Agree First, Read Never?: Ontology Development for Policy Document Analysis

Anonymous ACL submission

Abstract

Privacy policies and terms of service shape the collection, sharing, and protection of personal data—yet they remain some of the most opaque, lengthy, and complicated documents users encounter. Despite this urgent societal need, we currently lack a structured, formal, and scalable ontological framework to effectively analyze and interpret these texts across diverse platforms. In this work, we address this gap by introducing a four-layer hierarchical ontology for fine-grained policy document analysis. To construct this ontology, we start from 245 established cases from the Terms of Service; Didn't Read (ToS;DR) project and OPP-115 taxonomy. Using GPT-4 Turbo and a structured prompting pipeline, we map each policy case to a semantic hierarchy—spanning broad categories, sub-categories, fine-grained distinctions, and policy actions—and formalize the structure in Web Ontology Language (OWL). This novel ontology powers a staged, multi-tiered classification model that exploits semantic coherence at each level. Our empirical evaluation demonstrates that this approach massively outperforms a flat PrivBERT baseline, which struggles with the fine-grained label space (94.3% vs 73.4% F1). This superior performance holds consistently at finer levels, alongside an analysis that confirms practical inference efficiency for real-world deployment. Beyond raw metrics, our framework promotes interpretability, supports cross-document alignment via both hierarchical and non-hierarchical links, and ultimately helps shape more transparent, user-aligned privacy standards.

1 Introduction

Users typically agree to lengthy and complicated contracts without fully understanding them (Obar and Oeldorf-Hirsch, 2020). Further, due to the intricate structure of Terms of Service (ToS), users often struggle to understand and track changes when service providers revise policies (Adhikari, 2020). These documents often contain crucial information about data usage, user rights, and legal responsibilities. But, consumers often sign without fully understanding the terms and conditions, which can lead to unintentional breaches of personal privacy and data control. This is a major issue in society, and translating legalese into simple language should help (Bakos et al., 2014).

Pinterest's "Simply Put" feature (Pinterest), 500px's "Basically" summaries (500px), and Prolific's different access options (Prolific) for their Researcher Terms aim to simplify intricate legal terminology. These summaries are more approachable and comprehensible to users which enhances user empowerment and data literacy, promoting a more ethically responsible digital environment. Comprehending such complex privacy-related documents is crucial for users to make well-informed decisions regarding their data. The absence of a well-defined framework for examining these documents adds complexity to privacy assessments and adherence to regulations.

The challenges caused by lack of framework underscore the need for structured methodologies to analyze and interpret policy documents effectively, which is why we turn to ontological frameworks. Ontologies, due to their formal structure, may effectively handle the intricacies of privacy policy analysis by facilitating accurate definitions, affiliations, and hierarchies among concepts within a certain domain. This enables precise definitions and connections among the concepts (Gruber, 1993; Noy et al., 2001). While frameworks like PrivOnto (Oltramari et al., 2018) have made important strides in representing and analyzing privacy policies using ontologies, there is an ongoing opportunity to expand their capabilities to cover the diverse and evolving nature of privacy-related documents. This research aims to address the lack of an organized method to classify and examine privacy policy annotation by constructing an ontological framework from 245

concepts (cases) from Terms of Service; Didn't Read (2012) (ToS;DR).

To address this gap, our research constructs a hierarchical ontology from 245 canonical policy cases curated by ToS;DR (Terms of Service; Didn't Read, 2012). This ontology integrates multiple layers of semantic abstraction—including categories, sub-categories, fine-grained labels, and policy actions—to support both automated classification and human-centric interpretation. Our framework is designed not only to standardize and scale policy interpretation, but also to enhance explainability and cross-platform consistency in privacy assessments.

Guided by this motivation, our work is driven by the following research questions:

RQ1 How can we build a hierarchical ontology that captures both the meaning and structure of privacy policy documents across different services?

RQ2 To what extent does ontology-guided classification improve the accuracy, interpretability, and generalizability of automated privacy policy labeling compared to flat classification models?

RQ3 What types of relationships hierarchical and non-hierarchical are necessary to support fine-grained reasoning and clause-level alignment in privacy documents?

Contributions The main contributions of this work are threefold:

1. An **automated, repeatable pipeline** for ontology generation and evaluation, combining GPT-4–based structured prompting with expert validation to map raw policy clauses into a multi-level semantic hierarchy.
2. A **four-layer hierarchical ontology** of privacy policy concepts—anchored in the OPP-115 taxonomy, enriched with ToS;DR metadata, and formalized in OWL with both hierarchical and cross-cutting semantic relations.
3. A **comprehensive empirical evaluation** showing that our ontology-guided hierarchical classifier substantially outperforms a flat PrivBERT baseline across all tiers (Broad, Sub, Fine, Action), with accuracy gains of 8–17%, improved interpretability, and modest inference cost.

2 Background

Calpric (Qiu et al., 2023) is a tool that improves the process of labeling privacy policies by using automatic text selection and segmentation, which allows crowdsourced annotators to produce high-quality labels for different segments of privacy-related text more efficiently. As a result, the Calpric Privacy Policy Corpus contains 16,856 tagged text segments. This corpus supports the development of more precise models for analyzing privacy policies. The study Wilson et al. (2016a) conducted is similar in that it analyzes 115 website privacy policies and categorizes them into 23,194 data practices using skilled annotators and a customized annotation tool to simplify privacy policy analysis using natural language processing and machine learning.

Ontologies capture the relationships present in policies, hence improving the accuracy and comprehensiveness of automated privacy policy analysis. Placing terms inside individual policies and broader privacy regulations within local and global ontologies achieves this (Cui et al., 2023). The Privacy Policy Permission Model (PPPM) (Majedi and Barker, 2024) is a graphical representation of privacy regulations that is concise, consistent, and easy to understand. It helps to both understand how organizations use data and identify any flaws in privacy protections. This study also highlights the advantages of ontologies in policy abstraction, organization, and dissemination. It proposes more investigation into multiple inheritance, dispute resolution, and policy validation using POLICYTAB, a Protégé plug-in in open distributed settings such as the Web (Nejdl et al., 2005). Andow et al. (Andow et al., 2019) propose PolicyLint, a tool that identifies internal contradictions within individual privacy policies by analyzing negation and semantic hierarchies, highlighting misleading or conflicting statements often overlooked by existing summarization tools.

Esteves and Rodríguez-Doncel (2024) evaluated 22 privacy policy languages, vocabularies, and ontologies for GDPR compliance and develop a GDPR information flows ontology. In their 2018 publication, (Palmirani et al., 2018) presented PrOnto, an ontology designed to facilitate legal reasoning in the context of privacy. PrOnto achieves this by incorporating legal norms and privacy standards, thereby bridging the divide between legal texts and technical implementations. Pandit et al. (2019) created a uniform set of terms for data pri-

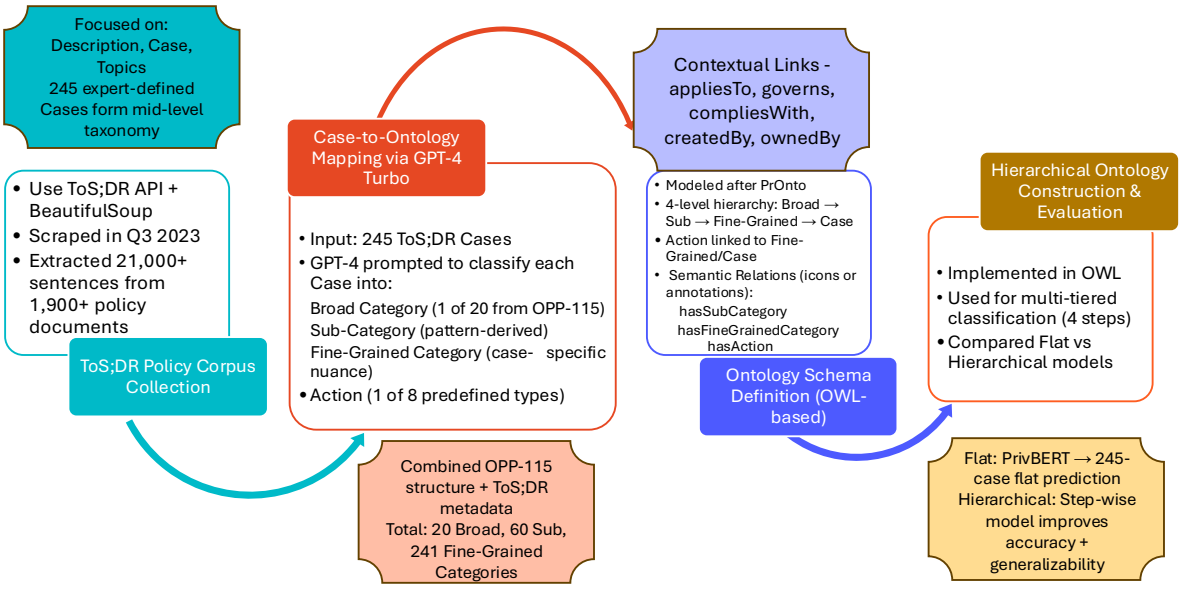

Figure 1: Pipeline architecture for ontology-guided hierarchical classification, highlighting the integration of community-curated ToS;DR cases, automated GPT-4 mapping, formal OWL schema modeling, and structured evaluation under a unified framework.

vacy in order to improve compatibility and comprehension among different systems. PrivOnto, a semantic framework developed by Oltramari et al. (2018), utilizes ontology-based methodologies to create a model for privacy rules. This framework enables automated compliance checks and enhances transparency in policies.

3 Ontology Construction Pipeline

Here we discuss the methodology in detail as shown in Figure 1:

3.1 Data Collection and Explanation

To initiate our ontology design, we collected annotated policy data from the Terms of Service; Didn't Read (2012) (ToS;DR) project, a long-running community-driven initiative that curates and categorizes user-facing policy documents. We used the ToS;DR streaming API along with BeautifulSoup (Richardson, 2023) to scrape the dataset in the third quarter of 2023.

The ToS;DR corpus comprises over 21,000 annotated sentence-level descriptions drawn from 1,900+ policy documents spanning major platforms and services. Each entry includes several metadata fields, including **Description** (a natural language snippet from a policy), **Case** (a normalized policy label), Title, Status, Comments, and Author ID.

Our work focuses on the **Case** attribute, which functions as a canonical label assigned by a team of approximately 12 legal and technical experts. These 245 manually-defined cases form a robust mid-level taxonomy of privacy-related practices across companies and services. Cases include both affirmative and negative versions of similar concepts (e.g., *"Data is sold"* vs. *"Data is not sold"*), offering comprehensive coverage of common policy patterns. This consistency across policies makes the set ideal for ontology construction, as the same conceptual unit (i.e., case) may occur across many services with comparable semantics.

In addition, ToS;DR annotators often associated cases with broader *topics* (e.g., `Personal data`, `Third parties`, `Tracking`), which we used as a weak signal for aligning cases with higher-level categories during the classification phase (see Section 3.2). This alignment, in turn, enabled us to map diverse textual policies to a unified, interpretable hierarchical structure.

3.2 Data Classification

The data classification stage instantiates a four-level taxonomy over the 245 ToS;DR cases and automates the assignment of each case into this hierarchy. We implemented our classification pipeline using publicly available libraries, including `NLTK`

Input: Cases from ToSDR

Description: Given the following sentence {case},
How will you map the case to a category from the list of 10 categories {Categories} and 8 actions {actions}?
How will you generate the new sub-category and fine-grained category for each case?

Context:
- Case: A one-line pre-defined sentence from a privacy policy or terms of service document.
- Categories: Type of Data Practice associated with the case.
- Sub-Categories: Types of sub-labels given to the categories for more granularity.
- Fine-Grained Categories: More specific divisions of sub-categories that provide detailed classifications.
- Actions: The specific action or outcome associated with the case.

Instruction:
1. You should first review all the cases and then match each case with a category and action.
2. Each case should be mapped to strictly one of the listed categories and actions from the provided lists and you should generate new sub-category and fine-category relevant to the category of its own.
3. Be consistent with your responses every time you run the prompt.
4. If the case doesn't fit into any category or action, you may leave it blank.

Few shot example:
{"Case": "Do Not Track (DNT) headers are ignored and you are tracked anyway even if you set this header.", "Category": "Do Not Track", "Sub-Category": "Violation", "Fine-Grained Category": "Tracking Ignored", "Action": "Assert" }

Figure 2: Structured prompt design for GPT-4 Turbo, guiding each ToS;DR case into a four-level ontology with predefined categories and actions, and generated sub- and fine-grained labels in a strict JSON format.

for text preprocessing, `spaCy` for linguistic parsing, the `transformers` library by Hugging Face for model integration, and `BERT` embeddings for sentence encoding.

3.2.1 Taxonomy

Broad Categories: 20 high-level data-practice categories from OPP-115 (Wilson et al., 2016b). (Data Retention, Data Security, Disclaimer, Dispute Resolution, Do Not Track, First Party Collection/use, Governing Law, Indemnification, International and Specific Audiences, Limitation of Liability, Misc, Policy Change, Registration & Account Security, Service Description, Third Party Sharing/Collection, User Access, Edit and Deletion, User Access, Edit and Deletion , User Choice, User Responsibility)

Sub-Categories: Approximately 60 semantically coherent groupings induced from recurring patterns and ToS;DR topics. Few examples:

1. Data Retention - " Some personal data may be kept for business interests or legal obligations"
2. User Identity - " Pseudonyms are allowed due to the service's billing policies"

Fine-Grained Categories: 241 specific labels capturing nuanced policy distinctions. Few examples are :

1. Allow Pseudonyms - " Pseudonyms are allowed"
2. Registration Requirements - " No need to register"

Actions: Eight canonical policy actions. (Ambiguous, Assert, Purpose, Denial, Personal Information Type, User Type, Choice Scope, Choice Type, Personal Information Type/Denial, Implicit)

3.2.2 Automation via LLM

A structured prompting pipeline assigned each case to the four taxonomy levels:

- *Prompt design:* Prompts list all 20 Broad Categories and specify JSON output keys (`broad_category`, `sub_category`, `fine_grained_category`, `action`), with two few-shot examples.
- *Batch inference:* GPT-4 Turbo processes cases in batches of 25 with `temperature=0.2` to encourage consistent, nearly-deterministic outputs.
- *Output parsing:* We parse responses as JSON; any schema violations trigger automated re-prompts until achieving a well-formed response.
- *Validation and refinement:* The researchers manually audited a random 10% subsample of mappings. We iteratively refined

prompt wording until manual-check accuracy exceeded 95%.

3.2.3 Outcome

The described procedure produced fully specified four-level labels for all 245 ToS;DR cases, yielding an ontology-ready dataset for downstream hierarchical classification and semantic reasoning. We passed each case through a structured prompting pipeline (provided in the supplemental materials) and asked the model to:

1. Assign the case to exactly one of the 20 **Broad Categories**.
2. Generate an appropriate **Sub-Category** based on the semantics of the case.
3. Create a more specific **Fine-Grained Category** that reflects the nuance of the case.
4. Select an **Action** from a predefined set of 8 canonical privacy policy actions (e.g., *Allow*, *Deny*, *Assert*).

We designed the prompt to enforce strict structure and consistent outputs, for example, the following classifications arose for the following case: Case: "`Alternative accounts are not allowed.`"

- **Broad Category:** `Registration & Account Security`
- **Sub-Category:** `Account Management`
- **Fine-Grained Category:** `Alternative Accounts Restriction`
- **Action:** `Deny`

Across all 245 cases, this process resulted in 20 broad categories, 60 subcategories, 241 unique fine-grained categories, and consistent mapping to 8 standardized actions. The resulting hierarchical structure reflects both legal policy expectations (from OPP-115) and practical document-level distinctions observed across services (from ToS;DR), enabling structured semantic comparisons between privacy documents.

4 The Hierarchical Ontology

To formalize the hierarchical structure for our classification, we developed an ontology schema inspired by the legal reasoning ontology *PrOnto* (Palmirani et al., 2018). Our schema captures the semantic relationships among privacy-related concepts across four levels: **Broad Categories**, **Sub-Categories**, **Fine-Grained Categories**, and individual **Cases**. These elements are connected through a series of structured relationships that support both hierarchical inheritance and cross-cutting associations.

4.1 Hierarchical Relationships

We define three key hierarchical links to organize categories:

- `hasSubCategory`: ...connects a Broad Category to its Sub-Categories.
- `hasFineGrainedCategory`: ...links a Sub-Category to one or more Fine-Grained Categories.
- `hasAction`: ...maps a Fine-Grained Category or Case to one of the eight standard Actions (e.g., `Allow`, `Deny`, `Assert`).

These relationships create a tree-like structure that supports fine-grained classification while maintaining semantic traceability from high-level categories down to individual labeled policy clauses.

4.2 Non-Hierarchical Semantic Links

Beyond taxonomic relationships, our ontology also incorporates a set of cross-cutting relationships to support contextual reasoning and regulatory interpretation:

- `appliesTo`: ...indicates the scope or entity targeted by a rule (e.g., a user, device, or account).
- `governs` and `compliesWith`: ...describe normative or legal constraints embedded within policy clauses.
- `ownedBy` and `createdBy`: ...denote provenance and authorship of specific policy concepts.

Altogether, our ontology schema serves as a flexible foundation for organizing, annotating, and reasoning about privacy policies and terms of service documents in a consistent and extensible manner.

Beyond taxonomic classification, these cross-cutting links are essential for enabling fine-grained semantic reasoning and compliance checking (RQ3). For example, the four-layer classification may label a clause as **Fine-Grained Category**: *Legal Name Required* and **Action**: *Assert*. However, adding the non-hierarchical link `governs` to Identity Verification and `appliesTo` to User Registration provides the necessary context to infer the legal constraint. Similarly, the compliesWith link connects policy concepts directly to regulatory metadata (e.g., GDPR requirements) which is crucial for cross-document alignment and automated auditing. This layered semantic structure moves the ontology beyond a pure classifier backbone toward a functional legal reasoning framework.

4.3 Ontology Construction

This work presents the development of a hierarchical ontology utilizing the Web Ontology Language (OWL) (Antoniou and Harmelen, 2009) to address the complexities involved in classifying 245 unique privacy-related cases. Each case links to a layered semantic structure consisting of **Categories**, **Sub-Categories**, **Fine-Grained Categories**, and **Actions**. These linkages encode formally defined relationships. The resulting ontology serves not only as a semantic reference, but also as a foundational supervision layer for downstream classification tasks.

We leverage this ontological structure to implement a multi-tiered classification framework designed to improve both predictive accuracy and interpretability. The core hypothesis is that enforcing a structured decomposition of the label space—guided by human-defined and LLM-refined ontological constraints—will result in more robust clause-level classification, particularly in domains where semantic classes are densely overlapping.

Overall, this construction process resulted in a reproducible and semantically rich ontology that serves as both a structured annotation framework and a reasoning backbone for automated policy analysis.

5 Hierarchical Classification and Evaluation

The ontology generation process is iterative, combining expert-defined categories from the ToS;DR dataset with LLM–assisted semantic refinement. Domain experts manually reviewed each mapping to ensure conceptual consistency and alignment with privacy law interpretations. Finally, we validated the ontology with reasoning tools in *Protégé* with the *HermiT* reasoner to verify logical coherence, eliminate redundancy, and ensure the absence of circular dependencies.

Our dual evaluation strategy allows us to directly compare flat and hierarchical paradigms under the same model architecture and dataset. We evaluate both systems using the following metrics:

- **Quantitative Evaluation** We compare the per-layer accuracy and macro-averaged F1 in the hierarchical model, measuring the accuracy at each classification stage (Broad Category, Sub-Category, Fine-Grained Category, Action). We did this to assess step-wise consistency and trace how errors propagate through the pipeline. In our analysis, we pay particular attention to the Fine-Grained Category level, selected because it measures end-to-end performance on the most detailed label set (241 classes), aligns directly with the granularity of the flat baseline, and avoids conflating errors across hierarchy layers that would occur if metrics were aggregated at higher or mixed levels. While class imbalance can influence accuracy, considering both metrics provides an intuitive overview of each tier's performance and complements the F1 results at finer levels.
- **Qualitative error analysis** of matched cases across services to uncover semantic drift and systematic misclassification patterns that are not evident from quantitative metrics alone.
- **Inference cost comparison** to estimate computational trade-offs, quantifying average inference time per instance and total model invocations required by hierarchical versus flat classification approaches, thereby evaluating scalability and real-world deployment feasibility.

This evaluation strategy emphasizes both quantitative performance and qualitative insight, enabling us to assess whether ontological structure improves not just accuracy, but also generalizability and interpretability in real-world privacy documents.

6 Empirical Results: Performance and Analysis

6.1 Example Structure

Figure 3 illustrates a sample ontological structure under the Broad Category *Registration & Account Security*, which contains Sub-Categories such as `User Identity`, `Account Management`, and `Payment`. The structure further refines these into Fine-Grained Categories (e.g., `Legal Name Required`, `Alternative Accounts Restriction`, `Credit Card Storage Policy`) and connected to appropriate Actions and metadata relationships.

6.2 Quantitative Evaluation

The hierarchical model consistently and substantially outperformed the flat PrivBERT baseline at every classification tier. As shown in Table 1, the Broad Category accuracy jumped from 75.5% (F1 = 73.4%) under the flat approach to 92.7% (F1 = 94.3%), a +17.2 percentage point gain in accuracy. Sub-Category and Fine-Grained levels each saw

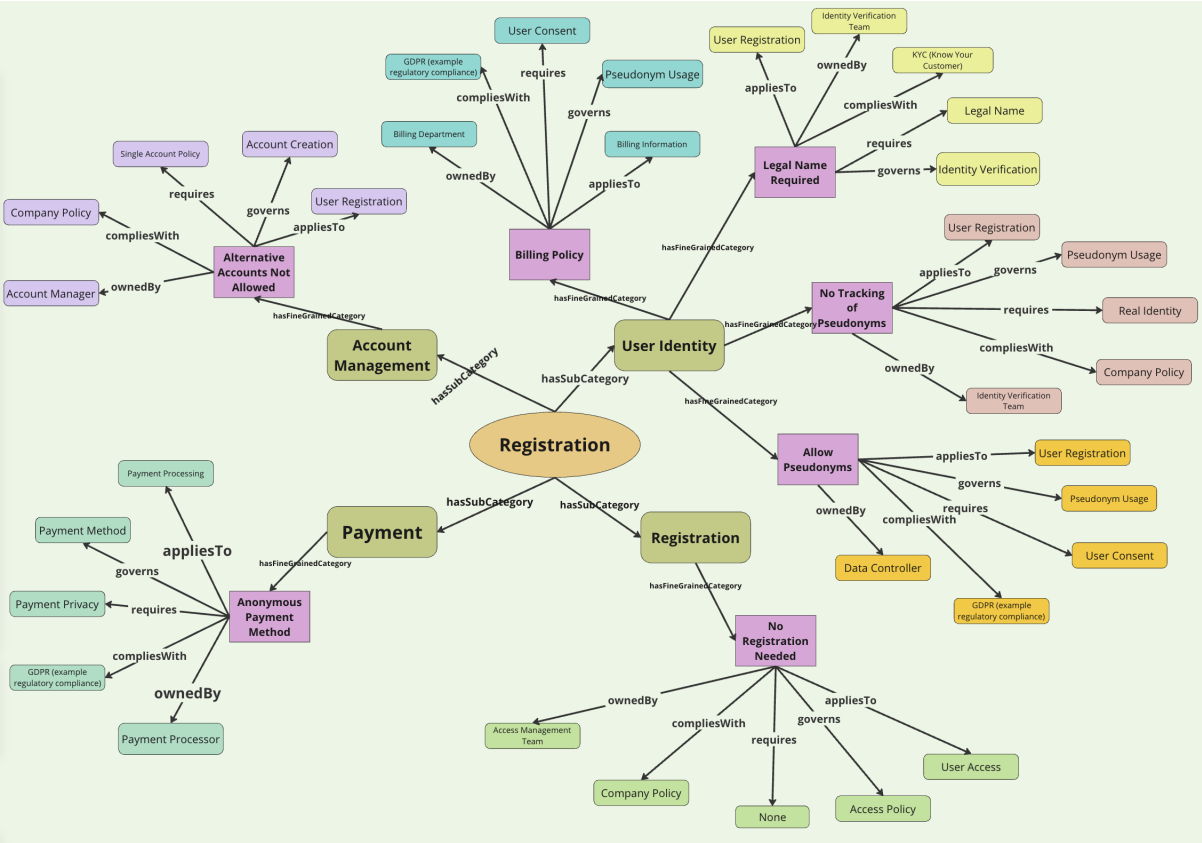

Figure 3: Hierarchical Ontological Structure for Registration: Categorization and Relationships for User Identity, Account Management, and Payment Methods

Label	Acc	F1	Prec.	Rec.
Category	0.927	0.943	0.960	0.927
Sub-Cat.	0.839	0.910	0.987	0.840
Fine-Grained	0.839	0.901	1.000	0.838
Action	0.863	0.883	1.000	0.860
Flat	0.755	0.734	0.739	0.755

Table 1: Macro Averaged Evaluation Metrics for Hierarchical and Flat Classification, both from PrivBERT

Evaluation Metric	Value
Total Unique Cases Evaluated	4
Total Instances Across Services	54
Broad Category Match Rate	94.3%
Sub-Category Match Rate	46.3%
Fine-Grained Category Match Rate	72.2%

Table 2: Summary of label agreement between ontology-based gold labels and predicted outputs across test services for four recurring cases. Note that we do not have a match rate for the action level because we did not train PrivBERT to predict actions.

accuracy rise to 83.9% (F1 $\approx$ 90.0%)—an +8.4 percentage point improvement—while Action prediction improved from 75.5% to 86.3% accuracy (F1 = 88.3%), a +10.8 percentage point increase. These results underscore that breaking the label space into semantically coherent, staged predictions dramatically reduces confusion and yields higher precision and recall across the board.

The metrics suggest that the hierarchical model demonstrates a superior capability in managing the complexity associated with classifying into 245 cases by systematically refining the prediction space at each stage. The hierarchical structure lets the model focus on fewer labels at each tier, enhancing precision and recall by using contextual links between categories and subcategories, making predictions more clear and interpretable. The structured ontology system improves model comprehension, ensures accurate labeling, and shows adaptability across websites, as demonstrated by the diversity of our dataset.

6.3 Qualitative Error Analysis

To better understand model behavior beyond aggregate performance metrics, we conducted a qualitative error analysis on four ontology cases that appeared across multiple services in our test set, reflected in Table 2. The model matched these cases to a total of 54 service-specific instances. We compared the predicted sub-category and fine-grained category with their gold-standard ontology labels to assess consistency and semantic fidelity.

The model preserved fine-grained semantic consistency in over 70% of test instances, even when higher-level sub-category labels diverged due to overlapping definitions (e.g., *Data Collection* vs. *Data Transparency*). This suggests the classifier successfully captures clause-level specificity, even when broader category scopes are ambiguous. These results affirm the utility of our ontology in supporting interpretable, generalizable classification of policy text.

Errors at the Sub-Category level mostly happen because of high-level semantic ambiguity, such as the overlap between categories like Data Collection and Data Transparency. For instance, a clause about third-party data access could be mistaken as Data Sharing (under Data Collection) or User Control Over Data (under Data Transparency). Crucially, despite this higher-level confusion, the hierarchical classifier still accurately predicts the Fine-Grained Category (e.g., Access by Third Parties). This pattern confirms our hypothesis (RQ2) that the staged structure successfully isolates high-level semantic errors from the crucial, granular distinctions needed for correct policy interpretation, proving the framework's practical utility.

6.4 Inference Efficiency

As we expected, the hierarchical model incurs a higher inference cost due to its multi-stage design, which Table 3 summarizes. While the flat PrivBERT model processed all 3,367 test instances in under 30 minutes using a single-pass prediction, the hierarchical model required approximately 45–50 minutes to complete the same set—reflecting the cost of making 3–4 sequential predictions per instance (Broad Category, Sub-Category, Fine-Grained Category, and Action).

Despite this increased cost, the runtime remains within a practical range, and the additional computational effort yields significantly higher accuracy and interpretability. Moreover, each stage of the hierarchical classifier operates over a smaller, semantically coherent label space, enabling more targeted and precise decisions at each level.

Model	Inference	Setup
Flat (PrivBERT)	≈30 min	≈$1.25/hr
Hierarchical (Ours)	≈50 min	≈$2.08/hr

Table 3: Inference time and setup cost comparison on 3,367 test cases.

Importantly, ontology construction and case labeling using GPT-4 Turbo required less than one hour—a one-time setup cost that is amortized over model training and evaluation. These findings demonstrate that ontology-guided models are not only more accurate, but also more computationally feasible for real-world policy classification pipelines.

7 Conclusion

We presented a repeatable pipeline for constructing a four-layer OWL ontology of privacy policy concepts—anchored in OPP-115, enriched with ToS;DR metadata, and populated via LLM prompting and manual validation. Empirical results show that our ontology-guided hierarchical classifier outperforms a flat PrivBERT baseline at every tier (broad, sub, fine, action), with accuracy gains of 8–17% and improved interpretability. Finally, combining hierarchical relations (`hasSubCategory`, `hasFineGrainedCategory`, `hasAction`) with cross-cutting links (`appliesTo`, `governs`, `compliesWith`, etc.) proved crucial for fine-grained reasoning and clause-level alignment. The ontology schema provides several key advantages:

1. It offers interpretable, layered categorization of complex privacy practices;
2. It mitigates label imbalance by distributing semantic granularity across successive levels; and
3. It supports robust generalization across diverse policy documents while preserving document-specific nuances.

This work establishes ontology-driven classification as both effective and scalable for real-world policy analysis.

8 Limitations

8.1 Complexity

The process of developing a complete ontology that appropriately describes the various components of terms of service agreements and privacy policies is a difficult and time-consuming proposition. It is possible that a smaller set of concepts would achieve better results, but we felt it worthwhile to appeal to ToS;DR for the initial cases. Similarly, it is possible that another hierarchical division would achieve better results. We believe our 4-level division manages breadth and depth complexity in an organized, complete manner.

8.2 Scalability

It is necessary to perform ongoing maintenance and refinement in order to guarantee that the ontology can scale to accept new regulations and adjustments. This scalability can handle increasing complexity and volume, making it possible to add future changes without affecting the system's overall structure.

8.3 Interoperability

It can be challenging to integrate the ontology with the systems and tools that are already in place to conduct an analysis of privacy policies. This requires careful consideration of compatibility and data standards. Use of OWL for ontologies and APIs for data sharing can simplify integration and scalability, making adapting to regulatory updates easier.

References

500px. 500px — 500px.com. https://500px.com/terms. [Accessed 16-06-2024].

Andrick Dev Adhikari. 2020. *Automated Change Detection in Privacy Policies*. Ph.D. thesis, University of Denver.

Benjamin Andow, Samin Yaseer Mahmud, Wenyu Wang, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Tao Xie. 2019. {PolicyLint}: investigating internal privacy policy contradictions on google play. In *28th USENIX security symposium (USENIX security 19)*, pages 585–602.

Grigoris Antoniou and Frank van Harmelen. 2009. Web ontology language: Owl. *Handbook on ontologies*, pages 91–110.

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.

Hao Cui, Rahmadi Trimananda, Athina Markopoulou, and Scott Jordan. 2023. PoliGraph: Automated privacy policy analysis using knowledge graphs. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 1037–1054, Anaheim, CA. USENIX Association.

Beatriz Esteves and Víctor Rodríguez-Doncel. 2024. Analysis of ontologies and policy languages to represent information flows in gdpr. *Semantic Web*, 15(3):709–743.

Thomas R Gruber. 1993. A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220.

Maryam Majedi and Ken Barker. 2024. The privacy policy permission model: A unified view of privacy policies. *Preprint*, arXiv:2403.17414.

Wolfgang Nejdl, Daniel Olmedilla, Marianne Winslett, and Charles C. Zhang. 2005. Ontology-based policy specification and management. In *The Semantic Web: Research and Applications*, pages 290–302, Berlin, Heidelberg. Springer Berlin Heidelberg.

Natalya F Noy, Deborah L McGuinness, and 1 others. 2001. Ontology development 101: A guide to creating your first ontology.

Jonathan A Obar and Anne Oeldorf-Hirsch. 2020. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, 23(1):128–147.

Alessandro Oltramari, Dhivya Piraviperumal, Florian Schaub, Shomir Wilson, Sushain Cherivirala, Thomas B Norton, N Cameron Russell, Peter Story, Joel Reidenberg, and Norman Sadeh. 2018. Privonto: A semantic framework for the analysis of privacy policies. *Semantic Web*, 9(2):185–203.

Monica Palmirani, Michele Martoni, Arianna Rossi, Cesare Bartolini, and Livio Robaldo. 2018. Pronto: Privacy ontology for legal reasoning. In *Electronic Government and the Information Systems Perspective: 7th International Conference, EGOVIS 2018, Regensburg, Germany, September 3–5, 2018, Proceedings 7*, pages 139–152. Springer.

Harshvardhan J Pandit, Axel Polleres, Bert Bos, Rob Brennan, Bud Bruegger, Fajar J Ekaputra, Javier D Fernández, Roghaiyeh Gachpaz Hamed, Elmar Kiesling, Mark Lizar, and 1 others. 2019. Creating a vocabulary for data privacy: The first-year report of data privacy vocabularies and controls community group (dpvcg). In *On the Move to Meaningful Internet Systems: OTM 2019 Conferences: Confederated International Conferences: CoopIS, ODBASE, C&TC 2019, Rhodes, Greece, October 21–25, 2019, Proceedings*, pages 714–730. Springer.

Pinterest. Pinterest Policy — policy.pinterest.com. https://policy.pinterest.com/en/terms-of-service. [Accessed 16-06-2024].

Prolific. Notion – The all-in-one workspace for your notes, tasks, wikis, and databases. — prolific.notion.site. https://prolific.notion.site/Researcher-Terms-7787f102f0c541bdbe2c04b5d3285acb. [Accessed 16-06-2024].

Wenjun Qiu, David Lie, and Lisa Austin. 2023. Calpric: Inclusive and fine-grain labeling of privacy policies with crowdsourcing and active learning. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 1055–1072.

Leonard Richardson. 2023. Beautiful soup 4.

Terms of Service; Didn't Read. 2012. Project website for 'terms of service; didn't read'.

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, Thomas B. Norton, Eduard Hovy, Joel Reidenberg, and Norman Sadeh. 2016a. 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, Berlin, Germany. Association for Computational Linguistics.

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, and 1 others. 2016b. 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.