Recovering Event actions in Privacy Policies: A Multi-Layered Approach to Structured Event Extraction

Abstract

Privacy policies describe data practices using verbal predicates, nominal expressions, and light-verb constructions, complicating structured event extraction. We present a multi-layer system that extends verbal predicate analysis with a nominal-event recovery module to extract structured event frames from both verbal and nominal expressions in privacy policies. Evaluated on 150 manually annotated privacy policy sentences, the full system achieves an action F1 score of 0.8641 and a field-level micro-F1 score of 0.6592. Disabling the nominal-event recovery module reduces action F1 to 0.8492. Enabling this module yields a net gain of 16 true positives at the cost of three additional false positives. Light-verb recovery has a small effect on overall action F1. These results indicate that nominal-event recovery improves recall and extends coverage beyond verbal predicates, while highlighting limitations in participant extraction and event linking. The extracted event frames provide an intermediate representation intended to support subsequent policy formalization and compliance verification, but their suitability for this task remains to be evaluated.

1. Introduction

Privacy policies describe how service providers collect, use, process and share personal information, but they often contain complex legal terms and expressions that make them difficult for users to read and understand. Moreover, it is also hard to verify whether long privacy policies do comply with relevant regulations, such as the General Data Protection Regulation (GDPR). Automatically transforming privacy policies into structured representations that capture the information relevant to policy interpretation can thus support downstream tasks such as formalization, compliance checking and question answering, benefiting both users and regulators.

Information extraction from textual privacy policies can provide a foundation to build structured representations of such policies. Many existing information extraction approaches usually rely on explicit verbal predicates and their arguments to identify policy event actions and related participants. However, policy event actions are not always expressed by direct verbs; they are also described through nominalizations, such as *'the collection of data'*, and light verb constructions such as *'make a request'*, in which the surface verb contributes little of the event meaning and the associated nominal expression carries the main semantic content. Approaches centered on verbal predicates may thus miss the real action, assign incorrect participant roles, or retain semantically weak verbs as event predicates.

To address this limitation, we develop a sentence-level extraction system that combines verbal predicate analysis with a nominal-event recovery module, and then represent the extracted information as structured policy-event frames. Each frame consists of an action and its semantic participants (agent, patient, and recipient). The definitions of action and its participants are based on the notion of thematic roles in predicate-argument structure [1]: action (what is done), agent (the entity performing the action), patient (the entity affected by, or acted upon in, the action), recipient (the entity receiving something through the action). It also contains extracted semantic information that we refer to as event qualifiers. This semantic information includes modality, negation, purpose, timing, condition, and embedded events.

The event frame schema can represent events expressed through verbal and nominal predicates. It allows actions recovered from nominal expressions to be represented as normalized predicates with their semantic participants and event qualifiers. The event frame schema also aims to provide an intermediate representation between natural-language privacy policy and future formalization, while

© 2026 This work is licensed under a "CC BY 4.0" license.

preserving relevant semantic information of natural policy languages and eliminating redundant and unnecessary expression.

To fill these event frames, we rely on a multi-layer extraction system that starts with a tool for extracting verbal predicate-argument structures and progressively incorporates modules to obtain information that the syntactic dependency parser and semantic role labeling fail to capture. A central component of this system is the *nominalization module*, which uses lexical resources and syntactic relations to identify event actions and participants expressed through nominal constructions. The light verb module handles light-verb constructions by recovering the event meaning expressed by the nominal component of a semantically weak verbal predicate. However, the system does not represent every nominal expression as an independent event. In some contexts, nominal expressions are part of a referent's descriptive content, and they serve to modify or restrict their referents, instead of expressing an independent policy event activity. To distinguish such nominal expressions and avoid over-extraction, our system applies a rule-based filter, although these rules cannot fully resolve the distinction between policy relevant events and events only mentioned within the description of referents.

The main contributions of this work are the following:

1. A nominal-event recovery module that combines lexical resources and syntactic relations to recover policy actions and participants expressed through nominal constructions, extending event extraction beyond explicit verbal predicates.

2. A sentence-level multi-layer extraction system that integrates a nominal-event recovery module with a complementary light-verb recovery module, and verbal predicate analysis, and represents the extracted actions, participants, deontic modality, negation, and event qualifiers in our proposed event-frame schema

3. An empirical evaluation of the proposed system through comparison with a reproduced predicate-extraction baseline, controlled ablation experiments, and error analysis. The results show that the nominal-event recovery module improves action recall while maintaining nearly unchanged precision.

2. Related Work

Research on automated privacy policy analysis spans several complementary tasks, including information extraction [2, 3, 4], classification [5], question answering [6], and summarization [7, 8]. Information extraction is most directly relevant to our work because it aims to recover structured semantic information from policy text. We therefore review existing approaches to structured policy representation and verb-centered extraction, followed by work on nominal predicates and light-verb constructions. We then discuss how these forms of extraction relate to policy formalization.

2.1. Structured Information Extraction from Privacy Policies

Early work on privacy policy information extraction typically maps policy segments into predefined privacy practice categories and category specific attributes. Wilson et al. [4] introduced the widely used OPP-115 corpus, in which policy segments are annotated with 10 privacy-practice categories. Because such representations remain relatively coarse-grained, subsequent work has developed more fine-grained structures for representing policy information.

PolicyIE [2] represents each policy sentence using an intent-slot structure, where the *intent* denotes the type of privacy practice (e.g. data collection), and the *slot* denotes specific information (e.g. data categories). Shvartzshanider et al. [9] use *contextual integrity* to represent privacy-policy information flows in terms of five parameters: sender, recipient, subject, information type and transmission principle. PolicyPulse [3] uses a more detailed verb-centered semantic representation in which clause-level predicate arguments are mapped to privacy-specific roles such as data collectors, trigger conditions, and retention periods. PoliGraph [10] extends the scope beyond individual sentences by constructing

a policy-level knowledge graph that links information distributed across sentences and paragraphs, thereby supporting cross-sentence reasoning.

Taken together, these studies show a progression from policy segment labels towards more structured and machine-interpretable representations, and emphasize different semantic dimensions depending on their downstream tasks. Our work follows this direction by representing policy statements as event frames containing actions, semantic participants, and event qualifiers.

2.2. Verb-Centered Semantic Extraction

Early extraction methods often relied on syntactic dependency parsing and manually defined rules. Previous work on privacy-policy reasoning [11] identifies policy actions and their participants primarily through syntactic dependency relations. PoliGraph [10], for example, similarly uses dependency relations to identify information-exchange predicates, entities, and data types. Bakker et al. [12] map dependency relations to semantic roles, such as mapping grammatical subjects to actors. These approaches do not require task-specific annotated policy data and are comparatively transparent and controllable. Their coverage, however, depends on predefined syntactic patterns, which can limit their ability to recover semantically relevant predicates across varied linguistic realizations.

Semantic Role Labeling (SRL) provides a more explicitly semantic representation by identifying predicates and their arguments, often with reference to lexical-semantic resources such as PropBank [13, 14]. Shvartzshanider et al. [9] use SRL to extract information-flow roles from policy sentences, while PolicyPulse [3] maps frequently occurring verbal predicates and their arguments to privacy-specific roles, including data collectors, trigger conditions, and retention periods. These approaches primarily extract semantic structure around explicit verbal predicates. They may therefore miss or misrepresent events expressed through nominalizations (such as *collection of data*), or light-verb constructions (such as *make a request*), in which the nominal component carries much of the event meaning.

An alternative is to rely on annotated legal data to guide semantic role extraction. In the same study, Bakker et al. [12] also fine-tuned BERT models on annotated legal sentences and showed that the domain-specific pre-trained models performed better than the method that combined general SRL and the mapping step. However, such an approach requires high-quality annotations, and any new role type requires retraining the entire model.

More recently, large language models (LLMs) have been used to extract semantic information from policy texts. PolicyGPT [15] uses GPT-4 and reports a high accuracy on the OPP-115 paragraph-level classification task, well above previous baselines. However, Bakker et al. [12] indicated that GPT-4's achieved lower accuracy than a fine-tuned LEGAL-BERT on a finer-grained annotation task. This comparison suggests that general LLMs could perform well on coarse-grained paragraph classification or extraction, but struggle with the extraction tasks that require more fine-grained semantic information.

Overall, the approaches reviewed above largely organize semantic structure around explicit verbal predicates or predefined policy categories. Such representations can provide limited coverage of policy actions expressed through nominal phrases or light-verb constructions. This motivates the work on nominal predicates and complex predicates reviewed next.

2.3. Nominal Expressions and Light Verb Constructions

Nominal expressions can also convey events through nouns formed from verbs (deverbal nouns), as in *the collection of data*. NomBank annotates the argument structure of nominal predicates, providing a noun-centered counterpart to PropBank's annotation of verbal predicates [16].

Bonial et al. [17] expanded PropBank to cover noun, adjective, and complex predicates, including light verb constructions, motivated by the need to represent semantically equivalent events across different surface realizations. Bonial et al. [18] further developed dedicated PropBank annotation guidelines for light-verb constructions, emphasizing that such constructions are semi-productive and that the nominal component often contributes the central event semantics. QANom [19] extends the QA-SRL formalism to nominal predicates and represents their arguments through question–answer pairs. NounAtlas [20]

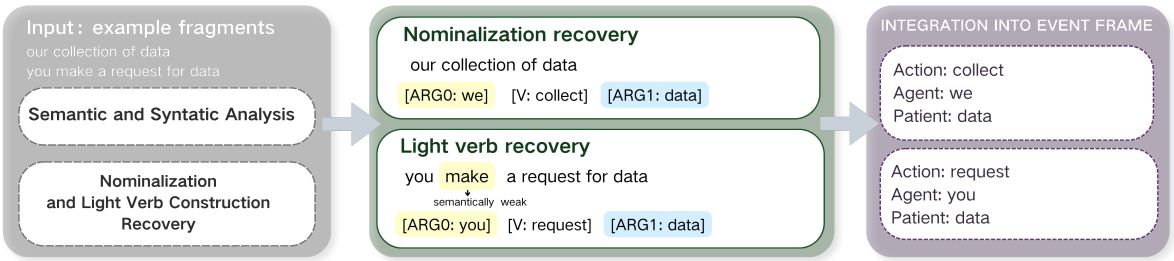


Figure 1: The overview of this proposed system

provides an inventory of nominal predicates and roles that can be integrated with verbal semantic role labeling.

These studies mainly focus on general-purpose semantic annotation rather than privacy policy information extraction. In privacy policies, policy actions are frequently realized through expressions such as *'collection of personal data'* or *'make a request'*. If an extraction system relies on explicit verbal predicates, these constructions may lead to missed events, weak predicate heads, or incomplete participant structures. Thus, our work uses NomBank and PropBank information within a privacy policy extraction pipeline to recover event predicates and their associated participants from nominal and light verb expressions, allowing them to be represented in the same event frame schema as common verbal events.

2.4. Formalization and Semantic Representation

Beyond predicate-argument structure, the modality of a policy sentence is also significant. For example, *'The company collects personal data'*, *'The company may collect personal data'*, and *'The company must not collect personal data'* refer to the same core action but express, respectively, a statement of fact, permission, and prohibition. Existing research on privacy policy information extraction focused on agent, action, and patient, and some also stressed other semantic information, including purpose and conditions. However, most of them did not explicitly distinguish between obligations, permissions, and prohibitions, which are core concepts in legal texts. Works on formalized language, such as FRETish [21], show that controlled language statements can be transformed into formal representations such as temporal logic, but this line of work assumes that the input is already in a controlled or structured language. Therefore, extracting predicate-argument structures, modality, negation, conditions, purpose, and temporal constraints from natural privacy policy texts is a necessary intermediate step towards formalization and compliance checking.

3. Methodology

3.1. Task and System Workflow Overview

Our system combines complementary NLP techniques to identify policy event actions, their relevant participants, and to capture event qualifiers, which modify policy events.

Our system combines Semantic Role Labeling (SRL), co-reference resolution module, syntactic analysis, Abstract Meaning Representation (AMR), and modules for addressing nominalizations. SRL and syntactic dependency analysis provide the basic predicate-argument structures, and complete predicate expressions, such as *participate in* and *be subject to*, are also preserved. Syntactic dependency analysis supplements implicit arguments and event qualifiers, and co-reference resolution links local referring expressions to their antecedents. Additional modules aim to handle nominal events and light verb constructions, such as *'the collection of data'* and *'make a request'* that are not identified by the basic verbal analysis. Then this system recovers event qualifiers, and links between governing and

embedded events before integrating all available information into the final structured event frames. An overview of this process is illustrated in Figure 1.

Given privacy policy sentences, this system outputs structured event frames of these sentences, containing an action with its participants, namely agents, patients and recipients. These event frames also include important event qualifiers, such as modality, negation, condition, purpose, manner and timing, which modify or restrict the scope of an event.

3.2. Semantic Analysis: Semantic Role Labeling

The first step of extraction uses HanLP ModernBERT SRL model [22, 23] to identify explicit verbal predicates and their semantic arguments. Rather than relying on syntactic dependency relations, SRL does assign semantic roles, allowing the proposed system to recover predicate-argument structure across both active and passive constructions. For example, in *'personal data shall be deleted by the controller'*, *Personal data* is the grammatical subject but the semantic patient, while *the controller* is the agent of *delete*, which means *the controller* does this action. By doing this, consistent event frames can be obtained from both active and passive voice sentences.

Then identified predicates are normalized through lemmatization, which transforms the inflected predicate to their basic verbal forms (e.g., *collected* to *collect*). When a verb forms a fixed phrase with a particle or preposition (as in *be subject to*, and *participate in*), the full expression is kept as an action, rather than reduced to the bare verb, because the bare verb would not convey the same meaning of the fixed phrase.

Moreover, SRL does not cover all the event expressions. When encountering nominal expressions, and implicit predicate-argument structures, only using SRL will fail to identify this information accurately. To avoid giving wrong information as the noise, when a field in the event frame cannot be filled, it remains *unresolved* until it can be recovered by another module. Thus additional modules are designed and applied to provide supplementary information and address these unresolved fields. If the system still cannot identify a field, it marks it as unresolved in the output.

3.2.1. Co-reference Resolution

Co-reference resolution is considered in this phase to recover local referring expressions. Privacy policies often use pronouns and possessive expressions such as *we*, *you*, *our*, and *your*. When references are not explicit, correctly extracting the referring expressions is not sufficient, because the semantic information needed for further formalization remains unresolved. For example, in *'we collect information about how our Services are performing when users use them'*, resolving the pronoun *them* to *our Services* can make the event frame clearer by explicitly connecting the extracted argument with its antecedent. In the current design, co-reference resolution mainly focuses on references within individual sentences. If a pronoun or an anaphoric expression cannot be resolved safely, it will be kept as *unresolved*.

3.3. Syntactic and AMR-Based Supplementation

Syntactic dependency analysis is used to supplement structures that are difficult for SRL to identify. Dependency relations are used to recognize inherited subjects in infinitival clauses and to identify event qualifiers such as condition, purpose, and temporal modification. Condition clauses are recognized by markers, such as *'if'* and *'unless'*, purpose clauses with expression such as *'for the purpose...'*, temporal with markers like *'when'*, and embedded events usually serve as modifying clauses, like relative clause. These clauses may also be an independent event when they contain policy propositions, and their relation to the governing event is represented by the corresponding markers. These syntactic dependency relations help to establish relations between governing events and embedded events. When both events have been found, the system links the event in main clause to the corresponding event in subordinate clause rather than representing the two events as unrelated.

AMR provides additional evidence only for selected cases of modality and negation in which a corresponding modal or polarity relation can be identified reliably. Negation records explicit negative

polarity, whereas modality captures the resulting deontic interpretation. Thus, expressions such as *must not* may be represented with both negation = true and modality = prohibition.

Our work uses AMR as semantic evidence for modality extraction. AMR abstracts modal concepts from predicates. For example, deontic expressions can be represented through predicates such as *obligate-01* or *permit-01* through the AMR analysis. Modal predicates are represented by AMR as modal predicates *(obligate-01, permit-01*) and mapped directly to the corresponding modalities. Other expressions, such as *mandatory* or *required*, have no such direct counterpart, but AMR can represent them as ordinary predicates such as *require-01*, and a controlled mapping table supported by PropBank frameset information, links these predicates to a modality. For example, the frameset *require-01* is associated with meanings such as *need* and *obligation*, and links to the FrameNet frame *Imposing-obligation*, which serves as the evidence for mapping *require-01* to *obligation*.

However, we have to add that modalities are not necessarily expressed by a small set of surface modal verbs. In addition to modal predicates such as *must*, *may,* and *shall*, policy rules also express deontic meaning through other adjectives, nouns, and verbal predicates, for example *mandatory*, *required*, and *the right to*. This issue will be considered in future work.

3.4. Nominal Event Recovery

This module focuses on nominal event expressions, such as nominalizations, possessive event nouns, and gerundive constructions. It aims to recover expressions' predicates and arguments, and then create an event frame. The current recovery rule mainly handles nominal expressions with an *of*-complement, and this kind of construction also provides the agent and patient information. The module uses information from NomBank to recover the verbal root of the nominal phrases [24]; for example, nominal heads such as *collection* can be mapped to the verbal root *collect*.

This module also handles possessive event nouns, and the possessive constituents and *of*-complements are considered as subject-like arguments and object-like arguments respectively [25]. In this example, *our* is recovered to *we*, so that the full expression can be represented as *we collect data*. If the agent is expressed by a *by*-phrase, such as in *the disclosure of data by the controller*, the module uses the *by*-phrase as the agent.

Gerundive nominals are treated in the same way. Not every -ing form denotes an event. For example, *the sharing of information* does express an event, but *a flying bird* is only a modifier. Therefore, the module only repairs gerundive nominals when there is enough syntactic evidence, such as an *of*-complement and a nominal use of the *-ing* form.

Furthermore, the existence of a verbal counterpart does not mean that a nominal expression should be represented as a policy event. In our annotation guideline, a nominal expression is recovered as an independent policy event only when it expresses a policy-event activity, either as an activity directly described by the policy sentence, or as an event connected with another policy event, such as appearing in a condition or a purpose. When nominal expressions function as modifiers to restrict or describe their referents, they should be preserved with their referents. For example, in *We provide information about the collection of personal data*, *collection* does denote an action, but the sentence does not state that the company collects data, and the *collection of data* serves to specify what the information is about. These criteria only guide the manual annotation, but the current system cannot fully determine whether a nominal expression should be represented as a separate policy event. In practice, rule-based filtering is applied to nominal-event candidates to reduce over-extraction. Filtered expressions may remain within the referent's descriptive content. However, these filters cannot cover all semantic contexts, as further discussed in Section 6.

3.5. Light Verb Constructions Recovery

Light-verb constructions pose a further challenge for event extraction because the surface verb is semantically weak, while an accompanying noun carries the core event meaning [26, 18]. Standard SRL and dependency-based analyses may identify the light verb as the predicate and treat the nominal

component as one of its arguments. For example, in *You make a request for the data*, a frame centered on *make* represents *a request for the data* as an argument, rather than identifying *request* as the underlying event predicate.

To address this problem, the light-verb recovery module re-centers the event frame on the nominal event. In the example above, *make a request* is normalized to the predicate *request*; *you* is retained as the agent, while *for the data* is associated with the recovered event as its patient. The resulting frame therefore represents the semantically informative event rather than the light verb.

The module relies on a predefined lexicon of light-verb–noun pairs, such as *make a request* and *have consent*. This lexicon was constructed from unlabeled text in OPP-115. Verb–noun collocations were extracted and ranked by frequency, after which combinations headed by common light verbs, including *make*, *have*, *give*, and *do*, were retained. The most frequent relevant combinations were then selected. Although these light verb pairs came from unlabeled text in OPP-115, which includes the evaluation sentences, no annotated standard information from the evaluation set was used. At inference time, spaCy is used to identify candidate verb–noun combinations and their dependency relations. If a detected pair occurs in the predefined lexicon, the module maps the nominal component to its corresponding verbal predicate and uses the dependency structure to recover the associated agent and arguments.

4. Experiment

4.1. Evaluation Data

We constructed a separate evaluation set with 150 policy sentences from the OPP-115 privacy policy corpus[4], and annotated these sentences using our proposed event-frame schema according to annotation guidelines provided in the Appendix. We first selected policy texts concerning data collection, use, sharing, user control, data access, and retention, and segmented them into sentences. Then we used linguistic rules to select candidate sentences and create a candidate pool that contains nominalizations, light-verb constructions, passive voice, or multiple events. We next used category-guided random sampling with a fixed seed to select 150 candidates, with 60 nominalization candidates and 30 candidates each for light-verb constructions, passive voice, and multiple events. We also manually reviewed the 150 candidate sentences and excluded the incomplete sentences cut by the selection script (e.g., missing punctuation and section headings incorrectly included within sentences), and replaced them with the sentences from the candidate pool. The evaluation set finally contains 31 non-challenge sentences and 119 challenge sentences. The sentences in challenge subset at least include one of nominalizations, light verb constructions, passive voice, and multiple events, whereas 31 non-challenge sentences were confirmed that they contained none of these linguistic phenomena during the manual review. Because the sampling procedure deliberately favors these linguistic phenomena, the evaluation is intended to assess the system on challenging extraction cases rather than estimate its performance on privacy policies in general. The 150 sentences were annotated by a single annotator following the event-frame annotation guidelines, containing 496 event frames with 532 atomic action units. The 150-sentence evaluation set was not used during the development, and a separate development set was used to refine the system.

4.2. Evaluation Settings

We compare our system's action identification performance with that of a reproduced version of PolicyPulse's SRL extraction component [3], because the classifier model required to reproduce the complete PolicyPulse is not available in their released resources. Such comparison focuses on identifying and recovering policy-event actions rather than privacy policy classification. The systems produce different output structures, so we compare only the lexical head of each extracted predicate (e.g., *give* for *give up*) with the annotated actions in the same sentence. This criterion is less restrictive than exact action matching, as it does not require the full predicate phrase to match the annotated action.

We evaluate the action identification and event-field level identification. When an annotated frame has multiple actions, each action is expanded into an atomic action unit for the action level evaluation, while its participants and event qualifiers are retained according to the original frame. The actions of the system and annotated standard are aligned one-to-one by exact string matching after case and whitespace normalization. Repeated actions are matched according to the input order. Each matched action is considered as a true positive (TP), each unmatched identified action counts as false positive (FP), and each missing action counts as false negative (FN). Precision, recall, and F1 are computed from the resulting true positive (TP), false positive (FP), and false negative (FN) counts.

To evaluate event-field extraction, we compute end-to-end micro-F1 across all event fields, including actions, their semantic participants, and event qualifiers. The extracted information in each field is normalized by using the same whitespace rules. Each field is compared with the annotated standard strictly, and partial text match is not accepted. After action alignment, true positives, false positives, and false negatives are aggregated across all event fields, including agent, patient, recipient, and event qualifiers, namely modality, negation, condition, purpose, temporal information, manner, duration, and embedded-event relations. Extracted information associated with a missed event contributes false negatives, whereas extracted information associated with a spurious event contributes false positives. Fields are identified as *null, none*, or *unresolved*, are excluded from evaluation, if in annotated standard these fields are also labeled as *null, none* or *unresolved*. If the annotated standard contains *null, none* or *unresolved* but the system gives an output, this causes a false positive (FP). Conversely, if the field in the annotated standard provides some information but the system fails to extract, this counts as a false negative (FN). For negation, both *true* and *false* are evaluated.

4.3. Ablation Study

We conduct an ablation study on the same evaluation set. We thus disable the nominalization recovery module and the light verb recovery module individually and jointly, while keeping all other components unchanged. This yields three configurations: without the nominalization recovery module, without the light verb recovery module, and without both modules. This comparison assesses the contribution of the proposed linguistic recovery modules to event extraction while keeping the syntactic dependency parser and semantic analysis tools fixed. Each configuration is evaluated using the metrics above in Section 4.2, and the full results are reported in the Appendix A.

5. Results

5.1. Overall Event Frame Extraction Performance

Figure 2 reports the overall extraction performance of the full pipeline on the 150-sentence evaluation set. The system produced 535 action units, of which 461 were aligned with actions in the annotated standard, with a precision of 0.8617, recall of 0.8665, and action F1 score of 0.8641. The system achieved an end-to-end micro-F1 of 0.6592 across all the evaluated event fields. Figure 3 also reports extraction performance for the semantic participants and qualifiers of each event. Negation achieved the highest F1 among all non-action fields (0.8491), followed by modality (0.6474), agent (0.6351), recipient (0.6071), and patient (0.5106). Among the remaining event qualifiers, purpose achieved an F1 of 0.4393, condition 0.3393, temporal information 0.3095, and manner 0.2621. Embedded-event linking achieved an F1 of 0.3411, whereas duration scored 0.0000, because the system failed to identify any of the three annotated duration examples. These results indicate that action identification is considerably more reliable than the extraction of complete semantic information. Nevertheless, the nonzero scores for event qualifiers and embedded-event relations show that the system can recover some of the relational structure required by the event representation.

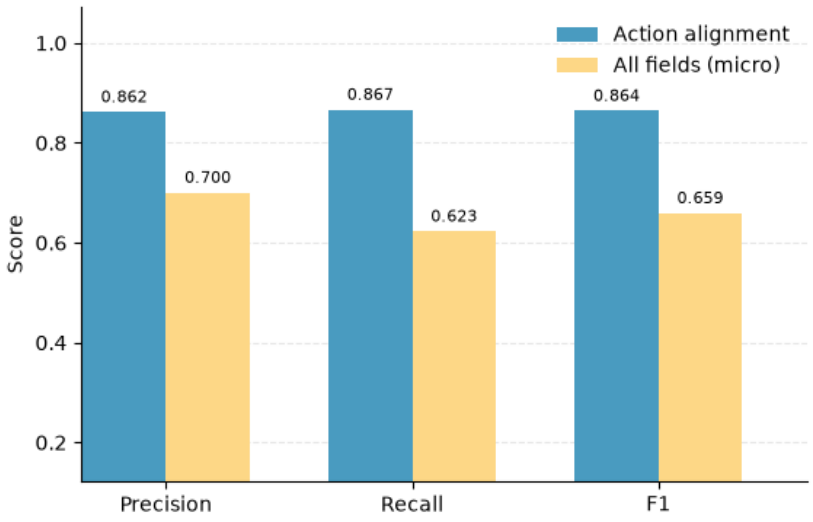


Figure 2: Overall performance of our system

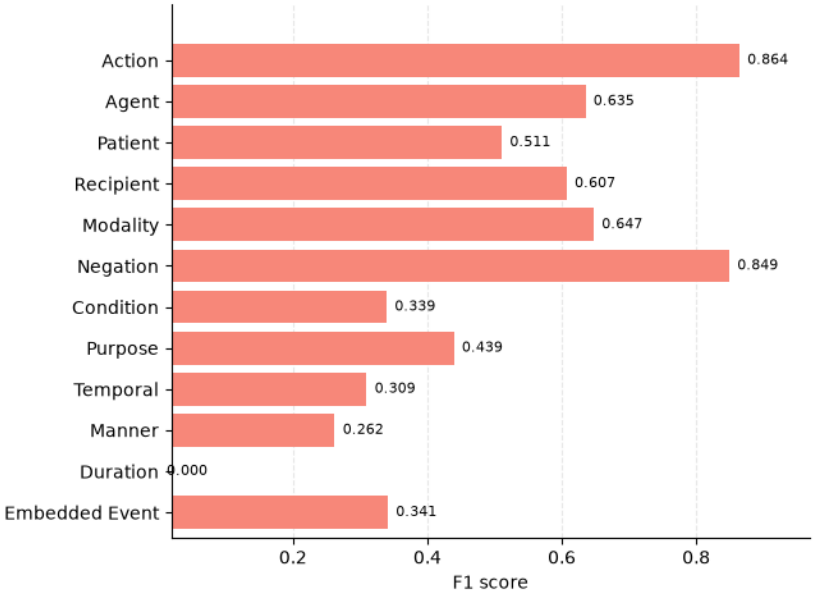


Figure 3: Field-level extraction performance

5.2. Comparison with the PolicyPulse SRL Baseline

We now compare our system with a reproduction of the SRL extraction component of PolicyPulse. In this evaluation, the reproduced PolicyPulse SRL component achieved a precision of 0.5659, recall of 0.8308, and an F1 score of 0.6733. Under the same lexical-head matching, our system achieved a precision of 0.8822, a recall of 0.8872, and an F1 score of 0.8847 on the same test set. This comparison matched actions using only their lexical heads and the resulting action F1 score (0.8847) is higher than the 0.8641 reported in Section 5.1, where the full normalized action must match exactly (Panel C of Table 2 in Appendix A). These results provide a comparison of action identification and should not be interpreted as an end-to-end comparison with the complete PolicyPulse system.

Configuration	# Act.	Act. F1	Micro-F1
Full system	535	**0.8641**	**0.6592**
− Nominal recovery	516	0.8492	0.6453
− Light-verb recovery	531	0.8636	0.6588
− Nominal + LVC	512	0.8487	0.6449

Table 1
Ablation results for nominal event and light verb recovery.

5.3. Ablation Experiment Results

Table 1 reports the ablation results. Removing the nominalization recovery module reduced the number of extracted actions from 535 to 516 and decreased action F1 from 0.8641 to 0.8492, with event-field level micro-F1 dropping from 0.6592 to 0.6453. Removing the light verb recovery module had a smaller effect, with action F1 decreasing from 0.8641 to 0.8636 and event-field level micro-F1 from 0.6592 to 0.6588. When both modules are removed, the system extracted only 512 actions, with an action F1 score of 0.8487 and event-field level micro-F1 of 0.6449. Enabling the Nominalization recovery module increased recall from 0.8365 to 0.8665, while precision remained nearly unchanged (0.8624 without the module and 0.8617 with it), at the cost of three additional unmatched predictions identified by the module.

Removing both modules results in the lowest action F1 and overall micro-F1, indicating that the two modules make contributions. The drop is larger on the challenge subset (0.8584 to 0.8419) than on the non-challenge subset (0.8941 to 0.8876), see in Table 3.

6. Discussion and Error Analysis

The ablation results confirm the contribution of the nominalization recovery module to identification of actions such as *collection/collect*, *processing/process*, and *disclosure/disclose*. When the nominalization module was used 21 additional actions were identified, 18 of these did match to the annotated standard and three did not. However, two previously matched *require* actions are no longer matched, resulting in a net gain of 16 true positives. Overall, the module improved recall while leaving precision almost unchanged.

Beyond action identification, the field-level results show that extracting complete event frames remains challenging. Negation, modality, agent, and recipient were recovered more reliably than purpose, condition, temporal information, and manner, while embedded-event relations proved particularly difficult. Even correctly identified actions could have missing or incorrect participants. The difference between action F1 (0.8641) and micro-F1 (0.6592) across all the fields also reflects the difficulty of extracting information beyond the event predicate. Identifying an action does not mean that its participants, event qualifiers, and semantic relations have been correctly represented.

The errors reveal two limitations. One limitation concerns expressions that can function as both predicates and indicators of deontic modality. For example, in *Collection of personal information required to access certain website services*, the system recovered *collect* but omitted *require*, treating it as an expression of obligation rather than as an independent action, as specified in the annotated standard.

Second, not every nominal expression should be recovered as an independent policy event, but the current nominal-recovery module still recovers these expressions as actions, leading to over-extraction. Of the three unmatched actions, two were cases of over-extraction: *Display of Data*, appearing as a quoted section title, was treated as an event, while *property in the rights* was incorrectly mapped to *possess*. The third involved *the use of information we received from third parties*, where the system failed to distinguish an independent *use* event from a manner qualifier of another *use* event. These cases highlight the need to understand contextual information to distinguish independent events from other uses of nominal expressions.

According to the error analysis above, a further question is whether all nominal expressions should be recovered as an independent policy event. We can divide this problem into two aspects: Does this nominal expression describe a policy event? For example, the expressions, such as *Display of Data* mentioned above, are just a quoted content within sentences and will not be considered as policy events.

Another question is whether these nominal expressions should always be recovered as independent policy events? These nominal expressions may serve different semantic functions according to the context. Consider the two following sentences:*We provide information about the collection of personal data* and *The methods used for the collection of personal data are described below*, in both sentences, the nominal expression *collection of personal data* denotes a *collection* event. However, the first *collection of personal data* specifies the topic of the *information*, and the second one is a purpose for which the methods are used.

These examples suggest that the challenge is not only to identify and recover nominal expressions, but also to determine whether they should be represented as independent events and how their semantic relations should be preserved. As mentioned above, extracting event qualifiers also remains a challenge. Therefore, our future work will investigate how semantic relations can be more accurately represented to capture a wider range of connections between each event, and preserve sufficient semantic information to support formalization of privacy policies, as discussed in the following section.

The evaluation annotations were produced by a single annotator, and inter-annotator agreement was not assessed. This limits our ability to measure annotation consistency, particularly for cases requiring judgments about nominal-event action recovery .

6.1. Future Work

To better support the further formalization of privacy policies, the next step is to refine the proposed event frame schema and ensure it can include the core semantic information, such as more explicit semantic relations among events. When designing this event frame schema, we aimed to capture as much relevant semantic information as possible. However, we have not yet evaluated whether it preserves sufficient semantic information required for downstream formalization. For the purpose of the refinement of our event schema, we need to further analyze how the same predicate can express different semantic functions across contexts, together with more precise annotation guidelines. For example, one phenomenon is that how same predicate should be represented in different context position. Consider the following sentences:

1. We require users to provide this information.
2. We are required to provide this information.
3. We provide information required to collect your data.

In the first sentence, *require* is annotated as an independent action and assigns the *obligation* modality to the *provide* event. In the second sentence, *are required to* express an obligation applying to the *provide* event.In the third sentence, *information required* indicates that the information is needed for the *collect* event. Although the current guidelines specify how these constructions are annotated, further work is still needed to establish more accurate criteria for distinguishing actions, modality, and participant-level descriptions across different contexts. Therefore, we plan to review the 150 sentences in the evaluation test, refine the annotation guidelines, and revise the event schema according to the review to make the event-frame schema preserve complete semantic information and relations.

After the revision, we will also use these 150 annotated sentences to guide further system development. We will construct a new independent evaluation set of 200-300 sentences from real-world privacy policies, which will be independently annotated by two annotators according to the revised annotation guidelines, and compute inter-annotator agreement to measure annotation consistency. The revised system will be evaluated on this evaluation set, and we will also include an LLM baseline.

Finally, we plan to formalize a small number of event frames of these sentences to assess whether the resulting event frames preserve the information needed for formalization. This attempt can help us identify whether there are still some limitations of the proposed event schema and guide us for further improvements towards formalization and compliance verification.

7. Conclusion

This work presents a multi-layer system for extracting structured event frames from privacy policies, with particular attention to the recovery of event actions expressed through nominalizations. By combining a nominal recovery module with verbal predicate analysis, our system identifies policy events expressed through different linguistic constructions and represents them in our proposed event-frame schema. The event frames are intended to serve as a structured intermediate representation for supporting future policy formalization.

On the 150-sentence evaluation set, the system achieved an action F1 of 0.8641 and an event-field level micro-F1 of 0.6592. Ablation experiments indicate that the nominalization recovery module provides the main contribution to the recovery of the actions expressed nominally by the increase of recall, whereas the light verb recovery module has a smaller effect on overall performance. The error analysis highlights remaining challenges in nominal event selection and the preservation of semantic relations between events. The future work will focus on the event selection and accurate semantic relation representation, followed by further evaluation and preliminary formalization experiments.

References

[1] D. Dowty, Thematic proto-roles and argument selection, language 67 (1991) 547–619.

[2] W. Ahmad, J. Chi, T. Le, T. Norton, Y. Tian, K.-W. Chang, Intent classification and slot filling for 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), 2021, pp. 4402–4417.

[3] A. Adhikari, S. Das, R. Dewri, Policypulse: Precision semantic role extraction for enhanced privacy policy comprehension., in: Proceedings of the Network and Distributed System Security Symposium (NDSS), 2025. doi:10.14722/ndss.2025.240916.

[4] S. Wilson, F. Schaub, A. A. Dara, F. Liu, S. Cherivirala, P. G. Leon, M. S. Andersen, S. Zimmeck, K. M. Sathyendra, N. C. Russell, et al., 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), 2016, pp. 1330–1340.

[5] F. Liu, S. Wilson, P. Story, S. Zimmeck, N. Sadeh, Towards automatic classification of privacy policy text, School of Computer Science Carnegie Mellon University (2018).

[6] W. Ahmad, J. Chi, Y. Tian, K.-W. Chang, Policyqa: A reading comprehension dataset for privacy policies, in: Findings of the Association for Computational Linguistics: EMNLP 2020, 2020, pp. 743–749.

[7] A. Adhikari, S. Das, R. Dewri, Natural language processing of privacy policies: A survey, arXiv preprint arXiv:2501.10319 (2025).

[8] M. Keymanesh, M. Elsner, S. Sarthasarathy, Toward domain-guided controllable summarization of privacy policies., in: NLLP@ KDD, 2020, pp. 18–24.

[9] Y. Shvartzshanider, A. Balashankar, T. Wies, L. Subramanian, Beyond the text: Analysis of privacy statements through syntactic and semantic role labeling, in: Proceedings of the Natural Legal Language Processing Workshop 2023, 2023, pp. 85–98.

[10] H. Cui, R. Trimananda, A. Markopoulou, S. Jordan, {PoliGraph}: Automated privacy policy analysis using knowledge graphs, in: 32nd USENIX Security Symposium (USENIX Security 23), 2023, pp. 1037–1054.

[11] Y. Huang, C. Perini Brogi, R. De Nicola, Reasoning on privacy policies, in: Proceedings of OVERLAY 2025, volume 4142 of *CEUR Workshop Proceedings*, 2025, pp. 209–218. URL: https://ceur-ws.org/Vol-4142/.

[12] R. M. Bakker, A. J. Schoevers, R. A. van Drie, M. P. Schraagen, M. H. de Boer, Semantic role extraction in law texts: a comparative analysis of language models for legal information extraction, Artificial Intelligence and Law (2025) 1–35.

[13] L. Màrquez, X. Carreras, K. C. Litkowski, S. Stevenson, Special issue introduction: Semantic role labeling: An introduction to the special issue, Computational Linguistics 34 (2008) 145–159. URL: https://aclanthology.org/J08-2001/. doi:10.1162/coli.2008.34.2.145.

[14] M. Palmer, D. Gildea, P. Kingsbury, The proposition bank: An annotated corpus of semantic roles, Computational linguistics 31 (2005) 71–106.

[15] C. Tang, Z. Liu, C. Ma, Z. Wu, Y. Li, W. Liu, D. Zhu, Q. Li, X. Li, T. Liu, et al., Policygpt: Automated analysis of privacy policies with large language models, arXiv preprint arXiv:2309.10238 (2023).

[16] A. Meyers, R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, R. Grishman, Annotating noun argument structure for NomBank, in: M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa, R. Silva (Eds.), Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04), European Language Resources Association (ELRA), Lisbon, Portugal, 2004. URL: https://aclanthology.org/L04-1228/.

[17] C. Bonial, J. Bonn, K. Conger, J. D. Hwang, M. Palmer, PropBank: Semantics of new predicate types, in: N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, S. Piperidis (Eds.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), European Language Resources Association (ELRA), Reykjavik, Iceland, 2014, pp. 3013–3019. URL: https://aclanthology.org/L14-1011/.

[18] C. Bonial, M. Palmer, Comprehensive and consistent PropBank light verb annotation, in: N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis (Eds.), Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), European Language Resources Association (ELRA), Portorož, Slovenia, 2016, pp. 3980–3985. URL: https://aclanthology.org/L16-1628/.

[19] A. Klein, J. Mamou, V. Pyatkin, D. Stepanov, H. He, D. Roth, L. Zettlemoyer, I. Dagan, QANom: Question-answer driven SRL for nominalizations, in: D. Scott, N. Bel, C. Zong (Eds.), Proceedings of the 28th International Conference on Computational Linguistics, International Committee on Computational Linguistics, Barcelona, Spain (Online), 2020, pp. 3069–3083. URL: https://aclanthology.org/2020.coling-main.274/. doi:10.18653/v1/2020.coling-main.274.

[20] R. Navigli, M. Lo Pinto, P. Silvestri, D. Rotondi, S. Ciciliano, A. Scirè, NounAtlas: Filling the gap in nominal semantic role labeling, in: L.-W. Ku, A. Martins, V. Srikumar (Eds.), Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 16245–16258. URL: https://aclanthology.org/2024.acl-long.857/. doi:10.18653/v1/2024.acl-long.857.

[21] J. J. Joosten, M. L. Chamosa, S. S. Fernández, Coherency through formalisations of structured natural language, a case study on fretish, arXiv preprint arXiv:2605.10462 (2026).

[22] H. He, J. D. Choi, The stem cell hypothesis: Dilemma behind multi-task learning with transformer encoders, in: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Online and Punta Cana, Dominican Republic, 2021, pp. 5555–5577. URL: https://aclanthology.org/2021.emnlp-main.451.

[23] B. Warner, A. Chaffin, B. Clavié, O. Weller, O. Hallström, S. Taghadouini, A. Gallagher, R. Biswas, F. Ladhak, T. Aarsen, G. T. Adams, J. Howard, I. Poli, Smarter, better, faster, longer: A modern bidirectional encoder for fast, memory efficient, and long context finetuning and inference, in: W. Che, J. Nabende, E. Shutova, M. T. Pilehvar (Eds.), Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vienna, Austria, 2025, pp. 2526–2547. URL: https://aclanthology.org/2025.acl-long.127/. doi:10.18653/v1/2025.acl-long.127.

[24] A. Meyers, R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, R. Grishman, The NomBank project: An interim report, in: Proceedings of the Workshop Frontiers in Corpus Annotation at HLT-NAACL 2004, Association for Computational Linguistics, Boston, Massachusetts, USA, 2004, pp. 24–31. URL: https://aclanthology.org/W04-2705/.

[25] J. Grimshaw, Argument structure., the MIT Press, 1990.

[26] M. Butt, The light verb jungle, in: G. Aygen, C. Bowern, C. Quinn (Eds.), Papers from the GSAS/Dudley House Workshop on Light Verbs, volume 9 of *Harvard Working Papers in Linguistics*,

2003, pp. 1–49.

A. Detailed Evaluation Results

Panel A. Overall evaluation and ablation results

Configuration	Action P	Action R	Action F1	Micro-P	Micro-R	Micro-F1
Full system	0.8617	0.8665	**0.8641**	0.7003	0.6227	**0.6592**
− Nominal recovery	0.8624	0.8365	0.8492	0.6957	0.6017	0.6453
− Light-verb recovery	0.8644	0.8628	0.8636	0.7017	0.6209	0.6588
− Nominal & light-verb recovery	0.8652	0.8327	0.8487	0.6972	0.5998	0.6449

Panel B. Field-level F1 scores

Field	Full	−Nom.	−LVC	−Nom. & LVC
Action	**0.8641**	0.8492	0.8636	0.8487
Agent	**0.6351**	0.6220	0.6333	0.6202
Patient	0.5106	0.4954	**0.5131**	0.4980
Recipient	**0.6071**	0.5893	**0.6071**	0.5893
Modality	**0.6474**	0.6322	0.6279	0.6127
Negation	**0.8491**	0.8340	0.8485	0.8333
Condition	**0.3393**	**0.3393**	**0.3393**	**0.3393**
Purpose	0.4393	0.4368	**0.4457**	0.4432
Temporal	0.3095	**0.3133**	0.3095	**0.3133**
Manner	0.2621	0.2483	**0.2721**	0.2585
Duration	0.0000	0.0000	0.0000	0.0000
Embedded event	0.3411	0.3359	**0.3438**	0.3386

Panel C. Shared lexical-head diagnostic

System	Precision	Recall	F1
Proposed system	**0.8822**	**0.8872**	**0.8847**
PolicyPulse-SRL	0.5659	0.8308	0.6733

Table 2
Overall evaluation, nominal and light-verb ablations, field-level F1 scores, and the shared lexical-head diagnostic.

Configuration	Non-challenge (31)	Challenge (119)
Full system	0.8941	0.8584
− Nominal recovery	0.8876	0.8419
− Light-verb recovery	0.8941	0.8578
− Both modules	0.8876	0.8411

Table 3
Action F1 by evaluation subset. The non-challenge and challenge subsets contain 84 and 448 gold atomic actions, respectively.

Table4 illustrates the proposed event frame schema through an annotated privacy policy sentence containing two events. The table shows how event actions, semantic participants, and event qualifiers are represented in separate frames.

B. Example Event-Frame Annotation

Field	S17-E01	S17-E02
ID	S17-E01	S17-E02
Sentence ID	S17	S17
Sentence	We shall collect your payment information if you choose to purchase access to Anthropic's products and services.	We shall collect your payment information if you choose to purchase access to Anthropic's products and services.
Clause type	main	condition
Action	collect	purchase
Action source	explicit	explicit
Action relation	–	–
Agent	We	you
Agent source	explicit	explicit
Patient	your payment information	access to Anthropic's products and services
Patient source	explicit	explicit
Recipient	–	–
Recipient source	–	–
Modality	obligation	none
Manner	–	–
Duration	–	–
Negation	false	false
Condition	if you choose to purchase access to Anthropic's products and services	–
Purpose	–	–
Temporal	–	–
Embedded event	–	–
Embedded event source	–	–

Table 4
Example event frame annotation illustrating the proposed schema for a privacy policy sentence containing a conditional event.

C. Annotation Guidelines

C.1. Annotation Unit and General Principles

This study performs information extraction at the sentence level. Annotation proceeds on an event-by-event basis, identifying each event's participants, core action, and associated information. Main clauses, subordinate clauses, and event-denoting nominalization are analyzed according to their syntactic structure and semantic relations.

Annotations must preserve the meaning of the original sentence. Co-reference resolution, nominalization reconstruction, and necessary expression normalization are permitted, but must not introduce information absent from the original sentence.

C.2. Agent, Action, and Patient

The three core roles are defined as follows:

- **Agent**: The actor or entity performing the action.

- **Action**: The core action expressed by the event.
- **Patient**: The entity affected by, involved in, or targeted by the action. This study uses Patient as a broad annotation category encompassing both patient and theme arguments.

In canonical active clauses, the subject, predicate, and object generally correspond to Agent, Action, and Patient, respectively. This mapping also applies to subordinate clauses. However, final assignments must reflect syntactic structure and semantic relations rather than surface word order alone.

If a role is neither expressed nor recoverable under these guidelines, it must not be supplied.

C.3. Explicit and Implicit Roles

A role is **explicit** when it has an overt linguistic realization within the relevant event structure. A role is **implicit** when it is not overtly expressed within that structure but can be recovered through syntactic or semantic relations within the sentence.

For example, in *We intend to collect data*, the Agent of *collect* is not overtly expressed in the infinitival clause but is recoverable from the main-clause subject *we*. It is therefore annotated as implicit. Recovery must follow the actual control relation; infinitival subjects must not automatically be assumed to inherit the main-clause subject.

The same principle applies to relative clauses. When an overt relative pronoun realizes a role, that role is annotated as explicit and the pronoun is resolved to its antecedent. When the relevant relative element is omitted and the role must be recovered from the antecedent, the role is annotated as implicit.

Coreference resolution and expression normalization do not change a role's explicit or implicit status. For example, replacing an overt *we* with *the company* does not make the Agent implicit.

C.4. Passive Clauses

Passive clauses are annotated according to semantic roles:

- The passive subject generally corresponds to Patient.
- The core action expressed by the passive predicate corresponds to Action.
- An overt actor introduced by a *by*-phrase or another construction corresponds to Agent.

If the Agent is omitted but recoverable from information within the sentence, it is annotated as implicit. If it cannot be determined, no Agent is supplied.

C.5. Nominalizations

For nominalizations that express actions or events, reconstruct the corresponding verb and argument relations before assigning Agent, Action, and Patient.

For the pattern **agentive possessive + event noun + *of*-complement**, this study applies the following rules:

- The agentive possessive preceding the event noun corresponds to Agent.
- The event noun corresponds to Action.
- The argument complement following *of* is annotated as Patient.

For example, *our collection of data* is reconstructed as *we collect data*. After coreference normalization, the annotation is:

- Agent: *the company*

- Action: *collect*
- Patient: *data*

Because *our* overtly expresses the Agent, this role is explicit. If the nominalization does not express an Agent, the presence of an *of*-complement alone does not justify supplying one.

This rule applies to argument complements within event nominalizations. It does not apply to non-eventive *of*-constructions expressing relations such as possession or part–whole relations.

C.6. Light-Verb Constructions

When a verb primarily serves a grammatical function and the accompanying noun carries the core action meaning, annotate the action expressed by that noun as Action. Identify the remaining roles accordingly.

For example, in an expression such as *make a request*, if *request* carries the core action meaning, annotate *request* as Action.

C.7. Condition

Condition specifies a prerequisite or circumstance under which an event occurs, a rule applies, or an action is permitted, required, or prohibited.

Common condition markers include *if*, *unless*, *only if*, *provided/providing that*, *as long as*, *so long as*, *on condition that*, *in the event that*, and *in case of*. Expressions such as *when*, *where*, and *subject to* may also introduce a Condition when they convey conditional meaning.

Annotation must be based on meaning rather than marker matching alone. For example, *if* introducing an embedded question with the meaning of *whether* does not establish a Condition.

Preserve the complete condition and its logical meaning, including the negation or restriction expressed by markers such as *unless* and *only if*.

C.8. Purpose

Purpose specifies the intended goal of an action.

Common purpose markers include *to*, *for*, *in order to*, *so as to*, *so that*, *in order that*, *for the purpose of*, *with the aim of*, and *with a view to*.

These expressions introduce Purpose only when they convey an intended goal. Fixed expressions, selected complements, and other constructions without purpose meaning must not be annotated as Purpose.

Examples:

- In *collect data to improve the service*, *to improve the service* expresses Purpose.
- In *provide information for verification*, *for verification* expresses Purpose.
- In *have access to data*, *to data* does not express Purpose.
- In *be responsible for processing data*, *for processing data* does not express Purpose.

C.9. Temporal

Temporal covers all temporal adverbial clauses and modifiers that constrain an event's timing, including time points, periods, durations, frequencies, start and end times, and temporal ordering.

Common expressions include *when*, *while*, *before*, *after*, *until*, *since*, *once*, *as soon as*, *during*, and *within*, as well as modifiers such as *for 30 days*, *annually*, and *at any time*.

Temporal meanings must be distinguished from other relations. For example, causal *since* is not annotated as Temporal. When *when* specifies a condition under which a rule applies, it is annotated as Condition.

C.10. Modality

To limit the number of modality categories and facilitate subsequent formalization, this study recognizes only the following three categories:

Category	Meaning	Restricted marker set
Prohibition	An action is prohibited	*must not, cannot*
Permission	An action is permitted	*can, may, could, might, would*
Obligation	An action is required	*must, require, mandate, shall*

Common grammatical variants of these markers belong to the corresponding categories, including *can't*, *required*, *is required to*, and *mandated*. When *require* functions as a separate verbal predicate, it is retained as an action, and its deontic modality scopes over the embedded event in its complement. For example, in *We require users to provide information*, *require* is annotated as a separate action, while *obligation* is assigned to the *provide* event.

Annotation must account for meaning and the scope of negation. A marker receives a modality label only when it conveys prohibition, permission, or obligation. Uses expressing ability, possibility, inference, or hypothetical situations are not automatically classified as Permission. For example, *may fail* does not express Permission when it indicates possibility. Likewise, *not required to* expresses neither Obligation nor Prohibition.

Modal expressions outside the restricted marker set do not introduce additional categories.

C.11. Co-reference Resolution

Pronouns and other referring expressions must be resolved to their antecedents. An explicit antecedent or specification within the sentence takes precedence over default mappings.

Because this study performs sentence-level information extraction, the following defaults apply unless otherwise specified:

Original expression	Normalized expression
we	*the company*
our	*the company's*; use *the companies'* when multiple companies are explicitly indicated
you	*the user*
your	*the user's*; use *the users'* when multiple users are explicitly indicated

When a possessive expression is reconstructed as an event participant, use the corresponding non-possessive form. For example, when *our* in *our collection* is recovered as Agent, normalize it to *the company*.

Apart from these default mappings, antecedents that cannot be determined from the sentence must not be supplied using information from outside the sentence.

C.12. Embedded Event

In this study, **Embedded Event** refers to a case in which a pronoun or referring expression within a sentence refers to the complete event expressed by another clause, rather than to an individual entity within that clause.

For example, in *We collect data, which allows us to improve the service*, if *which* refers to the complete event "the company collects data," resolve it to that event and preserve its relation to the subsequent event.

Apply Embedded Event treatment only when the referent is a complete event. If the expression refers only to an entity such as *data*, apply ordinary coreference resolution.

C.13. Normalization of Special Expressions

Complex expressions may be reformulated into equivalent forms that facilitate annotation and formalization, provided that their original meaning and semantic relations are preserved.

For example:

> *information about how the service performs*
> → *service performance information*

Such reformulations must preserve the original information target and its qualifying relations. They must not introduce new actions, participants, or additional meanings.