Building a Long Text Privacy Policy Corpus with Multi-Class Labels

Florencia Marotta-Wurgler[*]
NYU School of Law
wurglerf@exchange.law.nyu.edu

David Stein[*]
Khoury College of Computer Sciences
Northeastern University School of Law
(now at Vanderbilt Law School)
david.stein@vanderbilt.edu

Abstract

Legal text poses distinctive challenges for natural language processing. The legal meaning and effect of a term may be affected by interdependence, the role of defaults in the presence of silence, and terms incorporated by reference. Further, legal text is often susceptible to multiple valid, conflicting interpretations; perhaps the most common answer to a legal interpretation question is "it depends."

This work introduces a new, hand-coded dataset for the interpretation of privacy policies. It includes privacy policies from 149 firms, including documents incorporated by reference. The policies are annotated across 64 dimensions that map onto commonly included terms and applicable US and EU legal rules. Our annotation methodology is designed to capture the core challenges peculiar to legal language, including indeterminacy, interdependence between clauses, and the effects of legal default rules in the presence of contractual silence. We present a set of baseline results for the dataset using current large language models.

1 Introduction

Legal interpretation presents a particularly challenging interpretative task. The legal meaning and effect of a term in a legal document such as a contract may be affected by interdependence between clauses, the role of default rules in the presence of contractual silence, and terms incorporated by reference. These texts can also contain inconsistencies and language susceptible to multiple valid interpretations. Privacy policies are common and notable examples of this: They often contain interdependent clauses—sometimes spread across multiple documents—whose meaning is best understood when read in concert (Reidenberg et al., 2016; Marotta-Wurgler, 2016b,c). Like other legal texts, privacy policies must be interpreted in the context of applicable legal rules, which can define rights and obligations including on matters not explicitly addressed in the text. These interpretative tasks matter, as contract disputes often involve disagreements about the meaning of terms, including situations where the outcome of a case will depend on whether a term is or is not ambiguous. Some of these conflicts involve evaluating reasonable disagreements regarding the meaning of a term. Interpretative challenges are exacerbated by the length, complexity, and subdomain-specific language used in legal documents.

Yet few legal interpretation benchmarks capture these challenges. And—to the best of our knowledge—no legal benchmark attempts to detect reasonable interpretive disagreements. Instead, a common approach is to treat such disagreements as errors that require correction or removal. Not only do datasets that capture some of the challenges inherent in legal interpretation across multiple domains present a compelling NLP challenge; as LLMs begin to play a role in legal institutions and the practice of law, effective datasets are critical for tuning NLP systems to those contexts, and effective benchmarks are critical for assessing their real-world capability.

This work introduces a new, hand-coded dataset for interpreting Privacy Policies. These important legal documents govern the relationship between firms and individuals regarding the collection, use, sharing, and security of personal information, and are generally incorporated by reference in most Terms of Use. Like many consumer standard terms, they are rarely read (Bakos et al., 2014). Privacy policies are long, complex, and require legal expertise to understand. There are also stereotypical legal documents: they are drafted by experts, include domain-specific vocabulary and interpretative conventions (Zheng et al., 2021; Mellinkoff, 2004; Mertz, 2007). Importantly, privacy policies are generally publicly available and map to a con-

[*] Equal contribution

Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 8156–8219
July 27 - August 1, 2025 ©2025 Association for Computational Linguistics

sistent set of well-defined legal questions and applicable legal rules. This presents an opportunity to represent privacy policy content reasonably consistently against which automated interpreters can be tuned and measured.

Current privacy policy datasets either offer high-granularity labels for short samples of policy text or low-granularity classification of longer text. These approaches may not capture many domain-specific aspects of legal interpretation that are relevant to the expanding range of automated legal tasks. For example, neither approach accounts for how documents that are "incorporated by reference" may affect the way a policy restricts (or doesn't restrict) how a company can use user data. As legal interpretation increasingly becomes the target of automation, new datasets are needed. This paper aims to help address that need.

Our dataset of privacy policies is hand-coded by subject-matter experts. Our coding variables span a fairly exhaustive set of terms and capture a representative set of legal questions in this context. Our coding method is designed to capture common challenges of legal contract interpretation by addressing the inherent difficulty associated with interpreting terms that are characterized by inconsistency, ambiguity, or are subject to reasonable disagreements in interpretation. It is comprised of 149 privacy policies and the documents they incorporate by reference, including Terms of Use, Cookie Policies, California Consumers Privacy Act (CCPA) disclosures, and terms complying with the European Union's General Data Protection Regulation (GDPR). Our coding accounts for how applicable legal rules and referenced documents can affect the meaning of terms. Importantly, our approach incorporates relevant legal rules across the U.S. and the E.U., which guide the interpretation of the meaning of contractual silence. Using this dataset, we evaluate the performance of current LLMs. We find that their performance varies widely across tasks and models, but still remains low on many tasks. This suggests that the dataset provides a challenging benchmark for future research in textual interpretation.

2 Related Work

Prior work building datasets for privacy policies mostly focuses on expert annotation or classification of short text. (Lippi et al., 2019; Bui et al., 2021; Ahmad et al., 2021). Some research has also looked into crowd-sourcing annotation (Wilson et al., 2018). One privacy-policy-adjacent dataset involving classification of longer legal text labels the content of cookie banner disclosures with the stated purposes for data collection (Santos et al., 2021). In addition to annotated datasets, there are large-scale compilations of privacy policies scraped from the Internet and Internet Archive(Amos et al., 2021; Srinath et al., 2021).

The most widely-used privacy policy dataset is the OPP-115 dataset introduced by Wilson et. al. in 2016. OPP presented coders with paragraph-length excerpts from 115 privacy policies, to which coders added word-level annotations related to 36 data practices across 10 categories. The OPP dataset was used to train prominent tools used to pick out specific clauses from privacy policies (Harkous et al., 2018; Mousavi Nejad et al., 2020). It has also been used to generate related datasets, either by transforming its annotations for use in a new task like question-answering or GDPR compliance (Poplavska et al., 2020; Ahmad et al., 2020), or as an input into composite legal-task benchmarks like LegalBench and PrivacyGLUE (Guha et al., 2023; Chalkidis et al., 2022). The OPP taxonomy scheme has also been used to organize other privacy-related datasets (Ravichander et al., 2019). Another notable privacy policy dataset—the *unfair-TOS* dataset introduced by Lippi et. al.—annotates "potentially unfair" clauses in privacy practices and is also incorporated into some composite benchmarks, including the privacy-policy-specific PrivacyGlue benchmark (Shankar et al., 2023). Our dataset addresses three key limitations of existing approaches: First, by preserving complete document context, we capture cross-references and definitional relationships that are lost when policies are segmented. Second, our legal rules-based taxonomy reflects actual regulatory categories rather than data practices derived from policy content. Third, we explicitly preserve annotation disagreement as a meaningful signal rather than seeking to maximize inter-annotator agreement.

Benchmarking legal AIs goes beyond the traditional metrics-and-datasets approach. Alternative evaluation approaches include having NLP systems take the bar exam (Bommarito II and Katz, 2022; Katz et al., 2024) (though some have questioned the efficacy of that evaluation approach (Martínez, 2024)), grading LLM-generated law school exam answers (Choi et al., 2021), and measuring how law student performance is affected by LLM use (Choi and Schwarcz, 2023).

The use of coder disagreement as a meaningful signal has recently been introduced as a way to detect ambiguity or minority viewpoints (Jiang and de Marneffe, 2022; van der Meer et al., 2024). The presence of indeterminacy in legal text can affect how the text is interpreted and enforced by courts, meaning annotator disagreement an important feature of a legal annotation dataset. Cross-document relation extraction datasets (Jain et al., 2024) have been used to train models on groups of interrelated documents. In addition to the domain-specific language, legal text often contains document-specific definitions that may differ from ordinary usage—or from other similar documents. Legal interpretation involves holistically parsing complex, interconnected documents with meaningful silence and potentially indeterminate meaning to answer questions that may also require domain-specific expertise to understand; a particularly challenging task in a uniquely challenging context.

Our approach captures these complexities by including complete collections of related documents, preserving annotator disagreement, and using an annotation scheme covering the relevant legal frameworks. This provides the first full-document, disagreement-preserving dataset for privacy policy compliance analysis that addresses the inherent complexity and ambiguity characterizing real-world legal interpretation tasks.

3 Dataset Preparation

3.1 Document Selection

Privacy policies vary significantly between services and across markets (Marotta-Wurgler, 2016b). To account for this, we evaluate policies from multiple markets and across a range of traffic rankings. We grouped US-based, English-language websites by Tranco traffic rating (Pochat et al., 2018). We selected 40 English-language websites from each of six ranking ranges: $[1, 10]$; $[10, 1000]$; $[1000, 10k]$; $[10k, 100k]$; $[100k, 1M]$; and $[1M, \infty]$. For ease of comparison, we first sampled from the websites that either were used in OPP-115 (Wilson et al., 2016) or in the legal empirical study that provides a basis for our coding scheme (Marotta-Wurgler, 2016b). When necessary, we fell back to random selection from English-language websites within that range. We then discarded any websites that had no privacy policy or terms of service linked from their landing page. We also discarded all but the highest-ranking from any set of urls sharing a single policy (e.g., facebook.com and instagram.com linked to the same policy). The resulting sample contained 174 policies. Our final dataset contains all "top ten" websites in the sample, along with 142 additional policies selected at random from the sample. Our dataset includes websites tagged with 23 of the 26 "tier 1" market categories in the IAB Content Taxonomy 2.0 (IAB Tech Lab, 2024). The remaining three categories are illegal content, religious content, and websites primarily about pets. A full list of urls, category tags, and Tranco ranks is included in appendix D.

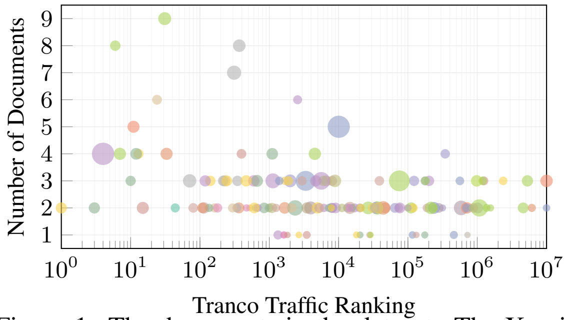

Figure 1: The documents in the dataset. The X axis shows the website's Tranco traffic rank. The Y axis reflects the number of documents collected for that site. Color and radius correspond to website industry (see Appendix D) and word count (ranging from 458 to $39,988$) respectively. An 18-document outlier is omitted from this chart.

Many privacy policies are spread across multiple linked documents. For example, the Amazon privacy policy reads: "*To enable our systems to recognize your browser or device and to provide and improve Amazon Services, we use cookies and other identifiers. For more information about cookies and how we use them, please read our Cookies Notice.*" The full terms regarding cookie use are found in the referenced "Cookies Notice." This is commonly referred to as “incorporation by reference.” Terms incorporated by reference are enforceable by courts as long as the referenced documents are clearly identified and accessible.(ALI, 2024). Failing to include those referenced documents can result in incomplete (and potentially incorrect) interpretations of the policy. To compile the full set of included terms, we downloaded each website's privacy policy and terms of service, then iteratively added any referenced documents. Figure 1 shows the distribution of policies by rank and the number of distinct documents associated with each policy after applying this procedure.

3.2 Coding Process

Each policy was coded independently by two law students who had completed relevant coursework on contracts and received training regarding legal interpretation of privacy policies. All coders were compensated for their time at the standard rate for J.D. student research assistants at their institution. We presented each coder with all relevant legal documents corresponding to a firms' website and multiple-choice questions about those documents. For each question, coders were asked to highlight any and all text (including text in documents incorporated by reference) they found relevant to answering the question. They answered each question by selecting from a set of choices describing possible policy content, including, when applicable, the possibility that the policy was silent on a question. They also recorded their confidence in their answer on a Likert scale.

Unlike many QA tasks, legal interpretation is inherently probabilistic. As the saying goes, a good lawyer always gives the same answer: "it depends" (a terrible QA task if taken literally). A single privacy policy may reasonably give rise to multiple valid interpretations. Accordingly, we intentionally preserve instances of inter-coder disagreement and low confidence, as these disagreements reflect real-world legal uncertainty rather than annotation error. Our approach ensures that AI models trained on this dataset will be evaluated based on their ability to navigate genuine ambiguity, rather than being forced into an artificial, deterministic labeling scheme. As described in the next section, this approach requires additional effort to ensure that disagreements stem from ambiguity in the policy text, and not from coder confusion with the question or the underlying regulatory context.

For the first ten weeks of coding, we held weekly overview meetings with each coder where we reviewed all codings. For each reviewed question where we identified a disagreement or error, we highlighted relevant text, and recorded our own answer, confidence, and notes. These reviews responses are included in the dataset for reference but are not used in this paper's summary statistics or benchmarks.

Coding, review, and project management were all performed using a suite of custom web tools we developed, as shown in the appendices. We provide a hosted version of the tools on our website. The tool includes several quality-of-life features, including offline mode, bookmarks and tabs in the document pane, and assignment and project management tools. We are releasing the tool as an open-source project in parallel with the dataset from this paper. The source code is available on our GitHub repository[1].

Coders self-directed their approach to coding. Most coders spread their coding tasks across multiple sessions; using a six-week sample, we can estimate the length of each session by observing the time that the tool was open and did not lose input focus for more than ten minutes. Ignoring three multi-day outliers, the median contiguous coding session lasted 19 minutes, and total time to code a policy ranged from 51 to 283 minutes over 1 to 13 sessions, the median coding time was 89 minutes per policy, and the median policy took 4.5 sessions to code.

3.3 Coding Schema

We generated our coding schema following the procedure developed by Marotta-Wurgler (Marotta-Wurgler, 2016a). The approach has been used in legal empirical scholarship to make quantitative comparisons of privacy policy content and compliance between industries and over time (Marotta-Wurgler, 2016b,c; Davis and Marotta-Wurgler, 2019). The schema represents a policy content as a set of labels derived from significant and influential privacy guidelines and applicable rules that have shaped the content and structure of privacy policies. These are: the 1973 HEW Fair Information Practice Principles, the 2012 Federal Trade Commission's Information Privacy Guidelines, the 2012 White House Privacy Bill of Rights, the GDPR of 2018, and the CCPA of 2020, and contract law—the background rules courts have employed to enforce privacy policy. The resulting coding provides a granular representation of privacy policy content mentioned in relevant guidelines and laws. For example, there are three labels that encode the rights users and firms have with respect to changes to the policy (can the firm make changes, does a user have to assent to that change before it takes effect, and are changes retroactive). Another label marks whether the set of documents includes a class action waiver. The goal of a granular approach was to minimize ambiguity in representation of policy content and enhance consistency among coders. We translate these variables into 64 multi-choice questions, which we

[1] https://github.com/document-coder/document-coder

group into 11 categories:

1. *CCPA* (10 labels): Tracks requirements unique to the California Consumer Protection Act, such as whether the subject can request that their personal information not be sold.
2. *GDPR* (5 labels): Tracks requirements unique to the GDPR, such as whether the entity has designated a Data Privacy Officer.
3. *Data Practices (DP)* (1 label): whether the firm has procedures to safely dispose of personal information.
4. *Enforcement (E)* (8 labels): Tracks mechanisms of legal redress.
5. *Notice (N)* (13 labels): Tracks notices pertaining to data collection and mandatory disclosures with state privacy laws.
6. *Contract (K)* (1 label): Tracks whether policy incorporates terms by reference.
7. *Privacy by Design (PBD)* (2 labels): Tracks general data practices in management and design.
8. *Security (SE)* (8 labels): Tracks information security practices.
9. *Sharing (SH)* (7 labels): Tracks sharing with third and other parties.
10. *User Control (UC)* (8 labels): Tracks user rights regarding personal information access and control.

For each question we drafted a set of answers designed to minimize ambiguity while providing granular representation of policy content. These include choice sets that are binary (*"Do the Terms of Use or Terms of Service incorporate the Privacy Policy by reference?" [yes/no]*), single-class (*"Does the Privacy Policy offer data requests by consumers explicitly free of charge?" [Not Applicable/Yes/No]*), and multi-class (*"Does the privacy policy provide means by which a user can contact the company with any privacy concerns or complaints? [select all that apply...]"*). Response options also include and distinguish between policies being silent regarding a term or the particular term being not applicable (e.g., a policy that states that no personal information is collected does not need to provide information about how such information is stored).

In contrast to other privacy policy datasets, which make efforts to maximize inter-coder agreement and often discard points of disagreement, we preserve disagreement and low-confidence coding. Because ambiguity is feature of many legal texts, the ground truth is effectively probabilistic, meaning disagreement and low confidence are expected features of a legal interpretation classification task. To ensure that disagreements and reported low confidence correspond to ambiguity in the policy rather than unclear coding instructions, we engaged in an iterative process to reduce exogenous sources of ambiguity from our coding. Note that the goal is to ensure annotators are consistent in their coding approach, which may not always imply consistency in their coding outcomes. Our goal is to maximize the likelihood that inconsistencies between annotations reflect ambiguities in the legal text, not in the annotators' understanding of their instructions.

Once a week during the 10-week iterative revision period, we met with coders and discussed each of their coding choices. The recorded highlights for each question helped coders recall and explain their decision-making. We qualitatively assessed the source of each instance of inter-coder disagreement or low reported confidence, choosing between five possible causes:

1. *Questions*: coders interpret the question in conflicting ways due to poor or confusing wording
2. *Choices*: the answer choice sets did not fully map onto the text and law
3. *Defaults*: the answer choice set does not properly account for the existence of default rules that alter the meaning of contractual silence
4. *Coder Error*: a coder made a mistake.
5. *Policy Text*: the text of the policy is ambiguous or susceptible to reasonable disagreements in interpretation

We addressed disagreements or confusion resulting from Categories 1, 2, and 3 by either adding clarifying details to the language of questions and answer choice sets, or adjusting the set of choices to better reflect the range of observed practices. When we detected coder error we corrected it and sent updated training guidance to other coders. We made no changes following disagreements and low confidence caused by unclear policy text. After any change to a question we removed any coding recorded using an outdated version of that question from the dataset. Coders relabeled those questions at the end of the iterative adjustment period.

Not every policy implicates the same questions or choices; some issues arose later in our revision

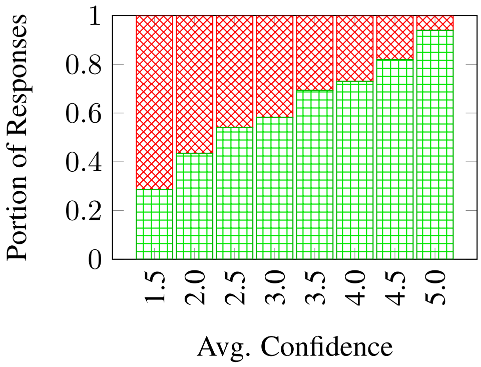

Figure 2: Pairwise agreement rate among coders by average self-reported confidence on a Likert scale.

window. After five rounds of iteration and revision, we stopped observing instances of the first three categories. The final set of coding instructions, including the history of changes corresponding to each variable, is included as an online appendix.

As an initial sanity check on our data, we compare confidence and inter-coder agreement rates, as shown in figure 2. We observe that inter-coder agreement has a roughly linear relationship with self-reported confidence. This suggests that both disagreement and confidence correspond with situations where coders see multiple potentially appropriate answers. Our efforts to remove other sources of confusion and our use of expert coders mean that this ambiguity should largely come from the text of the policy.

3.4 Insights from Iterative Schema Refinement

To test whether our coding scheme and iterative refinements actually reduce measurement errors by reducing exogenous sources of ambiguity, we included the OPP annotation scheme for the first five weeks of our iterative process. OPP is a natural baseline to measure against: it is cited as the current "gold standard" privacy policy annotation scheme (Mousavi Nejad et al., 2020), and has been used to train numerous automated privacy policy interpreters and included in the LEGALBENCH and PRIVACYGLUE composite legal reasoning benchmarks. While the OPP taxonomy was originally designed to annotate short phrases by their associated data practices, the taxonomy and dataset has been adapted for several other contexts and tasks.

While initial inter-coder agreement rates were similar between our questions and the OPP schema, we found that clarity issues in the OPP scheme held steady week-over-week while our scheme's error rate went down. This was true for both the classifications selected by coders, and the text coders marked as relevant to each question. Figure 3 shows the portion of questions changed after each week in our scheme and OPP. After four rounds of iterative adjustment, we removed the OPP annotations from our coding scheme. With enough iterative changes, both our scheme and OPP's would likely resolve; however, with each change we apply comparison with the original OPP-115 dataset and scheme becomes more difficult.

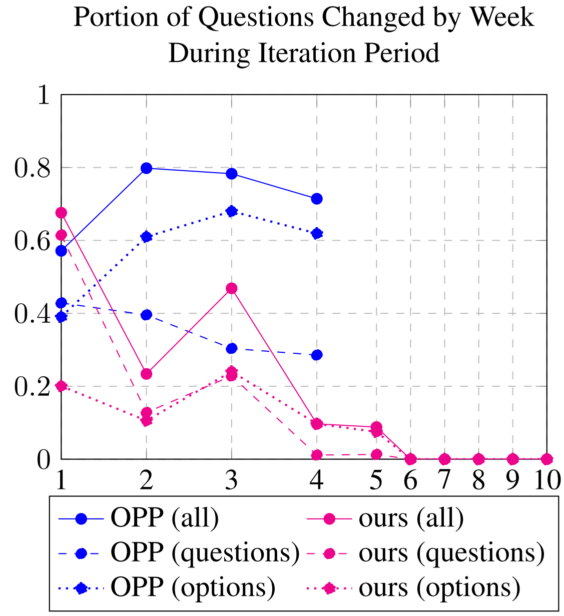

Figure 3: Portion of questions changed during iterative refinement, by coding scheme. OPP was removed from the scheme after week 5.

These results demonstrate how ambiguity and noise can come from the coding scheme or the way in which it is presented. While these results caution against uncritical reuse of the OPP taxonomy in new contexts, they do not apply directly to the OPP dataset. OPP was designed to pick out phrases disclosing data practices from short text samples; our scheme is designed to record the ways in which the full text of a privacy policy implicates the rights and responsibilities of firms and users.

We think that these differences in coder convergence between our scheme and OPP help justify our label selection method and iterative approach. There are two potential exogeneity worth addressing here. First, if training and regular meetings were only reason for convergence then we would expect all questions to converge. The OPP questions did not, which might suggest that our labels

Questions	64
Categories	11
Total Coders	18
Policies	149
Paragraphs	52,176
Words	1,524,570
Highlight Annotations	22,609
Policy Classifications	19,729
Confidence Scores	17,243

Table 1: Summary statistics on the corpus

correspond more closely to the terms contained in a privacy policy. Second, if the difference between the two schemes were the result of our initial drafting choices, we would expect to see lower rates of confusion at the outset. Instead, we see similar rates of initial confusion that improve once we detect and correct our drafting issues, further supporting the notion that our labels are better-aligned with policy content.

4 Dataset Contents

The dataset comprises annotations for 149 documents, each coded by two or more coders. For each document and question, the dataset contains classifications selected by each coder, a list of sentences the coder marked as relevant to answering the question, and their self-reported confidence ranked on a Likert scale. The dataset includes 64 variables motivated by 11 legal categories. For a subset of documents and questions, the dataset also includes the amount of time each coder spent answering the question. Coders were all upper-level law students who had completed course work covering the relevant topics in contract law. Table 1 provides additional descriptive statistics about the dataset.

Except for the CCPA category, correlation between question responses is low, as shown in Figure 4. Since answers to one question are not highly predictive of answers to another, we feel that this set of questions provides reasonable coverage over distinct legal interpretation tasks.

While achieving high inter-coder agreement was explicitly not our goal, the dataset exhibits a moderate level of inter-coder agreement, with Krippendorff's Alpha ranging from 0.40 to 0.96, averaging at 0.55. Absolute agreement rates across the dataset are 79%, with most disagreement concentrated in 7 questions. This may suggest that there are a few areas where firms are more likely to use opaque language or build flexibility into their terms.

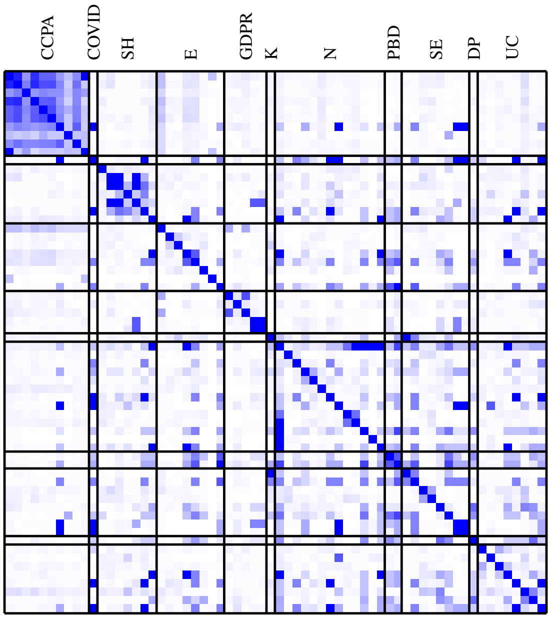

Figure 4: Correlation matrix showing correlation between question encodings. High correlation means that answers to questions covary by policy. For example, the high correlation between CCPA sections suggests that most policies are all-or-nothing along the CCPA-related dimensions we measure.

5 Results

Using our dataset, we evaluate current language models on their ability to perform two tasks. The first task ("holistic classification") is a multi-classification task that uses the entire policy as input: given our questions and a policy from our dataset, select the most likely answers for each question. The second task ("highlight prediction") is an annotation task that targets individual paragraphs: given a question from our dataset and a paragraph from one of the policies in our sample, predict whether a coder marked that paragraph as relevant to answering the question.

We evaluate the *holistic classification* task using batched cross-entropy loss. Each of the k policies in the dataset is associated with n sets of labels, $\{L_1, L_2, ..., L_n\}$, each corresponding to a question in our coding scheme. Each label L_i is associated with a set of options, $m_{i,j} \in M_i$. Given the ambiguity present in some privacy policies, the ground truth value of L_i^k may not be a single value, but rather a probability distribution over M_i. Coder responses are therefore definitionally noisy. We compute the goal probability distribution $\mathbf{y}_i^k$ as $([c_{i1}^k, ..., c_{im}^k])$ normalized

Average Cross Entropy Loss

Category	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	Gemini 2.0 Flash	random guesses
overall	0.174	0.309	0.175	0.217	0.177	0.456
CCPA	0.217	0.293	0.204	0.234	0.299	0.456
COVID	0.021	0.021	0.026	0.021	0.023	0.533
DP	0.084	0.293	0.214	0.266	0.060	0.497
E	0.092	0.215	0.118	0.158	0.148	0.464
GDPR	0.176	0.259	0.151	0.187	0.209	0.442
K	0.136	0.143	0.094	0.150	0.096	0.468
N	0.219	0.361	0.197	0.216	0.168	0.444
PBD	0.153	0.474	0.340	0.308	0.053	0.459
SE	0.154	0.295	0.134	0.233	0.137	0.445
SH	0.249	0.333	0.223	0.292	0.213	0.458
UC	0.125	0.377	0.160	0.196	0.133	0.467

Table 2: Average cross-entropy on holistic classification task, by category.

to sum to 1, where c_{ij}^k represents the number of coders who selected option j for question i on document k. We apply label smoothing to account for noise, as described in (Müller et al., 2020), setting $\alpha \in [.1, .3]$ as a function of average Likert score s_i^k, with greater smoothing for lower confidence: $\alpha_i^k = .3 - \frac{s_i^k}{25}$. We use LLMs to generate a probability distribution p_i^k over the set of options for each label. When logprobs are available, we generate the distribution by crawling the response tree of each branch until an answer is selected or the net probability is negligible. We evaluate model responses by computing cross-entropy loss between the model's response and the reference distribution, $\frac{1}{m_i}\sum_{j=1}^{m_i}(y_{ij}^k \log(p_{ij}^k)+(1-y_{ij}^k)\log(1-p_{ij}^k))$. We report average loss by question, category, and across the entire dataset.

Because some policies contain more than 32 thousand tokens, we can only test LLMs with sufficiently large context windows without resorting to context-expanding techniques or alternative models, which are out of scope for this project. The performance of some commercial LLMs with sufficiently large context windows appears in table 2. For some of the questions in our dataset, some models performed worse than random guessing, a result we found surprising. The errors appear to be caused by the models incorrectly selecting "not applicable" and "does not disclose" options far too often.

We evaluate *highlight predictions*, a binary classification task, by concatenating individual paragraphs with question text and option descriptions. $y_{ij} = 1$ if at least one coder flagged paragraph j as relevant when answering question i, and 0 otherwise. We tested several BERT models (Chalkidis et al., 2020; Devlin et al., 2018; Liu et al., 2019)—popular in legal application—using zero-shot labeling, prompting the model to answer whether the paragraph was relevant and computing the relative likelihood of an affirmative or negative response using the tree-crawling approach described above. We also tried cross-encoding and bi-encoding techniques using BGE, E5, and MiniLM (Xiao et al., 2023; Wang et al., 2020, 2022). The results are shown in table 3.

model	**prec.**	**recall**	f_1
Prompting			
BERT-BASE	1.00	18.05	1.89
RoBERTa-BASE	0.76	32.72	1.48
LEGAL-BERT	0.81	29.19	1.57
Cross-Encoding			
MiniLM L6	20.52	22.67	19.50
Bi-Encoding			
MiniLM L6	13.92	20.84	13.69
E5-base	10.76	20.00	10.84
BGE-base-en	14.45	20.82	14.07
GTE-base	6.70	17.36	7.54

Table 3: Average zero-shot performance on highlighting task, optimizing for f_1. Because highlighting is noisy and heavily skewed, we suspect a certain number of false positives are unavoidable. Current practical use-cases aim to identify relevant text from within policies (Lippi et al., 2019); setting a significantly higher β when computing thresholds (e.g., optimizing for f_{100} instead of f_1) can greatly increase recall.

We note that performance varies significantly across categories and questions, including questions within the same category. While differences in performance between models may be an artifact of our prompt design, we found the variance between similar questions about similar topics striking. At least for the systems we tested, an LLM's ability to answer one legal question appears to not be predictive of that LLM's ability to answer other questions, even within extremely narrow domains like "properties of sharing practices described within a privacy policy."

6 Future Directions

This project is designed to contribute to the growing body of legal task corpora. We plan to add it to open-source legal benchmarks, such as the LEGALBENCH consolidated corpus.

One of the challenges of analyzing legal documents is how work-intensive it is, how ephemeral some documents are, and how difficult it can be to comparing documents across time, especially if they incorporate changing (external) legal contexts by reference. By releasing these tools and putting greater emphasis on reproducibility, we plan to extend this dataset to observe how privacy policies change over time.

Our tools for classifying legal documents were intentionally designed to apply to other legal tasks, or to support future extensions of the question set and dataset. We hope to partner with other legal experts to expand this dataset to cover a broader range of legal questions and documents.

Finally, we have begun investigating the underlying cause of uneven rates of disagreement by question among coders (and, to a lesser extent, similarly uneven response rates between state-of-the-art LLMs). Future work may investigate whether firms are intentionally ambiguous to obscure practices or add flexibility, whether a mismatch between technology and law makes certain disclosure difficult or nonsensical, or whether some other factor is at play.

7 Conclusion

We have described a hand-coded dataset for the interpretation of privacy policies. It captures granular, multi-class data about 149 privacy policies and their associated documents along 64 dimensions, and is intended as a new resource for the development and benchmarking of NLP systems that interpret long legal text. Given the relatively poor performance of current state-of-the-art LLMs at task described in this dataset, we suspect that some aspects of the task that are not well represented by existing training data and benchmarks.

Our coding approach is designed to capture complexities inherent to the task of legal interpretation that are not present in current privacy policy datasets, such as addressing textual ambiguity, indeterminate meaning, interdependent clauses, contractual silence, and the effect of legal defaults. Along with our classification data, we include relevant-text annotations and confidence scores from each labeller. We supplement this dataset with our own coding of the questions where labellers disagree or report low confidence, which may provide additional insight into the textual ambiguities in the underlying policies. We include the tools we used to produce this dataset, including a hosted online tool that (non-technical) domain experts can use to produce similar classification datasets within their own areas of expertise.

Limitations

All but one of our population of coders learned consumer contracts (the relevant class for privacy policy interpretation) at the same law school from the same two law professors. They may have adopted some of those professor's biases, or approach contract interpretation in similar ways. That overlap may have obscured lingering ambiguities in our coding scheme. It may also have biased them towards understanding a coding scheme designed by one of those professors. We think the latter possibility is fairly remote–privacy policies are a sufficiently esoteric corner of contract law; it receives very little dedicated class time.

Decision fatigue is always a concern with labeling. Though our method does not take any explicit steps to mitigate fatigue, there are three reasons we don't think that particular issue is likely to meaningfully to effect data quality. First, reading and interpreting the clauses of a contract is a common task for a lawyer; one that our coders were recently evaluated on. Second, our coders spread most of their coding across multiple sessions, reducing the risk of decision fatigue. Third, in the weekly sessions where coders discussed the reasoning behind their decisions with us, we did not observe any qualitative difference in the level of care or depth of discussion surrounding any particular subset of the coding questions.

As noted above, the "ground truth" meaning of a contract can be probabilistic. Our coders effectively took a noisy sample of each contract with $n = 2$. We think this is reasonable for two reasons. First, the two-coder approach matches the current state of the art for privacy policy datasets. Second, confidence seems to be a decent predictor of disagreement, which opens mitigation options. We didn't, but potentially could, explore mitigation options.

The noisiness of our measurements also means that our benchmarks in part 5 necessarily contain a (somewhat arbitrary) smoothing factor. We suspect that specific tasks might be better measured using other metrics, and that the smoothing factor could be tuned to reflect confidence.

Likert scores are notoriously messy, meaning our confidence measurements may not contain as much information as we'd ideally like to capture.

Our reported benchmark performance rely on the quality of our prompt design. We have more experience designing prompts for LEGALBERT and GPT-4; our measurement of Claude 3 and other BERT models may be influenced by a prompt that is better suited for GPT. (our prompt generation script is included in the associated github repo).

References

Wasi Uddin Ahmad, Jianfeng Chi, Tu Le, Thomas Norton, Yuan Tian, and Kai-Wei Chang. 2021. Intent classification and slot filling for privacy policies.

Wasi Uddin Ahmad, Jianfeng Chi, Yuan Tian, and Kai-Wei Chang. 2020. Policyqa: A reading comprehension dataset for privacy policies.

ALI. 2024. Adoption of standard contract terms. In *Restatement of Consumer Contracts*. American Law Institute.

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*, pages 2165–2176.

Yannis Bakos, Florencia Marotta-Wurgler, and David R Trossen. 2014. Does anyone read the fine print? consumer attention to standard-form contracts. *The Journal of Legal Studies*, 43(1):1–35.

Michael Bommarito II and Daniel Martin Katz. 2022. Gpt takes the bar exam. *arXiv preprint arXiv:2212.14402*.

Duc Bui, Kang G. Shin, Jong-Min Choi, and Junbum Shin. 2021. Automated extraction and presentation of data practices in privacy policies. *Proceedings on Privacy Enhancing Technologies*, 2021(2):88–110.

Ilias Chalkidis, Manos Fergadiotis, Prodromos Malakasiotis, Nikolaos Aletras, and Ion Androutsopoulos. 2020. LEGAL-BERT: The muppets straight out of law school. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2898–2904, Online. Association for Computational Linguistics.

Ilias Chalkidis, Abhik Jana, Dirk Hartung, Michael Bommarito, Ion Androutsopoulos, Daniel Martin Katz, and Nikolaos Aletras. 2022. Lexglue: A benchmark dataset for legal language understanding in english.

Jonathan H Choi, Kristin E Hickman, Amy B Monahan, and Daniel Schwarcz. 2021. Chatgpt goes to law school. *J. Legal Educ.*, 71:387.

Jonathan H Choi and Daniel Schwarcz. 2023. Ai assistance in legal analysis: An empirical study. *Available at SSRN 4539836*.

Kevin E Davis and Florencia Marotta-Wurgler. 2019. Contracting for personal data. *NYUL Rev.*, 94:662.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.

Neel Guha, Julian Nyarko, Daniel E. Ho, Christopher Ré, Adam Chilton, Aditya Narayana, Alex Chohlas-Wood, Austin Peters, Brandon Waldon, Daniel N. Rockmore, Diego Zambrano, Dmitry Talisman, Enam Hoque, Faiz Surani, Frank Fagan, Galit Sarfaty, Gregory M. Dickinson, Haggai Porat, Jason Hegland, Jessica Wu, Joe Nudell, Joel Niklaus, John Nay, Jonathan H. Choi, Kevin Tobia, Margaret Hagan, Megan Ma, Michael Livermore, Nikon Rasumov-Rahe, Nils Holzenberger, Noam Kolt, Peter Henderson, Sean Rehaag, Sharad Goel, Shang Gao, Spencer Williams, Sunny Gandhi, Tom Zur, Varun Iyer, and Zehua Li. 2023. Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models.

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.

IAB Tech Lab. 2024. IAB Content Taxonomy 3.0.

Monika Jain, Raghava Mutharaju, Kuldeep Singh, and Ramakanth Kavuluru. 2024. Knowledge-driven cross-document relation extraction.

Nan-Jiang Jiang and Marie-Catherine de Marneffe. 2022. Investigating reasons for disagreement in natural language inference. *Transactions of the Association for Computational Linguistics*, 10:1357–1374.

Daniel Martin Katz, Michael James Bommarito, Shang Gao, and Pablo Arredondo. 2024. Gpt-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, 382(2270):20230254.

Marco Lippi, Przemysław Pałka, Giuseppe Contissa, Francesca Lagioia, Hans-Wolfgang Micklitz, Giovanni Sartor, and Paolo Torroni. 2019. Claudette: an automated detector of potentially unfair clauses in online terms of service. *Artificial Intelligence and Law*, 27:117–139.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692.

Florencia Marotta-Wurgler. 2016a. Self-regulation and competition in privacy policies. *The Journal of Legal Studies*, 45(S2):S13–S39.

Florencia Marotta-Wurgler. 2016b. Understanding Privacy Policies: Content, Self-Regulation, and Markets. SSRN Scholarly Paper ID 2736513, Rochester, NY.

Florencia Marotta-Wurgler. 2016c. Understanding privacy policies: Content, self-regulation, and markets. *NYU Law and Economics Research Paper*, (16-18).

Eric Martínez. 2024. Re-evaluating gpt-4's bar exam performance. *Artificial Intelligence and Law*, pages 1–24.

David Mellinkoff. 2004. *The language of the law*. Wipf and Stock Publishers.

Elizabeth Mertz. 2007. *The language of law school: learning to" think like a lawyer"*. Oxford University Press, USA.

Najmeh Mousavi Nejad, Pablo Jabat, Rostislav Nedelchev, Simon Scerri, and Damien Graux. 2020. Establishing a strong baseline for privacy policy classification. In *ICT Systems Security and Privacy Protection: 35th IFIP TC 11 International Conference, SEC 2020, Maribor, Slovenia, September 21–23, 2020, Proceedings 35*, pages 370–383. Springer.

Rafael Müller, Simon Kornblith, and Geoffrey Hinton. 2020. When does label smoothing help?

Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczyński, and Wouter Joosen. 2018. Tranco: A research-oriented top sites ranking hardened against manipulation. *arXiv preprint arXiv:1806.01156*.

Ellen Poplavska, Thomas B Norton, Shormir Wilson, and Norman Sadeh. 2020. From prescription to description: Mapping the gdpr to a privacy policy corpus annotation scheme. In *Legal Knowledge and Information Systems-JURIX 2020: 33rd Annual Conference*.

Abhilasha Ravichander, Alan W Black, Shomir Wilson, Thomas Norton, and Norman Sadeh. 2019. Question answering for privacy policies: Combining computational and legal perspectives.

Joel R Reidenberg, Jaspreet Bhatia, Travis D Breaux, and Thomas B Norton. 2016. Ambiguity in privacy policies and the impact of regulation. *The Journal of Legal Studies*, 45(S2):S163–S190.

Cristiana Santos, Arianna Rossi, Lorena Sanchez Chamorro, Kerstin Bongard-Blanchy, and Ruba Abu-Salma. 2021. Cookie banners, what's the purpose? analyzing cookie banner text through a legal lens. In *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society*, pages 187–194.

Atreya Shankar, Andreas Waldis, Christof Bless, Maria Andueza Rodriguez, and Luca Mazzola. 2023. Privacyglue: A benchmark dataset for general language understanding in privacy policies. *Applied Sciences*, 13(6):3701.

Mukund Srinath, Shomir Wilson, and C Lee Giles. 2021. Privacy at scale: Introducing the privaseer corpus of web privacy policies. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics.

Michiel van der Meer, Neele Falk, Pradeep K. Murukannaiah, and Enrico Liscio. 2024. Annotator-centric active learning for subjective NLP tasks. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 18537–18555, Miami, Florida, USA. Association for Computational Linguistics.

Liang Wang, Nan Yang, Xiaolong Huang, Binxing Jiao, Linjun Yang, Daxin Jiang, Rangan Majumder, and Furu Wei. 2022. Text embeddings by weakly-supervised contrastive pre-training. *arXiv preprint arXiv:2212.03533*.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. 2020. Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers.

Shomir Wilson, Florian Schaub, Aswarth Abhilash Dara, Frederick Liu, Sushain Cherivirala, Pedro Giovanni Leon, Mads Schaarup Andersen, Sebastian Zimmeck, Kanthashree Mysore Sathyendra, N Cameron Russell, et al. 2016. The creation and analysis of a website privacy policy corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1330–1340.

Shomir Wilson, Florian Schaub, Frederick Liu, Kanthashree Mysore Sathyendra, Daniel Smullen, Sebastian Zimmeck, Rohan Ramanath, Peter Story, Fei Liu, Norman Sadeh, and Noah A. Smith. 2018. Analyzing privacy policies at scale: From crowdsourcing to automated annotations. *ACM Trans. Web*, 13(1).

Shitao Xiao, Zheng Liu, Peitian Zhang, and Niklas Muennighoff. 2023. C-pack: Packaged resources to advance general chinese embedding.

Lucia Zheng, Neel Guha, Brandon R Anderson, Peter Henderson, and Daniel E Ho. 2021. When does pretraining help? assessing self-supervised learning for law and the casehold dataset of 53,000+ legal holdings. In *Proceedings of the eighteenth international* *conference on artificial intelligence and law*, pages 159–168.

A Question List

A.1 CCPA

Question ID	Question Text	Choices
CCPA-1	Is the CCPA section in a separate link (as opposed to in the same privacy policy?)	• No, it's part of the same privacy policy • Yes, it's on a separate link or separate document • Not applicable, there is no CCPA section or CCPA reference in the contract
CCPA-2	Does the Policy state the firms' CCPA policy only applies to California residents? (for example, Tinder's states that "This California section supplements the Privacy Policy and applies solely to California consumers (excluding our personnel). The Table below describes how we process California consumers' personal information (excluding our personnel), based on definitions laid out in the California Consumer Privacy Act ("CCPA")."	• No • Yes • Not applicable, there is no CCPA section or CCPA reference in the contract
CCPA-3	Has California Privacy Rights Section that explains all rights afforded under the CCPA? (The right to request disclosure of business' data collection and sales practices , the categories of personal information collected, the source of the information, use of the information and, if the information was disclosed or sold to third parties, the categories of personal information disclosed or sold to third parties and the categories of third parties to whom such information was disclosed or sold; The right to request a copy of the specific personal information collected about them during the 12 months before their request (together with right #1, a "personal information request"); The right to have such information deleted (with exceptions); he right to request that their personal information not be sold to third parties, if applicable; and The right not to be discriminated against because they exercised any of the new rights.]	• No • Yes, fully compliant • Partial compliance • Not applicable
CCPA-4	Directs CA Residents to that section when describing general (non-california exclusive) data practices?	• No • Yes • Not applicable, there is no CCPA section or CCPA reference in the contract
CCPA-5	Personal Information Request: Offers CA residents an opportunity to request all information shared with third parties in the last year?	• No or does not disclose • Yes • Not applicable
CCPA-6	Offers California residents a direct link via which to contact site and request information?	• contact info in different section • contact info in same section • Direct link included in same section) • Not applicable
CCPA-7	Data requests are explicitly free of charge?	• No • Yes • Not applicable
CCPA-8	Does the Policy list the categories of personal information sold in the past 12 months?	• No • Yes • Not applicable

CCPA-9	Policy identifies at least two methods for submitting a personal information or erasure request, in accordance with CCPA? (These must include, at a minimum, a web page and a toll-free telephone number)	• No • Two methods identified • One method identified • Two methods identified, but different from web page and toll-free number. • Not applicable
CCPA-10	Firm offers the right of opt-out of selling personal information to third parties with a visible, direct link to "Do Not Sell My Personal Information."	• No mention • Yes • Mentions the right but explains why it’s not available/not applicable (e.g., firm does not sell information to third parties) • Not applicable

A.2 Notice

Question ID	Question Text	Choices
N-1	Does the company have a cookie policy? (note "Yes" if the company has a stated cookie policy (e.g., a section in the privacy policy explaining its cookie policy) or offers link to a document with it, or if there is an attached cookie policy in the coding tool)	• No • Yes
N-2	Does the company explicitly state they use tracking elements other than cookies? (e.g. "local storage cookies", "browser fingerprints")?	• No • Yes • ’does not disclose’
N-3	Biometric Information Collected and Stored? (e.g., facial scans, fingerprints, facial patterns, voice or typing cadence)	• No • Yes • ’does not disclose’
N-4	Company commits that PII will be used internally only for business purposes (e.g., effecting, administering, or enforcing a transaction, sending future correspondence to user, research, internal database compilation, servicing website)? [Article 4 of the GDPR, personal data is "any information relating to an identified or identifiable natural person. In general, internal business purposes involve use of data for purposes that are in the service of the user)[Advertising is not considered internal business purposes; Article 6 of GDPR defines as legitimiate or "business purpose processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;	• No • Yes • Yes, but there are other terms in the policy that conflict with such commitment
N-5	Company commits to use PII (personally identifiable information) data only for stated context specific purposes. These are purposes that a user would expect in the context of the service provided. E.g., what you post to a message board must be made public for a message board to work)?	• No • Yes • Yes, but there are other terms in the policy that conflict with such commitment
N-6	Third party tracking: site allows third parties to place advertisements that may track user behavior?	• No • Yes • ’does not disclose’
N-7	Policy identifies third party recipients of shared or sold data? (General question)	• Generic identification ("trusted 3rd parties") • Specific/named entity or named category of third party • Does not disclose

N-8	Policy defines words such as "affiliates" or "third parties" if it uses them? (i.e.. the policy includes definitions of what it means by these categories; examples are not enough to classify as a definition)	• No • Yes • Not applicable
N-9	Policy has a Change of Terms (COT) provision that explicitly or implicitly allows entity to change agreement?	• No • Yes
N-10	Policy requires user to explicitly assent to material changes?	• No or does't say • Yes • Not applicable
N-11	Policy states that material changes are retroactive?	• No • Yes (applies to data made before the change)
N-12	Is there layered notice or short notice? (i.e., are the key Policy terms giving users the gist of the contract summarized on top of the document or somewhere noticeable? Table of Contents does not qualify)	• No • Yes
N-13	Policy provides notice of data procedures if company is sold or otherwise ceases to exist (e.g., goes bankrupt)?	• No • Yes

A.3 Sharing

Question ID	Question Text	Choices
SH-1	Are affiliates and/or subsidiaries (specifically) bound by the same privacy policy, confidentiality agreements, or have a contract with firm outlining how data will be used and secured?	• No • Yes • Does not disclose • Not applicable
SH-2	Are contractors/service providers (under CCPA)/processors (under GDPR) (e.g., payment process companies) bound by either the same privacy policy, confidentiality agreements, or have a contract with firm outlining how data will be used and secured? (CCPA, GDPR compliance requirement)	• No • Yes • Does not disclose • Not applicable
SH-3	Are third parties bound by the same privacy policy?	• No • Yes • Does not disclose • Not applicable [this would be applicable if, for example, the company did not share information with third parties]
SH-4	Entity performs due diligence to ensure legitimacy of 3rd parties that have access to data?	• No • Yes • Does not disclose
SH-5	Entity has contract with 3rd parties (excluding processors/service providers) establishing how disclosed data can be used?	• No • Yes • Does not disclose • Not applicable

SH-6	Policy provides links to relevant 3rd parties' privacy policies? (sometimes the Policy includes links to third party privacy policies, usually when it states that any engagement with third parties will be governed by third party privacy policies	• No • Yes • Not applicable (in cases where there are no relationships with third parties of any sort)
SH-7	What is consent mechanism for sharing/selling PII or sensitive information to entities that aren't service providers (except for the purpose of effecting, administering, or enforcing a transaction, sending future correspondence to user, research, internal database compilation, servicing website)?	• Mandatory or does not disclose • Opt-out [The default is for the user to share PII with third parties, but the Policy gives the user the opportunity to opt out] • Opt-in [User must consent before data can be shared with or collected/used by third party] • Not applicable (applicable if firm does not engage in this practice)

A.4 User Control

Question ID	Question Text	Choices
UC-1	Can the user request that incorrect data be either rectified, updated, or erased?	• No • Yes, within 30 days [in compliance with GDPR] • Yes, no time limit • Does not disclose
UC-2	Can users can adjust privacy settings? [Double check on the website; directing user to control cookies via browser setting doesn't count]	• No • Yes
UC-3	Are users allowed to access and correct/update personal data collected?	• No • Can access data • Can access and correct data • Can access and correct data, and 3rd parties notified of correction
UC-4	Can user request that information be deleted or anonymized?	• No • Yes (partial) [User can delete/anonymize account, but the company/organization may continue to keep some of the user's data] • Yes (full) [User can delete account and all of the user's information is removed from company/organization's servers/databases.]

UC-5	Ownership Rights of User Information Provided (look in terms of use)	• Company owns data • User owns data, but licenses data to entity in a non-exclusive, royalty-free, form, with no right to compensation for company's use of the data • User owns the data but the license is not so broad as to permit the company to use, share, and commercialize data/proprietary media as it sees fit • Does not disclose
UC-6	What happens to data if entity ceases to exist or is acquired?	• Sold with company or otherwise distributed/disclosed [e.g., when the company liquidates and the assets are sold piecemal] • Sold but given continued protection under same Policy, or transferred to acquiring entity • Destroyed, anonymized, etc. • User is given choice as to what happens with data • Does not disclose
UC-7	If company is sold or goes bankrupt, user is given choice as to what happens to their data?	• No • Yes
UC-8	If user quits site, what happens to personal data?	• Retained and treated as if user is still using service • Retained but modified • Deleted/anonymized • Does not disclose

A.5 Security

Question ID	Question Text	Choices
SE-1	Policy guarantees data accuracy (must say the word "data accuracy" for this to be relevant)?	• No • Yes
SE-2	Policy specifies reasonable procedures in place to ensure accuracy? (Even a little procedure qualifies)	• No • Yes
SE-3	Policy reserves right to disclose protected information to comply with law/prevent crime?	• No or doesn't say • Yes
SE-4	Policy reserves right to disclose protected information to protect its own rights?	• No or doesn't say • Yes
SE-5	Users will be given notice of government requests for information about the user.	• No or doesn't say • Yes
SE-6	User will be notified of data breach?	• No or doesn't say • Yes

SE-7	Policy describes substantive privacy and security protections incorporated into entity's managerial/structural procedures (e.g., limiting the number of employees who have access to data, allowing data access only for job-related functions, assigning employees to oversee privacy issues, employing Chief Privacy Officer, requiring periodic audits)? (General question)	• No • Yes
SE-8	Policy specifically identifies means of technological security (e.g., encryption)?	• No • Yes

A.6 Data Practices

Question ID	Question Text	Choices
DP-1	Does company have a procedure for safely disposing unused/no longer needed data?	• No or doesn't say • Yes

A.7 Enforcement

Question ID	Question Text	Choices
E-1	Policy provides means by which user can contact site with privacy concerns and/or complaints? [select all that apply]	• No • Yes, for all users • yes, and mentions it in accordance with CCPA [or listed under a CCPA section in Policy] • yes, and mentions it in accordance with GDPR [or listed under a GDPR section in Policy]
E-2	Policy has forum selection clause? If so, which forum?	• No • Yes
E-3	Policy has choice of law clause? If so, which law?	• No • Yes
E-4	Policy has arbitration clause?	• No • Yes • consumer may choose arbitration
E-5	Policy has class action waiver?	• No • Yes
E-6	Policy disclaims liability for failure of security measures?	• No • Yes
E-7	Policy provides link to FTC's Consumer Complaint Form and/or the FTC telephone number? [this just asks if the Policy mentions the FTC's consumer complaint form at all or provides links to it]	• No • Yes

E-8	What privacy seal/certification/industry oversight organization does Policy claim [other than mandatory international law (Swiss Privacy Law, etc)? [Privacy Seals are independent, third-party enforcement programs to monitor company practices and enforce privacy policies. They are designed to provide protection to consumers by allowing Web companies to standardize privacy policies. Privacy seal programs include, among others, TRUSTe, BBBOnline, and CPA Webtrust. These are different from regulatory compliance seals, such as those that the company complies with COPPA, the Children Online Privacy Protection Act).	• • Name of certification or seal

A.8 Privacy By Design

Question ID	Question Text	Choices
PBD-1	Policy requires periodic compliance review of structural and technological data security measures?	• No • Yes
PBD-2	Policy contains self-reporting measures in case of privacy violation (to a privacy seal organization, 3rd party consultant)?	• No • Yes

A.9 Contract

Question ID	Question Text	Choices
K-1	Is the Privacy Policy incorporated by reference in the Terms of use?	• No • Yes

A.10 GDPR

Question ID	Question Text	Choices
GDPR-1	Policy states that it is GDPR compliant or includes section on GDPR compliance	• No or no mention • Yes
GDPR-2	Does Policy state it complies with EU-US Privacy Shield?	• No • Yes
GDPR-3	Does Policy state that GDPR terms apply only and exclusively to EU residents?	• No • Yes • Not applicable (no GDPR terms)
GDPR-4	Can users object to processing or automated decision making that could impact them? (This is only applicable if company does profiling or any other automated decision making, such as algorithmic decision making, or any automated decisions that don't involve a human)	• No • Yes • Not applicable • Does not disclose
GDPR-5	If firm engages in automated decision making, does it provide meaningful information about the logic involved, or significance/effect of such decisions?	• No • Yes • Not applicable • Does not disclose

A.11 COVID

Question ID	Question Text	Choices

COVID-1	Does the Policy include any terms related to contact tracing, health tracking, or other terms in relationship to COVID?	• No • Yes

B LLM Performance on Holistic Reading Task Benchmark

The following tables shows the binary cross entropy between coding produced by several commercial LLMs and the coding provided in our dataset on a per-question and per-category basis.

We used the following prompt across all models, adjusting data format to match each LLM's API. As with any LLM, additional or model-specific prompt design might result in significantly different results.

```
{
  messages: {
    role: "system",
    content: [
      {
        type: "text",
        text: """You are an AI assistant tasked with answering multiple-choice
            questions about a collection of legal documents that comprise a website'
            s privacy policy.
        You respond with a JSON object containing two fields: "answer" (a character
            or list of characters) and "confidence" (an integer between 1-5, where 1
              is least confident, and 5 is most confident)."""
      }, {type: "text", text: (first document as markdown)}, {type: "text", text: (
          second document as markdown)}, ...
    ]
  }, {
    role: "user",
    content: [
      {
        type:"text",
        text: """<question>
        (content of the question)
         - "A": (first option)
         - "B": (second option)
          ...
        </question>
        <instructions>
        Select the best answer for the provided question
        </instructions>"""
      }
    ]
  },
  return_type: {
    type: "json_schema",
    json_schema: {
    schema: {
      answer: {type: "string", pattern: "^[A-Z]$"},
      "confidence": {"type": "integer", "enum": [1,2,3,4,5]}
    },
  },
  temperature: 0
}
```

B.1 Performance by Category

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
CCPA	0.217	0.293	0.204	0.234	0.299	0.456
COVID	0.021	0.021	0.026	0.021	0.023	0.533
DP	0.084	0.293	0.214	0.266	0.060	0.497
E	0.092	0.215	0.118	0.158	0.148	0.464
GDPR	0.176	0.259	0.151	0.187	0.209	0.442
K	0.136	0.143	0.094	0.150	0.096	0.468
N	0.219	0.361	0.197	0.216	0.168	0.444
PBD	0.153	0.474	0.340	0.308	0.053	0.459
SE	0.154	0.295	0.134	0.233	0.137	0.445
SH	0.249	0.333	0.223	0.292	0.213	0.458
UC	0.125	0.377	0.160	0.196	0.133	0.467

B.2 Performance by Question

CCPA

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
CCPA-1	0.063	0.108	0.073	0.091	0.206	0.436
CCPA-2	0.191	0.357	0.098	0.103	0.458	0.469
CCPA-3	0.226	0.246	0.200	0.219	0.234	0.456
CCPA-4	0.179	0.289	0.157	0.275	0.439	0.474
CCPA-5	0.314	0.314	0.311	0.312	0.417	0.421
CCPA-6	0.157	0.344	0.150	0.199	0.181	0.500
CCPA-7	0.291	0.347	0.288	0.303	0.239	0.456
CCPA-8	0.324	0.316	0.321	0.337	0.353	0.432
CCPA-9	0.187	0.273	0.185	0.212	0.291	0.468
CCPA-10	0.212	0.324	0.224	0.257	0.188	0.445

Enforcement

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
E-1	0.120	0.209	0.147	0.160	0.199	0.465
E-2	0.184	0.177	0.255	0.218	0.208	0.465
E-3	0.055	0.089	0.221	0.075	0.202	0.456
E-4	0.054	0.178	0.044	0.105	0.136	0.506
E-5	0.082	0.236	0.017	0.207	0.106	0.442
E-6	0.140	0.219	0.142	0.161	0.277	0.421
E-7	0.006	0.021	0.008	0.008	0.007	0.482
E-8	0.092	0.584	0.113	0.332	0.056	0.474

GDPR

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
GDPR-1	0.101	0.156	0.085	0.085	0.196	0.461
GDPR-2	0.038	0.038	0.031	0.036	0.044	0.413
GDPR-3	0.274	0.478	0.234	0.244	0.493	0.464
GDPR-4	0.252	0.347	0.223	0.314	0.133	0.427
GDPR-5	0.208	0.275	0.179	0.244	0.193	0.443

Notice

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
N-1	0.094	0.097	0.097	0.090	0.209	0.484
N-2	0.134	0.138	0.110	0.262	0.253	0.453
N-3	0.249	0.326	0.252	0.337	0.405	0.448
N-4	0.481	0.438	0.473	0.325	0.118	0.396
N-5	0.500	0.446	0.459	0.433	0.156	0.414
N-6	0.102	0.138	0.091	0.106	0.238	0.501
N-7	0.274	0.345	0.279	0.268	0.170	0.394
N-8	0.105	0.405	0.056	0.074	0.055	0.444
N-9	0.035	0.049	0.028	0.049	0.224	0.476
N-10	0.172	0.453	0.104	0.198	0.077	0.482
N-11	0.063	0.685	0.057	0.057	0.058	0.413
N-12	0.198	0.562	0.113	0.136	0.078	0.422
N-13	0.444	0.592	0.444	0.481	0.147	0.443

Privacy By Design

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
PBD-1	0.199	0.556	0.603	0.543	0.065	0.448
PBD-2	0.107	0.393	0.078	0.073	0.042	0.470

Security

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
SE-1	0.072	0.058	0.052	0.045	0.040	0.528
SE-2	0.632	0.673	0.388	0.619	0.162	0.351
SE-3	0.057	0.077	0.024	0.044	0.247	0.439
SE-4	0.085	0.154	0.085	0.132	0.220	0.452
SE-5	0.020	0.223	0.041	0.264	0.035	0.485
SE-6	0.066	0.612	0.093	0.335	0.054	0.437
SE-7	0.215	0.409	0.308	0.315	0.195	0.429
SE-8	0.078	0.152	0.078	0.112	0.139	0.440

Sharing

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4	GPT-4o	Gemini 2.0 Flash	random
SH-1	0.235	0.205	0.202	0.190	0.148	0.459
SH-2	0.153	0.212	0.163	0.197	0.159	0.460
SH-3	0.148	0.222	0.150	0.159	0.158	0.478
SH-4	0.173	0.561	0.176	0.433	0.317	0.458
SH-5	0.344	0.305	0.261	0.308	0.355	0.445
SH-6	0.374	0.485	0.294	0.428	0.178	0.464
SH-7	0.313	0.344	0.314	0.331	0.171	0.445

B.3 Performance by Category

Category	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
CCPA	0.217	0.293	0.204	0.234	0.447
COVID	0.021	0.021	0.026	0.021	0.463
DP	0.084	0.293	0.214	0.266	0.466
E	0.092	0.215	0.118	0.158	0.469
GDPR	0.176	0.259	0.151	0.187	0.443
K	0.136	0.143	0.094	0.150	0.427
N	0.219	0.361	0.197	0.216	0.457
PBD	0.153	0.474	0.340	0.308	0.505
SE	0.154	0.295	0.134	0.233	0.466
SH	0.249	0.333	0.223	0.292	0.457
UC	0.125	0.377	0.160	0.196	0.451

B.4 CCPA

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
CCPA-1	0.063	0.108	0.073	0.091	0.469
CCPA-2	0.191	0.357	0.098	0.103	0.527
CCPA-3	0.226	0.246	0.200	0.219	0.427
CCPA-4	0.179	0.289	0.157	0.275	0.455
CCPA-5	0.314	0.314	0.311	0.312	0.423
CCPA-6	0.157	0.344	0.150	0.199	0.444
CCPA-7	0.291	0.347	0.288	0.303	0.450
CCPA-8	0.324	0.316	0.321	0.337	0.404
CCPA-9	0.187	0.273	0.185	0.212	0.464
CCPA-10	0.212	0.324	0.224	0.257	0.420

B.5 Enforcement

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
E-1	0.120	0.209	0.147	0.160	0.436
E-2	0.184	0.177	0.255	0.218	0.471
E-3	0.055	0.089	0.221	0.075	0.532
E-4	0.054	0.178	0.044	0.105	0.478
E-5	0.082	0.236	0.017	0.207	0.476
E-6	0.140	0.219	0.142	0.161	0.401
E-7	0.006	0.021	0.008	0.008	0.472
E-8	0.092	0.584	0.113	0.332	0.489

B.6 GDPR

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
GDPR-1	0.101	0.156	0.085	0.085	0.462
GDPR-2	0.038	0.038	0.031	0.036	0.427
GDPR-3	0.274	0.478	0.234	0.244	0.423
GDPR-4	0.252	0.347	0.223	0.314	0.462
GDPR-5	0.208	0.275	0.179	0.244	0.436

B.7 Notice

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
N-1	0.094	0.097	0.097	0.090	0.458
N-2	0.134	0.138	0.110	0.262	0.444
N-3	0.249	0.326	0.252	0.337	0.423
N-4	0.481	0.438	0.473	0.325	0.486
N-5	0.500	0.446	0.459	0.433	0.447
N-6	0.102	0.138	0.091	0.106	0.464
N-7	0.274	0.345	0.279	0.268	0.405
N-8	0.105	0.405	0.056	0.074	0.466
N-9	0.035	0.049	0.028	0.049	0.499
N-10	0.172	0.453	0.104	0.198	0.489
N-11	0.063	0.685	0.057	0.057	0.464
N-12	0.198	0.562	0.113	0.136	0.447
N-13	0.444	0.592	0.444	0.481	0.444

B.8 Privacy By Design

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
PBD-1	0.199	0.556	0.603	0.543	0.522
PBD-2	0.107	0.393	0.078	0.073	0.489

B.9 Security

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
SE-1	0.072	0.058	0.052	0.045	0.481
SE-2	0.632	0.673	0.388	0.619	0.434
SE-3	0.057	0.077	0.024	0.044	0.470
SE-4	0.085	0.154	0.085	0.132	0.445
SE-5	0.020	0.223	0.041	0.264	0.520
SE-6	0.066	0.612	0.093	0.335	0.474
SE-7	0.215	0.409	0.308	0.315	0.440
SE-8	0.078	0.152	0.078	0.112	0.466

B.10 Contract

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
K-1	0.136	0.143	0.094	0.150	0.427

B.11 Sharing

Question ID	Claude 3.5 "Sonnet"	Claude 3 "Haiku"	GPT-4o	GPT-4	random guesses
SH-1	0.235	0.205	0.202	0.190	0.450
SH-2	0.153	0.212	0.163	0.197	0.454
SH-3	0.148	0.222	0.150	0.159	0.482
SH-4	0.173	0.561	0.176	0.433	0.487
SH-5	0.344	0.305	0.261	0.308	0.430
SH-6	0.374	0.485	0.294	0.428	0.436
SH-7	0.313	0.344	0.314	0.331	0.462

C Results for Highlighting Task

C.1 BERT

C.1.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	11.90	75.52	1.07	1.36
COVID (COVID)	0.87	72.88	0.09	0.16
Contract (K)	7.16	96.45	1.31	2.04
Data Practices (DP)	0.68	98.95	0.68	0.68
Enforcement (E)	22.05	66.11	1.29	2.01
GDPR (GDPR)	5.52	88.37	0.84	1.00
Notice (N)	35.57	54.21	1.67	2.77
Privacy By Design (PBD)	2.18	95.92	0.30	0.50
Security (SE)	21.60	74.69	1.37	1.93
Sharing (SH)	6.16	95.36	1.71	2.23
User Control (UC)	17.63	82.95	1.19	1.95

C.1.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	1.01	99.24	0.76	0.81
CCPA-2	17.49	68.96	0.37	0.72
CCPA-3	32.97	9.78	1.05	1.98
CCPA-4	34.98	8.96	0.23	0.45
CCPA-5	4.14	98.15	2.13	1.96
CCPA-6	6.62	95.17	1.33	1.80
CCPA-7	0.67	97.67	0.34	0.45
CCPA-8	13.81	84.77	1.04	1.80
CCPA-9	4.47	96.82	2.76	2.75
CCPA-10	2.82	95.71	0.65	0.86

C.1.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.87	72.88	0.09	0.16

C.1.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	0.68	98.95	0.68	0.68

C.1.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	41.56	69.30	2.73	4.86
E-2	4.79	96.54	1.47	2.17
E-3	12.90	90.14	1.03	1.79
E-4	54.11	0.56	0.56	1.09
E-5	39.04	0.25	0.25	0.50
E-6	18.10	89.53	3.26	4.39
E-7	1.34	96.77	0.75	0.81
E-8	4.52	85.82	0.26	0.46

C.1.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	14.73	60.31	0.52	0.96
GDPR-2	2.53	97.94	1.36	1.36
GDPR-3	1.88	98.49	1.69	1.55
GDPR-4	4.76	89.32	0.33	0.59
GDPR-5	3.69	95.80	0.31	0.56

C.1.7 Contract (K)

Question ID	r	a	p	f_1
K-1	7.16	96.45	1.31	2.04

C.1.8 Notice (N)

Question ID	r	a	p	f_1
N-1	22.24	86.37	3.18	4.76
N-2	14.95	94.53	4.26	5.88
N-3	4.14	96.60	1.16	1.73
N-4	66.21	30.03	2.38	4.33
N-5	28.83	63.00	1.68	2.95
N-6	28.57	88.11	2.96	5.02
N-7	90.10	2.61	1.95	3.75
N-8	27.92	6.56	0.43	0.82
N-9	50.06	59.04	1.33	2.45
N-10	56.60	21.14	0.67	1.29
N-11	51.35	3.80	0.41	0.79
N-12	16.50	62.15	0.75	1.37
N-13	4.92	90.78	0.50	0.85

C.1.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	1.34	97.45	0.34	0.54
PBD-2	3.02	94.39	0.26	0.46

C.1.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	2.04	95.83	1.04	1.25
SE-2	3.24	98.49	2.85	2.77
SE-3	88.59	0.98	0.98	1.92
SE-4	27.91	67.06	0.97	1.84
SE-5	16.38	55.57	0.31	0.60
SE-6	1.01	99.62	1.01	1.01
SE-7	12.11	94.21	2.39	3.62
SE-8	21.48	85.79	1.36	2.44

C.1.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	4.82	96.45	1.09	1.59
SH-2	4.50	97.14	1.63	2.08
SH-3	8.02	95.04	2.13	3.18
SH-4	0.68	99.59	0.68	0.68
SH-5	12.93	86.74	0.84	1.53
SH-6	4.70	98.71	3.21	3.51
SH-7	7.44	93.86	2.40	3.03

C.1.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	19.30	80.65	1.04	1.92
UC-2	16.76	75.88	0.94	1.59
UC-3	10.99	92.91	1.76	2.76
UC-4	40.99	68.18	1.39	2.64
UC-5	5.36	94.70	2.14	2.53
UC-6	17.23	83.10	0.69	1.30
UC-7	14.14	81.27	0.58	1.06
UC-8	16.26	86.90	0.99	1.79

C.1.13 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	1.01	99.24	0.76	0.81
CCPA-2	17.49	68.96	0.37	0.72
CCPA-3	32.97	9.78	1.05	1.98
CCPA-4	34.98	8.96	0.23	0.45
CCPA-5	4.14	98.15	2.13	1.96
CCPA-6	6.62	95.17	1.33	1.80
CCPA-7	0.67	97.67	0.34	0.45
CCPA-8	13.81	84.77	1.04	1.80
CCPA-9	4.47	96.82	2.76	2.75
CCPA-10	2.82	95.71	0.65	0.86

C.1.14 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.87	72.88	0.09	0.16

C.1.15 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	0.68	98.95	0.68	0.68

C.1.16 Enforcement (E)

Question ID	r	a	p	f_1
E-1	41.56	69.30	2.73	4.86
E-2	4.79	96.54	1.47	2.17
E-3	12.90	90.14	1.03	1.79
E-4	54.11	0.56	0.56	1.09
E-5	39.04	0.25	0.25	0.50
E-6	18.10	89.53	3.26	4.39
E-7	1.34	96.77	0.75	0.81
E-8	4.52	85.82	0.26	0.46

C.1.17 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	14.73	60.31	0.52	0.96
GDPR-2	2.53	97.94	1.36	1.36
GDPR-3	1.88	98.49	1.69	1.55
GDPR-4	4.76	89.32	0.33	0.59
GDPR-5	3.69	95.80	0.31	0.56

C.1.18 Contract (K)

Question ID	r	a	p	f_1
K-1	7.16	96.45	1.31	2.04

C.1.19 Notice (N)

Question ID	r	a	p	f_1
N-1	22.24	86.37	3.18	4.76
N-2	14.95	94.53	4.26	5.88
N-3	4.14	96.60	1.16	1.73
N-4	66.21	30.03	2.38	4.33
N-5	28.83	63.00	1.68	2.95
N-6	28.57	88.11	2.96	5.02
N-7	90.10	2.61	1.95	3.75
N-8	27.92	6.56	0.43	0.82
N-9	50.06	59.04	1.33	2.45
N-10	56.60	21.14	0.67	1.29
N-11	51.35	3.80	0.41	0.79
N-12	16.50	62.15	0.75	1.37
N-13	4.92	90.78	0.50	0.85

C.1.20 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	1.34	97.45	0.34	0.54
PBD-2	3.02	94.39	0.26	0.46

C.1.21 Security (SE)

Question ID	r	a	p	f_1
SE-1	2.04	95.83	1.04	1.25
SE-2	3.24	98.49	2.85	2.77
SE-3	88.59	0.98	0.98	1.92
SE-4	27.91	67.06	0.97	1.84
SE-5	16.38	55.57	0.31	0.60
SE-6	1.01	99.62	1.01	1.01
SE-7	12.11	94.21	2.39	3.62
SE-8	21.48	85.79	1.36	2.44

C.1.22 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	4.82	96.45	1.09	1.59
SH-2	4.50	97.14	1.63	2.08
SH-3	8.02	95.04	2.13	3.18
SH-4	0.68	99.59	0.68	0.68
SH-5	12.93	86.74	0.84	1.53
SH-6	4.70	98.71	3.21	3.51
SH-7	7.44	93.86	2.40	3.03

C.1.23 User Control (UC)

Question ID	r	a	p	f_1
UC-1	19.30	80.65	1.04	1.92
UC-2	16.76	75.88	0.94	1.59
UC-3	10.99	92.91	1.76	2.76
UC-4	40.99	68.18	1.39	2.64
UC-5	5.36	94.70	2.14	2.53
UC-6	17.23	83.10	0.69	1.30
UC-7	14.14	81.27	0.58	1.06
UC-8	16.26	86.90	0.99	1.79

C.1.24 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	1.01	99.24	0.76	0.81
CCPA-2	17.49	68.96	0.37	0.72
CCPA-3	32.97	9.78	1.05	1.98
CCPA-4	34.98	8.96	0.23	0.45
CCPA-5	4.14	98.15	2.13	1.96
CCPA-6	6.62	95.17	1.33	1.80
CCPA-7	0.67	97.67	0.34	0.45
CCPA-8	13.81	84.77	1.04	1.80
CCPA-9	4.47	96.82	2.76	2.75
CCPA-10	2.82	95.71	0.65	0.86

C.1.25 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.87	72.88	0.09	0.16

C.1.26 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	0.68	98.95	0.68	0.68

C.1.27 Enforcement (E)

Question ID	r	a	p	f_1
E-1	41.56	69.30	2.73	4.86
E-2	4.79	96.54	1.47	2.17
E-3	12.90	90.14	1.03	1.79
E-4	54.11	0.56	0.56	1.09
E-5	39.04	0.25	0.25	0.50
E-6	18.10	89.53	3.26	4.39
E-7	1.34	96.77	0.75	0.81
E-8	4.52	85.82	0.26	0.46

C.1.28 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	14.73	60.31	0.52	0.96
GDPR-2	2.53	97.94	1.36	1.36
GDPR-3	1.88	98.49	1.69	1.55
GDPR-4	4.76	89.32	0.33	0.59
GDPR-5	3.69	95.80	0.31	0.56

C.1.29 Contract (K)

Question ID	r	a	p	f_1
K-1	7.16	96.45	1.31	2.04

C.1.30 Notice (N)

Question ID	r	a	p	f_1
N-1	22.24	86.37	3.18	4.76
N-2	14.95	94.53	4.26	5.88
N-3	4.14	96.60	1.16	1.73
N-4	66.21	30.03	2.38	4.33
N-5	28.83	63.00	1.68	2.95
N-6	28.57	88.11	2.96	5.02
N-7	90.10	2.61	1.95	3.75
N-8	27.92	6.56	0.43	0.82
N-9	50.06	59.04	1.33	2.45
N-10	56.60	21.14	0.67	1.29
N-11	51.35	3.80	0.41	0.79
N-12	16.50	62.15	0.75	1.37
N-13	4.92	90.78	0.50	0.85

C.1.31 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	1.34	97.45	0.34	0.54
PBD-2	3.02	94.39	0.26	0.46

C.1.32 Security (SE)

Question ID	r	a	p	f_1
SE-1	2.04	95.83	1.04	1.25
SE-2	3.24	98.49	2.85	2.77
SE-3	88.59	0.98	0.98	1.92
SE-4	27.91	67.06	0.97	1.84
SE-5	16.38	55.57	0.31	0.60
SE-6	1.01	99.62	1.01	1.01
SE-7	12.11	94.21	2.39	3.62
SE-8	21.48	85.79	1.36	2.44

C.1.33 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	4.82	96.45	1.09	1.59
SH-2	4.50	97.14	1.63	2.08
SH-3	8.02	95.04	2.13	3.18
SH-4	0.68	99.59	0.68	0.68
SH-5	12.93	86.74	0.84	1.53
SH-6	4.70	98.71	3.21	3.51
SH-7	7.44	93.86	2.40	3.03

C.1.34 User Control (UC)

Question ID	r	a	p	f_1
UC-1	19.30	80.65	1.04	1.92
UC-2	16.76	75.88	0.94	1.59
UC-3	10.99	92.91	1.76	2.76
UC-4	40.99	68.18	1.39	2.64
UC-5	5.36	94.70	2.14	2.53
UC-6	17.23	83.10	0.69	1.30
UC-7	14.14	81.27	0.58	1.06
UC-8	16.26	86.90	0.99	1.79

C.2 ROBERTA

C.2.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	17.12	68.77	0.83	1.36
COVID (COVID)	1.30	91.32	0.32	0.50
Contract (K)	15.86	90.92	1.38	2.33
Data Practices (DP)	7.88	50.77	0.16	0.31
Enforcement (E)	42.60	36.84	0.95	1.71
GDPR (GDPR)	20.40	27.81	0.25	0.48
Notice (N)	47.60	33.23	1.14	2.00
Privacy By Design (PBD)	5.15	69.26	0.46	0.47
Security (SE)	37.01	42.46	0.71	1.21
Sharing (SH)	31.92	40.67	0.95	1.60
User Control (UC)	38.31	44.63	1.11	1.75

C.2.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	7.20	89.23	0.87	1.38
CCPA-2	3.66	96.75	1.71	2.06
CCPA-3	20.93	54.38	1.36	2.40
CCPA-4	5.75	92.67	0.39	0.64
CCPA-5	13.57	73.27	0.54	0.94
CCPA-6	29.68	61.16	0.73	1.39
CCPA-7	4.36	81.90	0.35	0.60
CCPA-8	44.52	9.18	0.55	1.06
CCPA-9	29.69	56.67	1.02	1.85
CCPA-10	11.86	72.53	0.73	1.23

C.2.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	1.30	91.32	0.32	0.50

C.2.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	7.88	50.77	0.16	0.31

C.2.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	52.18	61.99	3.51	6.06
E-2	76.03	1.15	0.73	1.40
E-3	83.56	0.99	0.74	1.43
E-4	54.11	1.26	0.56	1.10
E-5	7.36	87.14	0.41	0.74
E-6	50.45	44.11	1.36	2.41
E-7	1.01	96.83	0.15	0.24
E-8	16.11	1.22	0.15	0.30

C.2.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	39.12	1.81	0.48	0.93
GDPR-2	2.96	87.61	0.25	0.40
GDPR-3	27.88	21.21	0.29	0.57
GDPR-4	18.61	17.14	0.16	0.32
GDPR-5	13.42	11.28	0.08	0.15

C.2.7 Contract (K)

Question ID	r	a	p	f_1
K-1	15.86	90.92	1.38	2.33

C.2.8 Notice (N)

Question ID	r	a	p	f_1
N-1	46.96	48.99	1.57	2.82
N-2	64.54	15.30	1.07	2.04
N-3	20.47	37.93	0.70	1.13
N-4	79.62	18.31	2.44	4.52
N-5	10.83	86.05	2.17	2.93
N-6	68.84	24.89	1.14	2.18
N-7	90.60	1.93	1.93	3.72
N-8	25.94	11.50	0.42	0.80
N-9	87.56	9.94	0.99	1.92
N-10	56.04	22.79	0.67	1.28
N-11	36.49	24.14	0.43	0.82
N-12	2.48	89.71	0.92	1.07
N-13	28.41	40.48	0.39	0.77

C.2.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	8.39	46.11	0.12	0.24
PBD-2	1.90	92.42	0.81	0.69

C.2.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	30.78	26.10	0.39	0.76
SE-2	3.91	96.60	0.88	1.33
SE-3	88.48	1.55	0.98	1.91
SE-4	66.26	26.34	0.87	1.69
SE-5	28.67	1.53	0.25	0.50
SE-6	2.48	89.38	0.35	0.57
SE-7	72.15	2.26	0.76	1.49
SE-8	3.36	95.90	1.20	1.46

C.2.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	28.85	54.32	1.10	1.84
SH-2	7.77	95.28	1.70	2.36
SH-3	30.39	52.01	1.04	1.87
SH-4	22.30	1.15	0.16	0.31
SH-5	57.05	1.83	0.47	0.93
SH-6	15.17	72.32	1.26	2.05
SH-7	61.93	7.80	0.94	1.82

C.2.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	12.29	87.34	2.29	2.82
UC-2	29.20	60.40	1.01	1.87
UC-3	14.92	85.32	2.46	3.33
UC-4	25.14	61.20	1.09	1.99
UC-5	72.11	0.58	0.58	1.15
UC-6	61.45	17.31	0.49	0.96
UC-7	57.24	4.67	0.43	0.86
UC-8	34.12	40.21	0.51	0.99

C.2.13 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	7.20	89.23	0.87	1.38
CCPA-2	3.66	96.75	1.71	2.06
CCPA-3	20.93	54.38	1.36	2.40
CCPA-4	5.75	92.67	0.39	0.64
CCPA-5	13.57	73.27	0.54	0.94
CCPA-6	29.68	61.16	0.73	1.39
CCPA-7	4.36	81.90	0.35	0.60
CCPA-8	44.52	9.18	0.55	1.06
CCPA-9	29.69	56.67	1.02	1.85
CCPA-10	11.86	72.53	0.73	1.23

C.2.14 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	1.30	91.32	0.32	0.50

C.2.15 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	7.88	50.77	0.16	0.31

C.2.16 Enforcement (E)

Question ID	r	a	p	f_1
E-1	52.18	61.99	3.51	6.06
E-2	76.03	1.15	0.73	1.40
E-3	83.56	0.99	0.74	1.43
E-4	54.11	1.26	0.56	1.10
E-5	7.36	87.14	0.41	0.74
E-6	50.45	44.11	1.36	2.41
E-7	1.01	96.83	0.15	0.24
E-8	16.11	1.22	0.15	0.30

C.2.17 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	39.12	1.81	0.48	0.93
GDPR-2	2.96	87.61	0.25	0.40
GDPR-3	27.88	21.21	0.29	0.57
GDPR-4	18.61	17.14	0.16	0.32
GDPR-5	13.42	11.28	0.08	0.15

C.2.18 Contract (K)

Question ID	r	a	p	f_1
K-1	15.86	90.92	1.38	2.33

C.2.19 Notice (N)

Question ID	r	a	p	f_1
N-1	46.96	48.99	1.57	2.82
N-2	64.54	15.30	1.07	2.04
N-3	20.47	37.93	0.70	1.13
N-4	79.62	18.31	2.44	4.52
N-5	10.83	86.05	2.17	2.93
N-6	68.84	24.89	1.14	2.18
N-7	90.60	1.93	1.93	3.72
N-8	25.94	11.50	0.42	0.80
N-9	87.56	9.94	0.99	1.92
N-10	56.04	22.79	0.67	1.28
N-11	36.49	24.14	0.43	0.82
N-12	2.48	89.71	0.92	1.07
N-13	28.41	40.48	0.39	0.77

C.2.20 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	8.39	46.11	0.12	0.24
PBD-2	1.90	92.42	0.81	0.69

C.2.21 Security (SE)

Question ID	r	a	p	f_1
SE-1	30.78	26.10	0.39	0.76
SE-2	3.91	96.60	0.88	1.33
SE-3	88.48	1.55	0.98	1.91
SE-4	66.26	26.34	0.87	1.69
SE-5	28.67	1.53	0.25	0.50
SE-6	2.48	89.38	0.35	0.57
SE-7	72.15	2.26	0.76	1.49
SE-8	3.36	95.90	1.20	1.46

C.2.22 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	28.85	54.32	1.10	1.84
SH-2	7.77	95.28	1.70	2.36
SH-3	30.39	52.01	1.04	1.87
SH-4	22.30	1.15	0.16	0.31
SH-5	57.05	1.83	0.47	0.93
SH-6	15.17	72.32	1.26	2.05
SH-7	61.93	7.80	0.94	1.82

C.2.23 User Control (UC)

Question ID	r	a	p	f_1
UC-1	12.29	87.34	2.29	2.82
UC-2	29.20	60.40	1.01	1.87
UC-3	14.92	85.32	2.46	3.33
UC-4	25.14	61.20	1.09	1.99
UC-5	72.11	0.58	0.58	1.15
UC-6	61.45	17.31	0.49	0.96
UC-7	57.24	4.67	0.43	0.86
UC-8	34.12	40.21	0.51	0.99

C.2.24 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	7.20	89.23	0.87	1.38
CCPA-2	3.66	96.75	1.71	2.06
CCPA-3	20.93	54.38	1.36	2.40
CCPA-4	5.75	92.67	0.39	0.64
CCPA-5	13.57	73.27	0.54	0.94
CCPA-6	29.68	61.16	0.73	1.39
CCPA-7	4.36	81.90	0.35	0.60
CCPA-8	44.52	9.18	0.55	1.06
CCPA-9	29.69	56.67	1.02	1.85
CCPA-10	11.86	72.53	0.73	1.23

C.2.25 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	1.30	91.32	0.32	0.50

C.2.26 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	7.88	50.77	0.16	0.31

C.2.27 Enforcement (E)

Question ID	r	a	p	f_1
E-1	52.18	61.99	3.51	6.06
E-2	76.03	1.15	0.73	1.40
E-3	83.56	0.99	0.74	1.43
E-4	54.11	1.26	0.56	1.10
E-5	7.36	87.14	0.41	0.74
E-6	50.45	44.11	1.36	2.41
E-7	1.01	96.83	0.15	0.24
E-8	16.11	1.22	0.15	0.30

C.2.28 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	39.12	1.81	0.48	0.93
GDPR-2	2.96	87.61	0.25	0.40
GDPR-3	27.88	21.21	0.29	0.57
GDPR-4	18.61	17.14	0.16	0.32
GDPR-5	13.42	11.28	0.08	0.15

C.2.29 Contract (K)

Question ID	r	a	p	f_1
K-1	15.86	90.92	1.38	2.33

C.2.30 Notice (N)

Question ID	r	a	p	f_1
N-1	46.96	48.99	1.57	2.82
N-2	64.54	15.30	1.07	2.04
N-3	20.47	37.93	0.70	1.13
N-4	79.62	18.31	2.44	4.52
N-5	10.83	86.05	2.17	2.93
N-6	68.84	24.89	1.14	2.18
N-7	90.60	1.93	1.93	3.72
N-8	25.94	11.50	0.42	0.80
N-9	87.56	9.94	0.99	1.92
N-10	56.04	22.79	0.67	1.28
N-11	36.49	24.14	0.43	0.82
N-12	2.48	89.71	0.92	1.07
N-13	28.41	40.48	0.39	0.77

C.2.31 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	8.39	46.11	0.12	0.24
PBD-2	1.90	92.42	0.81	0.69

C.2.32 Security (SE)

Question ID	r	a	p	f_1
SE-1	30.78	26.10	0.39	0.76
SE-2	3.91	96.60	0.88	1.33
SE-3	88.48	1.55	0.98	1.91
SE-4	66.26	26.34	0.87	1.69
SE-5	28.67	1.53	0.25	0.50
SE-6	2.48	89.38	0.35	0.57
SE-7	72.15	2.26	0.76	1.49
SE-8	3.36	95.90	1.20	1.46

C.2.33 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	28.85	54.32	1.10	1.84
SH-2	7.77	95.28	1.70	2.36
SH-3	30.39	52.01	1.04	1.87
SH-4	22.30	1.15	0.16	0.31
SH-5	57.05	1.83	0.47	0.93
SH-6	15.17	72.32	1.26	2.05
SH-7	61.93	7.80	0.94	1.82

C.2.34 User Control (UC)

Question ID	r	a	p	f_1
UC-1	12.29	87.34	2.29	2.82
UC-2	29.20	60.40	1.01	1.87
UC-3	14.92	85.32	2.46	3.33
UC-4	25.14	61.20	1.09	1.99
UC-5	72.11	0.58	0.58	1.15
UC-6	61.45	17.31	0.49	0.96
UC-7	57.24	4.67	0.43	0.86
UC-8	34.12	40.21	0.51	0.99

C.3 LEGALBERT

C.3.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	24.39	52.87	0.68	1.10
COVID (COVID)	4.03	1.83	0.04	0.09
Contract (K)	46.91	37.83	0.74	1.42
Data Practices (DP)	16.22	0.79	0.13	0.25
Enforcement (E)	37.92	50.50	0.98	1.67
GDPR (GDPR)	4.89	97.02	1.30	1.69
Notice (N)	37.91	41.06	1.20	1.99
Privacy By Design (PBD)	1.68	89.18	0.14	0.24
Security (SE)	30.77	54.70	0.94	1.45
Sharing (SH)	37.81	41.77	0.93	1.61
User Control (UC)	27.80	64.96	1.28	2.10

C.3.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	10.70	72.81	0.38	0.70
CCPA-2	12.75	80.07	0.49	0.93
CCPA-3	28.24	17.43	1.03	1.92
CCPA-4	1.06	99.58	1.41	1.17
CCPA-5	4.27	94.59	0.67	1.01
CCPA-6	49.20	18.13	0.50	0.97
CCPA-7	3.36	95.95	0.68	1.05
CCPA-8	48.32	0.76	0.52	1.01
CCPA-9	53.46	33.27	0.89	1.71
CCPA-10	32.55	16.06	0.26	0.52

C.3.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	4.03	1.83	0.04	0.09

C.3.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	16.22	0.79	0.13	0.25

C.3.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	71.68	29.93	2.09	3.96
E-2	4.93	97.53	2.30	2.93
E-3	71.35	14.03	0.77	1.47
E-4	21.65	68.99	0.73	1.36
E-5	39.73	0.77	0.26	0.52
E-6	89.86	1.73	1.14	2.21
E-7	0.67	96.15	0.07	0.12
E-8	3.47	94.85	0.49	0.81

C.3.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	10.94	95.94	3.37	3.92
GDPR-2	2.26	98.29	1.00	1.28
GDPR-3	7.14	94.29	1.20	1.92
GDPR-4	3.45	97.55	0.79	1.11
GDPR-5	0.67	99.03	0.13	0.22

C.3.7 Contract (K)

Question ID	r	a	p	f_1
K-1	46.91	37.83	0.74	1.42

C.3.8 Notice (N)

Question ID	r	a	p	f_1
N-1	17.31	83.89	2.14	3.10
N-2	73.28	1.14	0.93	1.80
N-3	3.69	90.39	0.42	0.74
N-4	90.66	3.06	2.24	4.22
N-5	64.36	5.09	1.38	2.59
N-6	82.55	1.17	1.02	1.98
N-7	32.08	70.54	2.25	3.95
N-8	28.86	0.60	0.40	0.77
N-9	3.38	96.32	1.97	2.23
N-10	14.43	84.85	1.18	1.99
N-11	4.84	84.78	0.87	1.03
N-12	24.41	11.48	0.41	0.80
N-13	53.02	0.45	0.35	0.69

C.3.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	2.35	79.79	0.17	0.30
PBD-2	1.01	98.57	0.11	0.19

C.3.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	1.93	97.70	1.44	1.49
SE-2	27.99	54.67	0.54	1.04
SE-3	87.79	3.07	1.00	1.95
SE-4	83.22	1.05	0.76	1.49
SE-5	29.37	1.77	0.26	0.51
SE-6	2.36	95.75	0.50	0.80
SE-7	7.89	93.89	2.14	2.94
SE-8	5.59	89.71	0.94	1.38

C.3.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	68.03	1.00	0.76	1.47
SH-2	72.97	0.89	0.67	1.32
SH-3	15.95	86.43	1.75	2.62
SH-4	15.54	30.35	0.17	0.33
SH-5	56.38	4.65	0.47	0.93
SH-6	5.26	95.62	1.13	1.77
SH-7	30.55	73.44	1.55	2.79

C.3.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	24.89	82.20	2.26	3.17
UC-2	38.86	29.30	0.70	1.34
UC-3	13.69	90.19	2.45	3.85
UC-4	13.48	92.16	1.87	3.01
UC-5	13.04	91.10	1.32	2.29
UC-6	53.85	40.74	0.59	1.15
UC-7	59.31	0.43	0.43	0.84
UC-8	5.26	93.57	0.64	1.11

C.3.13 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	10.70	72.81	0.38	0.70
CCPA-2	12.75	80.07	0.49	0.93
CCPA-3	28.24	17.43	1.03	1.92
CCPA-4	1.06	99.58	1.41	1.17
CCPA-5	4.27	94.59	0.67	1.01
CCPA-6	49.20	18.13	0.50	0.97
CCPA-7	3.36	95.95	0.68	1.05
CCPA-8	48.32	0.76	0.52	1.01
CCPA-9	53.46	33.27	0.89	1.71
CCPA-10	32.55	16.06	0.26	0.52

C.3.14 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	4.03	1.83	0.04	0.09

C.3.15 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	16.22	0.79	0.13	0.25

C.3.16 Enforcement (E)

Question ID	r	a	p	f_1
E-1	71.68	29.93	2.09	3.96
E-2	4.93	97.53	2.30	2.93
E-3	71.35	14.03	0.77	1.47
E-4	21.65	68.99	0.73	1.36
E-5	39.73	0.77	0.26	0.52
E-6	89.86	1.73	1.14	2.21
E-7	0.67	96.15	0.07	0.12
E-8	3.47	94.85	0.49	0.81

C.3.17 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	10.94	95.94	3.37	3.92
GDPR-2	2.26	98.29	1.00	1.28
GDPR-3	7.14	94.29	1.20	1.92
GDPR-4	3.45	97.55	0.79	1.11
GDPR-5	0.67	99.03	0.13	0.22

C.3.18 Contract (K)

Question ID	r	a	p	f_1
K-1	46.91	37.83	0.74	1.42

C.3.19 Notice (N)

Question ID	r	a	p	f_1
N-1	17.31	83.89	2.14	3.10
N-2	73.28	1.14	0.93	1.80
N-3	3.69	90.39	0.42	0.74
N-4	90.66	3.06	2.24	4.22
N-5	64.36	5.09	1.38	2.59
N-6	82.55	1.17	1.02	1.98
N-7	32.08	70.54	2.25	3.95
N-8	28.86	0.60	0.40	0.77
N-9	3.38	96.32	1.97	2.23
N-10	14.43	84.85	1.18	1.99
N-11	4.84	84.78	0.87	1.03
N-12	24.41	11.48	0.41	0.80
N-13	53.02	0.45	0.35	0.69

C.3.20 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	2.35	79.79	0.17	0.30
PBD-2	1.01	98.57	0.11	0.19

C.3.21 Security (SE)

Question ID	r	a	p	f_1
SE-1	1.93	97.70	1.44	1.49
SE-2	27.99	54.67	0.54	1.04
SE-3	87.79	3.07	1.00	1.95
SE-4	83.22	1.05	0.76	1.49
SE-5	29.37	1.77	0.26	0.51
SE-6	2.36	95.75	0.50	0.80
SE-7	7.89	93.89	2.14	2.94
SE-8	5.59	89.71	0.94	1.38

C.3.22 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	68.03	1.00	0.76	1.47
SH-2	72.97	0.89	0.67	1.32
SH-3	15.95	86.43	1.75	2.62
SH-4	15.54	30.35	0.17	0.33
SH-5	56.38	4.65	0.47	0.93
SH-6	5.26	95.62	1.13	1.77
SH-7	30.55	73.44	1.55	2.79

C.3.23 User Control (UC)

Question ID	r	a	p	f_1
UC-1	24.89	82.20	2.26	3.17
UC-2	38.86	29.30	0.70	1.34
UC-3	13.69	90.19	2.45	3.85
UC-4	13.48	92.16	1.87	3.01
UC-5	13.04	91.10	1.32	2.29
UC-6	53.85	40.74	0.59	1.15
UC-7	59.31	0.43	0.43	0.84
UC-8	5.26	93.57	0.64	1.11

C.3.24 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	10.70	72.81	0.38	0.70
CCPA-2	12.75	80.07	0.49	0.93
CCPA-3	28.24	17.43	1.03	1.92
CCPA-4	1.06	99.58	1.41	1.17
CCPA-5	4.27	94.59	0.67	1.01
CCPA-6	49.20	18.13	0.50	0.97
CCPA-7	3.36	95.95	0.68	1.05
CCPA-8	48.32	0.76	0.52	1.01
CCPA-9	53.46	33.27	0.89	1.71
CCPA-10	32.55	16.06	0.26	0.52

C.3.25 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	4.03	1.83	0.04	0.09

C.3.26 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	16.22	0.79	0.13	0.25

C.3.27 Enforcement (E)

Question ID	r	a	p	f_1
E-1	71.68	29.93	2.09	3.96
E-2	4.93	97.53	2.30	2.93
E-3	71.35	14.03	0.77	1.47
E-4	21.65	68.99	0.73	1.36
E-5	39.73	0.77	0.26	0.52
E-6	89.86	1.73	1.14	2.21
E-7	0.67	96.15	0.07	0.12
E-8	3.47	94.85	0.49	0.81

C.3.28 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	10.94	95.94	3.37	3.92
GDPR-2	2.26	98.29	1.00	1.28
GDPR-3	7.14	94.29	1.20	1.92
GDPR-4	3.45	97.55	0.79	1.11
GDPR-5	0.67	99.03	0.13	0.22

C.3.29 Contract (K)

Question ID	r	a	p	f_1
K-1	46.91	37.83	0.74	1.42

C.3.30 Notice (N)

Question ID	r	a	p	f_1
N-1	17.31	83.89	2.14	3.10
N-2	73.28	1.14	0.93	1.80
N-3	3.69	90.39	0.42	0.74
N-4	90.66	3.06	2.24	4.22
N-5	64.36	5.09	1.38	2.59
N-6	82.55	1.17	1.02	1.98
N-7	32.08	70.54	2.25	3.95
N-8	28.86	0.60	0.40	0.77
N-9	3.38	96.32	1.97	2.23
N-10	14.43	84.85	1.18	1.99
N-11	4.84	84.78	0.87	1.03
N-12	24.41	11.48	0.41	0.80
N-13	53.02	0.45	0.35	0.69

C.3.31 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	2.35	79.79	0.17	0.30
PBD-2	1.01	98.57	0.11	0.19

C.3.32 Security (SE)

Question ID	r	a	p	f_1
SE-1	1.93	97.70	1.44	1.49
SE-2	27.99	54.67	0.54	1.04
SE-3	87.79	3.07	1.00	1.95
SE-4	83.22	1.05	0.76	1.49
SE-5	29.37	1.77	0.26	0.51
SE-6	2.36	95.75	0.50	0.80
SE-7	7.89	93.89	2.14	2.94
SE-8	5.59	89.71	0.94	1.38

C.3.33 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	68.03	1.00	0.76	1.47
SH-2	72.97	0.89	0.67	1.32
SH-3	15.95	86.43	1.75	2.62
SH-4	15.54	30.35	0.17	0.33
SH-5	56.38	4.65	0.47	0.93
SH-6	5.26	95.62	1.13	1.77
SH-7	30.55	73.44	1.55	2.79

C.3.34 User Control (UC)

Question ID	r	a	p	f_1
UC-1	24.89	82.20	2.26	3.17
UC-2	38.86	29.30	0.70	1.34
UC-3	13.69	90.19	2.45	3.85
UC-4	13.48	92.16	1.87	3.01
UC-5	13.04	91.10	1.32	2.29
UC-6	53.85	40.74	0.59	1.15
UC-7	59.31	0.43	0.43	0.84
UC-8	5.26	93.57	0.64	1.11

C.4 MiniLM L6 (cross encoding)

C.4.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	28.92	99.34	22.64	24.54
Notice (N)	22.31	98.07	23.16	18.64
Sharing (SH)	15.86	99.44	15.86	15.11
User Control (UC)	24.53	99.22	24.08	23.61
Security (SE)	20.21	99.43	18.15	17.27
Data Practices (DP)	7.41	99.88	14.29	9.76
Enforcement (E)	20.20	98.92	14.01	15.29
Privacy By Design (PBD)	23.48	97.62	8.08	7.67
Contract (K)	29.55	99.44	34.51	31.84
GDPR (GDPR)	31.42	99.73	31.97	29.63
COVID (COVID)	0.00	99.81	0.00	0.00

C.4.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	14.67	99.30	7.01	9.48
CCPA-2	33.78	99.67	33.78	33.78
CCPA-3	27.37	98.40	22.29	24.57
CCPA-4	53.57	99.47	18.40	27.40
CCPA-5	32.94	99.53	25.45	28.72
CCPA-6	17.24	99.15	11.17	13.56
CCPA-7	30.00	99.88	39.13	33.96
CCPA-8	24.78	99.39	22.40	23.53
CCPA-9	10.37	98.84	7.87	8.95
CCPA-10	44.44	99.74	38.89	41.48

C.4.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.00	99.81	0.00	0.00

C.4.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	7.41	99.88	14.29	9.76

C.4.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	12.85	96.86	11.55	12.16
E-2	0.88	99.57	5.26	1.52
E-3	0.00	99.58	0.00	0.00
E-4	31.25	99.38	34.62	32.85
E-5	49.18	99.70	34.09	40.27
E-6	34.45	97.96	13.21	19.10
E-7	12.50	98.55	0.23	0.46
E-8	20.51	99.72	13.11	16.00

C.4.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	10.34	99.52	23.08	14.29
GDPR-2	52.78	99.76	26.03	34.86
GDPR-3	22.45	99.71	18.64	20.37
GDPR-4	26.09	99.73	20.69	23.08
GDPR-5	45.45	99.95	71.43	55.56

C.4.7 Contract (K)

Question ID	r	a	p	f_1
K-1	29.55	99.44	34.51	31.84

C.4.8 Notice (N)

Question ID	r	a	p	f_1
N-1	20.38	97.96	11.62	14.80
N-2	32.82	99.27	24.86	28.29
N-3	9.23	99.80	85.71	16.67
N-4	32.60	86.83	3.53	6.37
N-5	4.80	98.76	6.75	5.61
N-6	27.53	99.07	24.75	26.06
N-7	20.00	97.53	16.33	17.98
N-8	11.58	99.21	6.79	8.56
N-9	30.25	98.50	12.73	17.92
N-10	21.05	99.66	42.55	28.17
N-11	44.26	99.74	38.57	41.22
N-12	1.75	99.03	1.10	1.36
N-13	33.80	99.61	25.81	29.27

C.4.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	13.64	99.88	15.79	14.63
PBD-2	33.33	95.35	0.36	0.71

C.4.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	27.42	99.33	9.94	14.59
SE-2	14.77	99.59	21.67	17.57
SE-3	24.61	99.18	31.97	27.81
SE-4	12.59	99.43	28.12	17.39
SE-5	25.00	99.49	9.30	13.56
SE-6	33.33	99.70	15.58	21.24
SE-7	9.52	99.24	12.96	10.98
SE-8	14.44	99.51	15.66	15.03

C.4.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	27.73	99.32	21.85	24.44
SH-2	12.71	99.41	17.05	14.56
SH-3	15.33	99.36	22.11	18.10
SH-4	22.22	99.76	10.53	14.29
SH-5	16.47	99.62	24.56	19.72
SH-6	11.27	99.39	6.35	8.12
SH-7	5.30	99.24	8.60	6.56

C.4.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	28.97	99.24	25.30	27.01
UC-2	12.40	99.03	8.25	9.91
UC-3	22.22	98.50	13.86	17.07
UC-4	26.88	99.21	33.56	29.85
UC-5	21.09	99.13	14.44	17.14
UC-6	35.79	99.71	56.67	43.87
UC-7	34.25	99.66	31.65	32.89
UC-8	14.61	99.30	8.97	11.11

C.5 MiniLM L6 (bi-encoding)

C.5.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	20.84	99.27	15.59	16.52
Notice (N)	21.42	98.37	15.73	13.51
Sharing (SH)	12.63	99.03	6.25	7.27
User Control (UC)	21.22	99.15	18.10	19.18
Security (SE)	11.65	99.27	15.85	11.38
Data Practices (DP)	37.04	99.54	7.75	12.82
Enforcement (E)	38.51	90.31	10.52	14.06
Privacy By Design (PBD)	4.55	99.28	1.05	1.71
Contract (K)	36.36	97.28	6.18	10.56
GDPR (GDPR)	20.93	99.72	22.96	20.07
COVID (COVID)	0.00	99.63	0.00	0.00

C.5.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	13.33	99.31	6.67	8.89
CCPA-2	31.08	99.72	41.82	35.66
CCPA-3	26.67	98.46	23.17	24.80
CCPA-4	28.57	99.46	11.68	16.58
CCPA-5	25.88	99.40	16.06	19.82
CCPA-6	12.93	99.09	8.11	9.97
CCPA-7	16.67	99.68	6.49	9.35
CCPA-8	7.08	99.54	20.00	10.46
CCPA-9	4.88	98.53	2.75	3.52
CCPA-10	41.27	99.51	19.12	26.13

C.5.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.00	99.63	0.00	0.00

C.5.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	37.04	99.54	7.75	12.82

C.5.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	20.75	88.15	3.24	5.60
E-2	67.26	69.94	0.84	1.66
E-3	16.96	98.73	6.21	9.09
E-4	69.44	99.29	37.31	48.54
E-5	37.70	99.63	24.21	29.49
E-6	33.49	97.87	12.32	18.02
E-7	62.50	69.02	0.05	0.11
E-8	0.00	99.83	0.00	0.00

C.5.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	8.62	99.46	15.38	11.05
GDPR-2	16.67	99.78	14.29	15.38
GDPR-3	34.69	99.67	20.00	25.37
GDPR-4	17.39	99.84	44.44	25.00
GDPR-5	27.27	99.87	20.69	23.53

C.5.7 Contract (K)

Question ID	r	a	p	f_1
K-1	36.36	97.28	6.18	10.56

C.5.8 Notice (N)

Question ID	r	a	p	f_1
N-1	23.08	98.26	15.83	18.78
N-2	26.72	99.06	15.84	19.89
N-3	4.62	99.79	75.00	8.70
N-4	20.44	95.02	6.75	10.14
N-5	4.80	98.31	3.68	4.17
N-6	21.35	99.04	20.43	20.88
N-7	20.49	96.29	9.55	13.03
N-8	5.26	99.30	4.00	4.55
N-9	52.47	96.21	7.46	13.06
N-10	24.21	98.89	8.13	12.17
N-11	42.62	99.55	20.97	28.11
N-12	0.00	99.59	0.00	0.00
N-13	32.39	99.46	16.91	22.22

C.5.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	9.09	99.62	2.11	3.42
PBD-2	0.00	98.93	0.00	0.00

C.5.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	1.61	99.75	7.14	2.63
SE-2	13.64	99.41	10.62	11.94
SE-3	10.47	99.31	35.71	16.19
SE-4	18.18	99.11	14.86	16.35
SE-5	6.25	99.69	5.88	6.06
SE-6	22.22	99.87	42.11	29.09
SE-7	19.73	97.39	4.20	6.92
SE-8	1.11	99.65	6.25	1.89

C.5.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	29.41	97.55	5.12	8.72
SH-2	11.02	99.40	15.12	12.75
SH-3	29.20	98.32	8.99	13.75
SH-4	3.70	99.78	2.50	2.99
SH-5	11.76	99.30	6.90	8.70
SH-6	0.00	99.63	0.00	0.00
SH-7	3.31	99.20	5.10	4.02

C.5.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	17.93	99.13	15.57	16.67
UC-2	4.65	99.22	5.13	4.88
UC-3	17.87	98.42	10.95	13.58
UC-4	20.43	99.15	26.39	23.03
UC-5	9.38	98.62	3.88	5.49
UC-6	47.37	99.66	45.92	46.63
UC-7	15.07	99.53	12.36	13.58
UC-8	37.08	99.47	24.63	29.60

C.6 E5 Base (bi-encoding)

C.6.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	24.92	98.40	8.65	11.74
Notice (N)	19.61	97.92	12.35	10.41
Sharing (SH)	16.11	98.39	5.13	6.87
User Control (UC)	21.31	98.92	16.66	17.42
Security (SE)	12.27	98.97	8.24	7.53
Data Practices (DP)	7.41	99.77	4.35	5.48
Enforcement (E)	29.77	96.19	14.12	13.30
Privacy By Design (PBD)	0.00	99.62	0.00	0.00
Contract (K)	32.58	94.96	2.95	5.40
GDPR (GDPR)	19.92	99.15	17.16	15.07
COVID (COVID)	16.67	98.80	0.28	0.56

C.6.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	2.67	99.52	2.74	2.70
CCPA-2	51.35	99.38	20.43	29.23
CCPA-3	27.72	98.26	20.10	23.30
CCPA-4	41.07	98.34	4.74	8.50
CCPA-5	43.53	97.86	5.87	10.35
CCPA-6	16.38	97.84	3.35	5.56
CCPA-7	26.67	99.57	6.90	10.96
CCPA-8	10.62	98.90	5.00	6.80
CCPA-9	8.54	94.68	0.96	1.73
CCPA-10	20.63	99.61	16.46	18.31

C.6.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	16.67	98.80	0.28	0.56

C.6.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	7.41	99.77	4.35	5.48

C.6.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	17.98	92.98	5.13	7.98
E-2	24.78	88.43	0.82	1.59
E-3	13.39	98.01	2.92	4.80
E-4	47.92	99.25	31.65	38.12
E-5	36.07	99.77	41.51	38.60
E-6	45.45	94.11	5.45	9.74
E-7	50.00	97.12	0.47	0.92
E-8	2.56	99.86	25.00	4.65

C.6.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	15.52	97.11	2.30	4.00
GDPR-2	25.00	99.87	40.91	31.03
GDPR-3	12.24	99.73	13.33	12.77
GDPR-4	19.57	99.79	26.47	22.50
GDPR-5	27.27	99.25	2.79	5.06

C.6.7 Contract (K)

Question ID	r	a	p	f_1
K-1	32.58	94.96	2.95	5.40

C.6.8 Notice (N)

Question ID	r	a	p	f_1
N-1	49.62	96.23	11.47	18.63
N-2	16.03	99.25	15.44	15.73
N-3	3.08	99.79	66.67	5.88
N-4	13.87	96.38	7.27	9.54
N-5	16.16	93.61	2.11	3.73
N-6	32.02	97.11	7.12	11.64
N-7	16.30	95.97	7.10	9.89
N-8	2.11	99.44	2.67	2.35
N-9	35.19	97.83	9.50	14.96
N-10	14.74	98.91	5.43	7.93
N-11	26.23	99.36	9.88	14.35
N-12	0.00	99.61	0.00	0.00
N-13	29.58	99.46	15.91	20.69

C.6.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	0.00	99.92	0.00	0.00
PBD-2	0.00	99.31	0.00	0.00

C.6.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	8.06	99.21	2.73	4.08
SE-2	12.50	98.91	4.25	6.34
SE-3	8.38	99.26	25.40	12.60
SE-4	25.87	97.52	5.51	9.08
SE-5	0.00	99.82	0.00	0.00
SE-6	16.67	99.78	14.29	15.38
SE-7	24.49	97.64	5.72	9.28
SE-8	2.22	99.63	8.00	3.48

C.6.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	9.24	98.31	2.69	4.17
SH-2	16.10	99.12	10.38	12.62
SH-3	35.77	97.01	5.73	9.88
SH-4	7.41	99.56	1.83	2.94
SH-5	8.24	99.46	7.78	8.00
SH-6	26.76	96.50	1.88	3.51
SH-7	9.27	98.75	5.60	6.98

C.6.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	19.31	99.11	15.73	17.34
UC-2	23.26	97.24	3.97	6.79
UC-3	12.08	99.09	21.93	15.58
UC-4	28.49	98.89	20.95	24.15
UC-5	13.28	98.49	4.76	7.01
UC-6	32.63	99.60	36.05	34.25
UC-7	17.81	99.58	16.46	17.11
UC-8	23.60	99.32	13.46	17.14

C.7 GTE Base (bi-encoding)

C.7.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	18.35	99.01	8.59	11.33
Notice (N)	17.60	96.43	4.18	5.94
Sharing (SH)	16.54	95.29	2.91	4.14
User Control (UC)	22.66	97.70	7.47	9.26
Security (SE)	11.34	98.31	3.59	4.26
Data Practices (DP)	18.52	97.88	0.81	1.55
Enforcement (E)	26.37	96.17	9.39	10.83
Privacy By Design (PBD)	0.00	99.32	0.00	0.00
Contract (K)	3.79	98.27	1.27	1.90
GDPR (GDPR)	15.50	99.65	20.56	12.98
COVID (COVID)	0.00	99.35	0.00	0.00

C.7.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	10.67	98.58	2.20	3.64
CCPA-2	36.49	99.48	20.00	25.84
CCPA-3	38.95	97.88	19.44	25.93
CCPA-4	25.00	99.14	6.09	9.79
CCPA-5	21.18	99.45	15.52	17.91
CCPA-6	8.62	98.62	3.15	4.62
CCPA-7	13.33	99.39	2.50	4.21
CCPA-8	15.93	99.06	8.78	11.32
CCPA-9	0.61	99.07	0.85	0.71
CCPA-10	12.70	99.48	7.34	9.30

C.7.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.00	99.35	0.00	0.00

C.7.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	18.52	97.88	0.81	1.55

C.7.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	19.17	84.52	2.25	4.02
E-2	17.70	93.18	1.02	1.92
E-3	7.14	98.85	3.21	4.43
E-4	51.39	98.67	18.41	27.11
E-5	21.31	99.74	30.95	25.24
E-6	21.53	98.02	9.49	13.18
E-7	62.50	96.62	0.49	0.98
E-8	10.26	99.75	9.30	9.76

C.7.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	3.45	99.42	6.15	4.42
GDPR-2	25.00	99.87	42.86	31.58
GDPR-3	26.53	99.42	8.72	13.13
GDPR-4	4.35	99.84	40.00	7.84
GDPR-5	18.18	99.69	5.06	7.92

C.7.7 Contract (K)

Question ID	r	a	p	f_1
K-1	3.79	98.27	1.27	1.90

C.7.8 Notice (N)

Question ID	r	a	p	f_1
N-1	33.85	97.78	15.20	20.98
N-2	35.11	97.17	5.69	9.80
N-3	7.69	99.61	8.06	7.87
N-4	15.09	95.20	5.41	7.96
N-5	10.48	96.13	2.46	3.98
N-6	6.18	98.80	5.39	5.76
N-7	25.19	92.40	4.93	8.25
N-8	2.11	99.02	0.99	1.34
N-9	56.17	84.40	1.94	3.76
N-10	7.37	96.14	0.65	1.20
N-11	0.00	99.48	0.00	0.00
N-12	0.00	99.51	0.00	0.00
N-13	29.58	97.94	3.57	6.37

C.7.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	0.00	99.73	0.00	0.00
PBD-2	0.00	98.91	0.00	0.00

C.7.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	22.58	94.85	0.93	1.79
SE-2	4.55	98.54	1.13	1.81
SE-3	3.14	99.18	9.09	4.67
SE-4	20.28	98.08	5.94	9.19
SE-5	0.00	99.79	0.00	0.00
SE-6	5.56	99.53	1.85	2.78
SE-7	10.20	98.82	6.36	7.83
SE-8	24.44	97.69	3.42	5.99

C.7.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	15.97	97.94	3.54	5.80
SH-2	12.71	98.98	6.94	8.98
SH-3	5.84	98.69	2.96	3.93
SH-4	51.85	77.51	0.21	0.41
SH-5	16.47	96.57	1.45	2.66
SH-6	5.63	98.64	1.17	1.93
SH-7	7.28	98.68	4.12	5.26

C.7.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	10.34	98.74	5.75	7.39
UC-2	22.48	95.46	2.25	4.10
UC-3	14.49	98.15	7.41	9.80
UC-4	8.06	99.14	15.00	10.49
UC-5	44.53	94.42	3.44	6.39
UC-6	24.21	99.32	14.84	18.40
UC-7	17.81	99.21	6.84	9.89
UC-8	39.33	97.17	4.23	7.63

C.8 BGE 1.5 Base (bi-encoding)

C.8.1 Performance by Category

Category	r	a	p	f_1
CCPA (CCPA)	18.92	99.36	17.32	16.31
Notice (N)	22.05	98.15	10.19	12.39
Sharing (SH)	17.26	98.05	8.86	9.61
User Control (UC)	21.93	99.23	22.66	19.91
Security (SE)	11.07	99.26	8.90	8.77
Data Practices (DP)	7.41	99.84	7.69	7.55
Enforcement (E)	38.20	96.10	11.57	15.53
Privacy By Design (PBD)	15.91	99.60	5.47	8.14
Contract (K)	5.30	99.11	4.70	4.98
GDPR (GDPR)	24.38	99.70	37.72	25.37
COVID (COVID)	0.00	99.93	0.00	0.00

C.8.2 CCPA (CCPA)

Question ID	r	a	p	f_1
CCPA-1	1.33	99.59	2.00	1.60
CCPA-2	33.78	99.63	28.41	30.86
CCPA-3	34.74	98.17	21.52	26.58
CCPA-4	28.57	99.49	12.60	17.49
CCPA-5	23.53	99.47	17.70	20.20
CCPA-6	11.21	99.22	9.15	10.08
CCPA-7	3.33	99.90	33.33	6.06
CCPA-8	23.01	99.34	18.98	20.80
CCPA-9	4.27	99.11	5.98	4.98
CCPA-10	25.40	99.67	23.53	24.43

C.8.3 COVID (COVID)

Question ID	r	a	p	f_1
COVID-1	0.00	99.93	0.00	0.00

C.8.4 Data Practices (DP)

Question ID	r	a	p	f_1
DP-1	7.41	99.84	7.69	7.55

C.8.5 Enforcement (E)

Question ID	r	a	p	f_1
E-1	28.06	82.09	2.77	5.04
E-2	21.24	94.29	1.46	2.74
E-3	10.71	99.07	6.32	7.95
E-4	57.64	99.20	31.68	40.89
E-5	47.54	99.55	21.97	30.05
E-6	27.27	98.13	12.26	16.91
E-7	87.50	96.76	0.72	1.43
E-8	25.64	99.72	15.38	19.23

C.8.6 GDPR (GDPR)

Question ID	r	a	p	f_1
GDPR-1	10.34	99.38	12.90	11.48
GDPR-2	30.56	99.85	36.67	33.33
GDPR-3	36.73	99.49	12.86	19.05
GDPR-4	26.09	99.84	46.15	33.33
GDPR-5	18.18	99.94	80.00	29.63

C.8.7 Contract (K)

Question ID	r	a	p	f_1
K-1	5.30	99.11	4.70	4.98

C.8.8 Notice (N)

Question ID	r	a	p	f_1
N-1	28.08	98.22	17.51	21.57
N-2	19.85	99.00	11.87	14.86
N-3	4.62	99.76	21.43	7.59
N-4	19.22	94.24	5.39	8.41
N-5	7.86	98.36	6.06	6.84
N-6	30.90	98.63	16.08	21.15
N-7	24.94	96.17	10.71	14.99
N-8	4.21	98.75	1.39	2.09
N-9	48.77	95.98	6.59	11.62
N-10	14.74	98.50	3.66	5.87
N-11	42.62	99.45	16.67	23.96
N-12	0.00	99.53	0.00	0.00
N-13	40.85	99.31	15.10	22.05

C.8.9 Privacy By Design (PBD)

Question ID	r	a	p	f_1
PBD-1	31.82	99.76	10.94	16.28
PBD-2	0.00	99.44	0.00	0.00

C.8.10 Security (SE)

Question ID	r	a	p	f_1
SE-1	9.68	99.36	4.23	5.88
SE-2	3.41	99.29	2.31	2.75
SE-3	14.66	99.18	25.45	18.60
SE-4	15.38	99.15	14.19	14.77
SE-5	2.08	99.80	7.14	3.23
SE-6	25.00	99.70	12.68	16.82
SE-7	18.37	97.96	5.21	8.12
SE-8	0.00	99.62	0.00	0.00

C.8.11 Sharing (SH)

Question ID	r	a	p	f_1
SH-1	25.21	99.08	13.95	17.96
SH-2	21.19	99.09	12.32	15.58
SH-3	8.03	99.35	13.75	10.14
SH-4	40.74	90.84	0.40	0.80
SH-5	14.12	99.32	8.39	10.53
SH-6	4.23	99.56	4.48	4.35
SH-7	7.28	99.15	8.73	7.94

C.8.12 User Control (UC)

Question ID	r	a	p	f_1
UC-1	23.45	99.06	16.59	19.43
UC-2	5.43	99.29	7.29	6.22
UC-3	27.54	98.43	15.16	19.55
UC-4	32.80	98.82	21.03	25.63
UC-5	9.38	99.30	11.54	10.34
UC-6	52.63	99.57	37.04	43.48
UC-7	9.59	99.75	46.67	15.91
UC-8	14.61	99.62	26.00	18.71

D Websites in Dataset

Tranco Range	Tranco Rank	URL	IAB 2.0 Top-Level Categories	Document Count	Word Count
1-10	1	google.com	IAB19 (Technology & Computing)	2	12379
1-10	2	facebook.com	IAB14 (Society)	18	39915
1-10	3	youtube.com	IAB25 (Non-Standard Content) IAB1 (Arts & Entertainment)	2	12390
1-10	4	microsoft.com	IAB9 (Hobbies & Interests) IAB19 (Technology & Computing)	4	39988
1-10	6	twitter.com	IAB14 (Society)	8	10495
1-10	7	instagram.com	IAB19 (Technology & Computing) IAB14 (Society)	4	14897
1-10	10	netflix.com	IAB25 (Non-Standard Content) IAB1 (Arts & Entertainment)	3	10139
1-1000	11	linkedin.com	IAB4 (Careers) IAB14 (Society)	5	14495
1-1000	12	qq.com	IAB1 (Arts & Entertainment) IAB12 (News)	4	13692
1-1000	13	apple.com	IAB19 (Technology & Computing) IAB22 (Shopping)	4	9236
1-1000	15	wikipedia.org	IAB5 (Education)	2	14536
1-1000	24	amazon.com	IAB22 (Shopping)	6	9034
1-1000	31	pinterest.com	IAB18 (Style & Fashion) IAB14 (Society)	9	16726
1-1000	33	adobe.com	IAB19 (Technology & Computing) IAB3 (Business)	4	14684
1-1000	44	reddit.com		2	7229
1-1000	71	tumblr.com	IAB25 (Non-Standard Content)	3	17869
1-1000	80	msn.com	IAB25 (Non-Standard Content) IAB12 (News)	2	8635
1-1000	109	myshopify.com	IAB3 (Business) IAB22 (Shopping)	2	10987
1-1000	114	cnn.com	IAB12 (News)	2	13017
1-1000	119	twitch.tv	IAB25 (Non-Standard Content) IAB9 (Hobbies & Interests)	3	12413
1-1000	136	imdb.com	IAB25 (Non-Standard Content) IAB1 (Arts & Entertainment) IAB9 (Hobbies & Interests)	2	5318
1-1000	141	stackoverflow.com	IAB19 (Technology & Computing)	3	10484
1-1000	163	aliexpress.com	IAB22 (Shopping)	2	7345
1-1000	180	washingtonpost.com	IAB11 (Law, Gov't & Politics) IAB12 (News)	2	7565
1-1000	216	chaturbate.com	IAB25 (Non-Standard Content)	3	10693
1-1000	229	amazon.co.uk	IAB22 (Shopping)	3	12583
1-1000	245	researchgate.net	IAB19 (Technology & Computing)	3	11118
1-1000	295	walmart.com	IAB22 (Shopping)	2	7210
1-1000	311	pornhub.com	IAB25 (Non-Standard Content)	7	19749
1-1000	340	ted.com	IAB25 (Non-Standard Content)	2	7215
1-1000	347	livejasmin.com	IAB25 (Non-Standard Content)	3	10441
1-1000	365	okta.com	IAB3 (Business) IAB19 (Technology & Computing)	2	11675
1-1000	369	xvideos.com	IAB25 (Non-Standard Content)	8	15822
1-1000	397	instructure.com	IAB5 (Education)	4	8597
1-1000	464	theverge.com	IAB19 (Technology & Computing)	3	12697
1-1000	469	loc.gov	IAB11 (Law, Gov't & Politics)	2	2898
1-1000	524	craigslist.org	IAB25 (Non-Standard Content) IAB21 (Real Estate) IAB22 (Shopping)	2	2063
1-1000	598	homedepot.com	IAB10 (Home & Garden)	3	12885
1-1000	609	stumbleupon.com	IAB19 (Technology & Computing)	2	11417
1-1000	641	pbs.org	IAB25 (Non-Standard Content)	2	4196
1-1000	666	hulu.com	IAB25 (Non-Standard Content) IAB1 (Arts & Entertainment)	3	14169

1-1000	683	steampowered.com	IAB9 (Hobbies & Interests)	2	4574
1-1000	843	fortune.com	IAB1 (Arts & Entertainment) IAB13 (Personal Finance)	2	10844
1-1000	993	arstechnica.com	IAB1 (Arts & Entertainment)	2	6233
1000-10000	1090	fool.com	IAB13 (Personal Finance)	2	9979
1000-10000	1104	barnesandnoble.com	IAB1 (Arts & Entertainment)	4	13471
1000-10000	1128	corp.ign.com	IAB9 (Hobbies & Interests)	3	21645
1000-10000	1254	thehill.com	IAB12 (News)	2	5021
1000-10000	1337	dictionary.com	IAB9 (Hobbies & Interests)	1	7335
1000-10000	1396	usa.healthcare.siemens.com	IAB15 (Science) IAB7 (Health & Fitness)	3	5799
1000-10000	1621	usa.gov	IAB11 (Law, Gov't & Politics)	1	778
1000-10000	1628	archives.gov	IAB11 (Law, Gov't & Politics)	1	3153
1000-10000	1671	adweek.com	IAB3 (Business)	2	13495
1000-10000	1678	lonelyplanet.com	IAB20 (Travel)	3	5271
1000-10000	1831	wordreference.com	IAB19 (Technology & Computing) IAB5 (Education)	1	572
1000-10000	1889	amd.com	IAB19 (Technology & Computing)	3	8043
1000-10000	1999	allrecipes.com	IAB8 (Food & Drink)	3	16320
1000-10000	2146	slickdeals.net	IAB22 (Shopping)	2	6797
1000-10000	2363	reverbnation.com	IAB1 (Arts & Entertainment)	2	22859
1000-10000	2564	dpreview.com	IAB9 (Hobbies & Interests) IAB19 (Technology & Computing)	6	6980
1000-10000	2672	macrumors.com	IAB19 (Technology & Computing)	1	1272
1000-10000	2976	freep.com	IAB12 (News)	3	7884
1000-10000	3343	everydayhealth.com	IAB15 (Science) IAB7 (Health & Fitness)	3	34103
1000-10000	3474	uh.edu	IAB5 (Education)	1	4190
1000-10000	3510	edmunds.com	IAB2 (Automotive)	2	12616
1000-10000	3962	babycenter.com	IAB15 (Science) IAB6 (Family & Parenting) IAB7 (Health & Fitness)	2	16777
1000-10000	4541	match.com	IAB14 (Society)	4	15183
1000-10000	4683	ubi.com	IAB9 (Hobbies & Interests)	3	12588
1000-10000	4937	geocaching.com	IAB19 (Technology & Computing)	2	12973
1000-10000	5283	sltrib.com	IAB12 (News)	2	9042
1000-10000	5615	wizards.com	IAB9 (Hobbies & Interests)	3	28462
1000-10000	6002	thermofisher.com	IAB15 (Science)	3	9510
1000-10000	6145	newgrounds.com	IAB9 (Hobbies & Interests)	2	5409
1000-10000	6658	allstate.com	IAB13 (Personal Finance)	3	11136
1000-10000	7223	purevolume.com	IAB1 (Arts & Entertainment)	2	4547
1000-10000	7868	simplemachines.org	IAB19 (Technology & Computing)	2	2218
1000-10000	7951	basketball-reference.com	IAB17 (Sports)	2	3787
1000-10000	7973	gamepedia.com	IAB9 (Hobbies & Interests)	2	9944
1000-10000	8064	dailynews.com	IAB12 (News)	3	9532
1000-10000	8146	ebaumsworld.com	IAB1 (Arts & Entertainment)	2	8453
1000-10000	8851	moneysavingexpert.com	IAB13 (Personal Finance)	3	15539
1000-10000	9067	gaiaonline.com	IAB9 (Hobbies & Interests)	2	10800
10000-100000	10053	23andme.com	IAB7 (Health & Fitness) IAB15 (Science)	5	39584
10000-100000	11173	zacks.com	IAB13 (Personal Finance)	2	5880
10000-100000	13372	dailystrength.org	IAB7 (Health & Fitness)	2	6009
10000-100000	14512	valvesoftware.com	IAB10 (Home & Garden)	2	5320
10000-100000	14565	signonsandiego.com	IAB12 (News)	2	14777
10000-100000	15650	soundclick.com	IAB1 (Arts & Entertainment)	2	3699
10000-100000	17000	airliners.net	IAB9 (Hobbies & Interests) IAB19 (Technology & Computing) IAB20 (Travel) IAB25 (Non-Standard Content)	2	6880
10000-100000	18163	videohelp.com	IAB19 (Technology & Computing)	1	1963
10000-100000	20256	filefront.com	IAB9 (Hobbies & Interests)	2	6510
10000-100000	20440	somethingawful.com	IAB1 (Arts & Entertainment)	1	1612

10000-100000	20895	yardbarker.com	IAB17 (Sports)	2	5239
10000-100000	26711	eharmony.com	IAB14 (Society)	2	16146
10000-100000	28046	afterdawn.com	IAB12 (News)	1	847
10000-100000	28890	sci-news.com	IAB15 (Science)	1	569
10000-100000	30950	namepros.com	IAB19 (Technology & Computing)	2	8491
10000-100000	32720	us.mouthshut.com	IAB24 (Uncategorized)	2	7554
10000-100000	34840	drinksmixer.com	IAB8 (Food & Drink)	2	14662
10000-100000	36320	ashleymadison.com	IAB25 (Non-Standard Content) IAB14 (Society)	2	17552
10000-100000	37097	taylorswift.com	IAB19 (Technology & Computing)	2	12151
10000-100000	38498	kraftrecipes.com	IAB8 (Food & Drink)	2	7630
10000-100000	38793	cbsinteractive.com	IAB12 (News)	3	8446
10000-100000	44194	bolt.com	IAB3 (Business)	2	18194
10000-100000	47174	bio-rad.com	IAB3 (Business)	2	10793
10000-100000	64822	tgifridays.com	IAB8 (Food & Drink)	2	9020
10000-100000	74083	twoplustwo.com	IAB9 (Hobbies & Interests)	2	6218
10000-100000	75247	christianmingle.com	IAB14 (Society)	3	36186
10000-100000	76486	primagames.com	IAB9 (Hobbies & Interests)	2	11455
10000-100000	99193	dailyillini.com	IAB12 (News)	2	2275
100000-1000000	100437	listography.com	IAB25 (Non-Standard Content)	2	2925
100000-1000000	110999	friendfinder.com	IAB14 (Society)	2	12337
100000-1000000	116128	chicagomarathon.com	IAB7 (Health & Fitness)	1	3317
100000-1000000	118881	hardwarezone.com	IAB19 (Technology & Computing)	2	7314
100000-1000000	122817	cariboucoffee.com	IAB8 (Food & Drink)	3	10040
100000-1000000	132336	restaurantnews.com	IAB12 (News)	1	836
100000-1000000	175175	cincymuseum.org	IAB1 (Arts & Entertainment)	1	2323
100000-1000000	176865	myriad.com	IAB15 (Science) IAB7 (Health & Fitness)	3	6778
100000-1000000	178874	opendiary.com	IAB14 (Society)	3	6293
100000-1000000	187525	fightingillini.com	IAB12 (News)	2	5768
100000-1000000	201854	aq.com	IAB9 (Hobbies & Interests)	3	10208
100000-1000000	214643	mate1.com	IAB14 (Society)	2	15435
100000-1000000	230902	india-forums.com	IAB14 (Society)	2	8802
100000-1000000	237561	helix.com	IAB15 (Science) IAB7 (Health & Fitness)	2	14241
100000-1000000	239486	abita.com	IAB8 (Food & Drink)	2	2467
100000-1000000	277961	dnacenter.com	IAB15 (Science) IAB6 (Family & Parenting) IAB7 (Health & Fitness)	2	8662
100000-1000000	316184	coffeereview.com	IAB8 (Food & Drink)	2	2137
100000-1000000	344874	communitycoffee.com	IAB8 (Food & Drink)	4	8127
100000-1000000	459935	bgi.com	IAB7 (Health & Fitness)	1	4800
100000-1000000	562225	ambrygen.com	IAB7 (Health & Fitness)	3	6169
100000-1000000	585888	bigtent.com	IAB11 (Law, Gov't & Politics) IAB6 (Family & Parenting)	2	20840
100000-1000000	705374	orig3n.com	IAB15 (Science) IAB7 (Health & Fitness)	2	7119
100000-1000000	707387	enpnetwork.com	IAB4 (Careers)	2	13180
100000-1000000	722795	aboutus.disaboom.com	IAB24 (Uncategorized)	1	458
100000-1000000	814897	nygenome.org	IAB15 (Science) IAB7 (Health & Fitness)	2	1666
100000-1000000	879524	true.com	IAB13 (Personal Finance)	2	8363
100000-1000000	982740	sequencing.com	IAB15 (Science)	3	13097
1000000+	1015718	gays.com	IAB25 (Non-Standard Content)	2	11301
1000000+	1071352	spark.com	IAB14 (Society)	2	27220
1000000+	1201138	greensingles.com	IAB14 (Society)	3	6187
1000000+	1248563	completegenomics.com	IAB5 (Education)	3	6525
1000000+	1363985	veggiedate.org	IAB14 (Society)	2	3974
1000000+	1557614	metrodate.com	IAB14 (Society)	2	2316
1000000+	2364598	my.opera.com	IAB19 (Technology & Computing)	3	5731
1000000+	4599194	wealthymen.com	IAB14 (Society)	2	11487
1000000+	5301232	heremedia.com	IAB14 (Society)	3	12608

1000000+	6127774	epernicus.com	IAB3 (Business)	2	4917
1000000+	10000000	sediabio.com	IAB7 (Health & Fitness)	2	3323
1000000+	10000000	webmediabrands.com	IAB9 (Hobbies & Interests) IAB4 (Careers)	3	15587