. This reflects an industry-specific appreciation for formal, structured language that aligns with traditional practices and legal expectations within these domains. In contrast, other fields, such as Healthcare, Non-Profit, and Education & Training, consistently show a stronger preference for Casual and Empathetic tones, as indicated by higher comprehension and trust indices for these tones. This divergence underscores the importance of tailoring the tone of privacy policies to the audience's expectations and the sector's unique characteristics, enhancing the effectiveness of communication and policy engagement. Furthermore, through the detailed analysis of the Trust-Comprehension Score (TCS), it is evident that certain tone-domain combinations yield the highest TCS score, indicating the presence of "sweet spots." For example, the highest score observed was for the Healthcare focus group under the Empathetic tone, with a TCS score of 19.09. This is significantly higher than any TCS score within the Banking & Financial domain, where the highest score was 16.92 for the Empathetic tone. These numbers strongly support the **H3** that there are optimal combinations of tone and domain that significantly outperform others, reinforcing the goal of identifying specific pairings that maximize trust and comprehension.

7 CONCLUSION

This study highlights the significant impact of empathetic language in privacy policies on user trust and comprehension, contrasting sharply with traditional formal legalistic tones. Our findings reveal that users significantly prefer privacy policies written with an empathetic tone, as evidenced by 40.4% of participants favouring this approach, indicating its potential to bridge the comprehension gap between users and legal discourse. Additionally, the development of a model web application for analyzing privacy policy tones represents a step forward in systematically understanding the nuances of policy communication and its effects on user perception across various sectors.

The research further underscores the ethical considerations in crafting privacy policies, especially in sensitive areas like healthcare (highest TCS score at 19.09), where empathy can both engage users and risk misuse of trust. While empathetic language has shown to enhance user engagement and trust, the potential for manipulation necessitates a careful balance between improving comprehensibility and maintaining ethical standards. This work lays the groundwork for future exploration into creating privacy policies that not only meet legal requirements but also genuinely resonate with and protect users, fostering a digital environment where privacy and security are both robust and user-centric.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation award CNS-2301014

REFERENCES

[1] Esma Aïmeur, Oluwa Lawani, and Kimiz Dalkir. 2016. When changing the look of privacy policies affects user trust: An experimental study. *Computers in Human Behavior* 58 (2016), 368–379.

[2] Orlando Amaral, Muhammad Ilyas Azeem, Sallam Abualhaija, and Lionel C Briand. 2023. Nlp-based automated compliance checking of data processing agreements against gdpr. *IEEE Transactions on Software Engineering* (2023).

[3] Ryan Amos, Gunes Acar, Eli Lucherini, Mihir Kshirsagar, Arvind Narayanan, and Jonathan Mayer. 2021. Privacy policies over time: Curation and analysis of a million-document dataset. In *Proceedings of the Web Conference 2021*. 2165–2176.

[4] Jordan M Blanke. 2020. Protection for 'Inferences drawn': A comparison between the general data protection regulation and the california consumer privacy act. *Global Privacy Law Review* 1, 2 (2020).

[5] Peter Broeder. 2020. Culture, Privacy, and Trust in E-commerce. *Marketing from Information to Decision Journal* 3, 1 (2020), 14–26.

[6] Carl Magnus Bruhner. 2022. Bridging the Privacy Gap: a proposal for enhanced technical mechanisms to strengthen users' privacy control online in the age of GDPR and CCPA.

[7] Rochelle A Cadogan et al. 2004. An imbalance of power: the readability of internet privacy policies. *Journal of Business & Economics Research (JBER)* 2, 3 (2004).

[8] Julie A Carlson. 2010. Avoiding traps in member checking. *Qualitative Report* 15, 5 (2010), 1102–1113.

[9] Avishek Choudhury and Hamid Shamszare. 2023. Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. *Journal of Medical Internet Research* 25 (2023), e47184.

[10] Sophie Cockcroft and Saphira Rekker. 2016. The relationship between culture and information privacy policy. *Electronic Markets* 26 (2016), 55–72.

[11] Lorrie Faith Cranor, Praveen Guduru, and Manjula Arjula. 2006. User interfaces for privacy agents. *ACM Transactions on Computer-Human Interaction (TOCHI)* 13, 2 (2006), 135–178.

[12] Prashant S Dhotre, Anurag Bihani, Samant Khajuria, and Henning Olesen. 2022. "Take It or Leave It": Effective Visualization of Privacy Policies. In *Cybersecurity and Privacy-Bridging the Gap*. River Publishers, 39–64.

[13] Berkeley J Dietvorst and Uri Simonsohn. 2019. Intentionally "biased": People purposely use to-be-ignored information, but can be persuaded not to. *Journal of Experimental Psychology: General* 148, 7 (2019), 1228.

[14] Eric Enge. 2012. *The art of SEO*. " O'Reilly Media, Inc.".

[15] Tatiana Ermakova, Annika Baumann, Benjamin Fabian, and Hanna Krasnova. 2014. Privacy policies and users' trust: does readability matter?. In *AMCIS*.

[16] Michael P Evans. 2007. Analysing Google rankings through search engine optimization data. *Internet research* 17, 1 (2007), 21–37.

[17] Benjamin Fabian, Tatiana Ermakova, and Tino Lentz. 2017. Large-scale readability analysis of privacy policies. In *Proceedings of the international conference on web intelligence*. 18–25.

[18] Jinjuan Feng, Jonathan Lazar, and Jenny Preece. 2004. Empathy and online interpersonal trust: A fragile relationship. *Behaviour & Information Technology* 23, 2 (2004), 97–106.

[19] ATCC Finance. 2015. Industry 4.0 Challenges and solutions for the digital transformation and use of exponential technologies. *Finance, audit tax consulting corporate: Zurich, Swiss* (2015), 1–12.

[20] Mehmet Firat. 2023. How chat GPT can transform autodidactic experiences and open education. *Department of Distance Education, Open Education Faculty, Anadolu Unive* (2023).

[21] Carlos Flavián and Miguel Guinalíu. 2006. Consumer trust, perceived security and privacy policy: three basic elements of loyalty to a web site. *Industrial management & data Systems* 106, 5 (2006), 601–620.

[22] Brian J Fogg, Cathy Soohoo, David R Danielson, Leslie Marable, Julianne Stanford, and Ellen R Tauber. 2003. How do users evaluate the credibility of Web sites? A study with over 2,500 participants. In *Proceedings of the 2003 conference on Designing for user experiences*. 1–15.

[23] Samuel L Gaertner, Jeffrey Mann, Audrey Murrell, and John F Dovidio. 1989. Reducing intergroup bias: The benefits of recategorization. *Journal of personality and social psychology* 57, 2 (1989), 239.

[24] Lei Gao and Alisa G Brink. 2019. A content analysis of the privacy policies of cloud computing services. *Journal of Information Systems* 33, 3 (2019), 93–115.

[25] A Shaji George and AS Hovan George. 2023. A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal* 1, 1 (2023), 9–23.

[26] Kambiz Ghazinour, Maryam Majedi, and Ken Barker. 2009. A model for privacy policy visualization. In *2009 33rd Annual IEEE International Computer Software and Applications Conference*, Vol. 2. IEEE, 335–340.

[27] Mark A Graber, Donna M D Alessandro, and Jill Johnson-West. 2002. Reading level of privacy policies on internet health web sites. *Journal of Family Practice* 51, 7 (2002), 642–642.

[28] Hamza Harkous, Kassem Fawaz, Rémi Lebret, Florian Schaub, Kang G Shin, and Karl Aberer. 2018. Polisis: Automated analysis and presentation of privacy policies using deep learning. In *27th USENIX Security Symposium (USENIX Security 18)*. 531–548.

[29] Martin Henze, Jens Hiller, Sascha Schmerling, Jan Henrik Ziegeldorf, and Klaus Wehrle. 2016. CPPL: Compact privacy policy language. In *Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society*. 99–110.

[30] Tun-Min Catherine Jai and Nancy J King. 2016. Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers? *Journal of Retailing and Consumer Services* 28 (2016), 296–303.

[31] Dietmar Jannach, Ahtsham Manzoor, Wanling Cai, and Li Chen. 2021. A survey on conversational recommender systems. *ACM Computing Surveys (CSUR)* 54, 5 (2021), 1–36.

[32] MNA Khan and A Mahmood. 2018. A distinctive approach to obtain higher page rank through search engine optimization. *Sādhanā* 43, 3 (2018), 43.

[33] Barbara Krumay and Jennifer Klar. 2020. Readability of privacy policies. In *Data and Applications Security and Privacy XXXIV: 34th Annual IFIP WG 11.3 Conference, DBSec 2020, Regensburg, Germany, June 25–26, 2020, Proceedings 34*. Springer, 388–399.

[34] Meira Levy and Irit Hadar. 2018. The importance of empathy for analyzing privacy requirements. In *2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE)*. IEEE, 9–13.

[35] Trina J Magi. 2010. A content analysis of library vendor privacy policies: do they meet our standards? *College & Research Libraries* 71, 3 (2010), 254–272.

[36] Sunil Manandhar, Kaushal Kafle, Benjamin Andow, Kapil Singh, and Adwait Nadkarni. 2022. Smart Home Privacy Policies Demystified: A Study of Availability, Content, and Coverage. In *31st USENIX Security Symposium (USENIX Security 22)*. 3521–3538.

[37] Aleecia M McDonald, Robert W Reeder, Patrick Gage Kelley, and Lorrie Faith Cranor. 2009. A comparative study of online privacy policies and formats. In *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 37–55.

[38] Matthew Newman, Mike Swift, and Vesela Gladicheva. 2020. GDPR and CCPA Start to Bare Teeth as Privacy Protection Goes Global. *Bus. L. Int'l* 21 (2020), 267.

[39] Kristen O'Loughlin, Martha Neary, Elizabeth C Adkins, and Stephen M Schueller. 2019. Reviewing the data security and privacy policies of mobile apps for depression. *Internet interventions* 15 (2019), 110–115.

[40] Irene Pollach. 2007. What's wrong with online privacy policies? *Commun. ACM* 50, 9 (2007), 103–108.

[41] Stephen A Rains and Leslie A Bosch. 2009. Privacy and health in the information age: A content analysis of health web site privacy policy statements. *Health communication* 24, 5 (2009), 435–446.

[42] Ian Reay, Patricia Beatty, Scott Dick, and James Miller. 2013. Privacy policies and national culture on the internet. *Information Systems Frontiers* 15 (2013), 279–292.

[43] Lior Jacob Strahilevitz and Matthew B Kugler. 2016. Is privacy policy language irrelevant to consumers? *The Journal of Legal Studies* 45, S2 (2016), S69–S95.

[44] Ahmad Syafi'i. 2020. Grammarly: An online EFL writing companion. *Journal of English Language Teaching and English Linguistics* 5, 2 (2020).

[45] Khadija Ali Vakeel, Saini Das, Godwin J Udo, and Kallol Bagchi. 2017. Do security and privacy policies in B2B and B2C e-commerce differ? A comparative study using content analysis. *Behaviour & Information Technology* 36, 4 (2017), 390–403.

[46] Svitlana Vakulenko, Vadim Savenkov, and Maarten de Rijke. 2020. Conversational browsing. *arXiv preprint arXiv:2012.03704* (2020).

[47] Samsudin Wahab, Ahmad Suffian Mohd Zahari, Khaled Al Momani, and Nor Azila Mohd Nor. 2011. The influence of perceived privacy on customer loyalty in mobile phone services: An Empirical Research in Jordan. *International Journal of Computer Science Issues (IJCSI)* 8, 2 (2011), 45.

[48] Ari Ezra Waldman. 2018. Privacy, notice, and design. *Stan. Tech. L. Rev.* 21 (2018), 74.

[49] Michael D Winans. 2021. Grammarly's tone detector: Helping students write pragmatically appropriate texts. *Relc Journal* 52, 2 (2021), 348–352.

[50] Kuang-Wen Wu, Shaio Yan Huang, David C Yen, and Irina Popova. 2012. The effect of online privacy policy on consumer privacy concern and trust. *Computers in human behavior* 28, 3 (2012), 889–897.

[51] Le Yu, Xiapu Luo, Xule Liu, and Tao Zhang. 2016. Can we trust the privacy policies of android apps?. In *2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 538–549.

[52] Sebastian Zimmeck and Steven M Bellovin. 2014. Privee: An architecture for automatically analyzing web privacy policies. In *23rd USENIX Security Symposium (USENIX Security 14)*. 1–16.