PropBank goes Public: Incorporation into Wikidata

Elizabeth Spaulding[1], Kathryn Conger[1], Anatole Gershman[2], Mahir Morshed[3], Susan Windisch Brown[1], James Pustejovsky[4], Rosario Uceda-Sosa[5], Sijia Ge[1] and **Martha Palmer[1]**

[1]University of Colorado Boulder, [2]Language Technologies Institute, Carnegie Mellon University, [3]University of Illinois Urbana-Champaign, [4]Brandeis University, [5]IBM Research, T.J.Watson

{elizabeth.spaulding,kathryn.conger,martha.palmer}@colorado.edu

Abstract

This paper presents the first integration of PropBank role information into Wikidata, in order to provide a novel resource for information extraction, one combining Wikidata's ontological metadata with PropBank's rich argument structure encoding for event classes. We discuss a technique for PropBank augmentation to existing eventive Wikidata items, as well as identification of gaps in Wikidata's coverage based on manual examination of over 11,300 PropBank rolesets. We propose five new Wikidata properties to integrate PropBank structure into Wikidata so that the annotated mappings can be added en masse. We then outline the methodology and challenges of this integration, including annotation with the combined resources.

1 Introduction

Recent work (Spaulding et al., 2023) has explored how Wikidata (Vrandečić and Krötzsch, 2014) can be adapted as an ontology for information extraction by: (1) providing an external resource that augments existing eventive Wikidata items with PropBank (Kingsbury and Palmer, 2002; Gildea and Palmer, 2002; Palmer et al., 2005) role information; and (2) documenting gaps of event coverage in Wikidata, where a PropBank concepts have no corresponding Wikidata items. For each PropBank predicate, annotators have been asked to identify a matching concept in Wikidata. All 11,300+ PropBank rolesets have been subject to at least one pass of this task, so that each PropBank predicate has either a Wikidata mapping or a flag indicating a missing concept in Wikidata. This paper outlines the methodology and challenges of integrating the project into Wikidata itself, including filling those gaps in event coverage. The paper will also document the unique annotation challenges faced in attempting to join the two different resources.

2 Background

Wikidata[1] is a large, crowd-sourced, multilingual knowledge base hosted by the Wikimedia Foundation. Originally created to support Wikipedia by serving as a repository linking matching articles of different languages together, it proliferated into a true knowledge graph of its own. It is now the centralized location for data in Wikipedia infoboxes, and is used for various natural language processing applications requiring the use of real-world knowledge. For example, Wikidata is often used as a source of facts to probe (Petroni et al., 2019) or update the knowledge (Meng et al., 2023) of large language models (LLMs). Knowledge bases generally have been used for search and question answering in large, industry-scale search engines (Noy et al., 2019). A major issue in using knowledge bases is quality and long-term maintenance (Piscopo and Simperl, 2019). While Wikidata is relatively reliable compared to other similar knowledge bases, there is still room for improvement in quality control (Shenoy et al., 2022) and coverage of event concepts (Spaulding et al., 2023).

Each item in Wikidata refers to either a concept ("planet") or a real-world instantiation of a concept ("Earth"), and is uniquely identified by a *Q* followed by a string of digits. Relations are called *properties* in Wikidata, and are uniquely identified by a *P* followed by a string of digits. Wikidata items can have statements, which are a property and a value (usually another item or a static data point such as a string or numerical quantity). Thus, <subject, relation, object> triples are <item, property, value> triples in Wikidata nomenclature. Statements can be annotated with qualifiers—for example, specifying the start and end time that a relation holds—and references, making Wikidata an extremely rich, detailed, and verifiable source of knowledge for natural language processing ap-

[1] https://www.wikidata.org/

Proceedings of The 18th Linguistic Annotation Workshop (LAW-XVIII), pages 166–175
March 22, 2024 ©2024 Association for Computational Linguistics

plications. Additionally, Wikidata is a multilingual resource. Users can browse and add data to Wikidata in their preferred language, so concepts can be described by labels written in a broad range of languages, with an average of around 8 labels[2] per Wikidata Item.

2.1 The DWD Overlay

The DARPA Wikidata Overlay (DWD overlay; Spaulding et al. 2023) serves as an external resource augmenting Wikidata items with event role information drawn from PropBank (Kingsbury and Palmer, 2002; Gildea and Palmer, 2002; Palmer et al., 2005), natural language templates for how each event can prototypically be used in a sentence, and temporal relation mappings for event-to-event relations. The overlay currently contains 5,764 eventive Wikidata Items, 2,224 of which have hand-curated PropBank roleset mappings (the rest have semi-automatic PropBank mappings). 9,011 PropBank rolesets are currently documented as having no matching Wikidata Item, and thus, are not in the overlay. It is contained in a JSON file and hosted on Github, and must be continually updated as Wikidata changes. Because it can be edited by anyone with an internet connection, Wikidata is constantly changing, making the maintenance of the overlay, which relies on Wikidata, untenable in the long-term. Wikidata is also far more desirable as a host of this information for computing applications, as Wikidata provides support for large-scale projects with a SPARQL query service, and already has a large community of dedicated contributors which maintain it. Once the mappings contained in the overlay are fully integrated into Wikidata, overlay users can reconstruct the overlay from Wikidata itself by querying the PropBank additions.

2.2 Role of PropBank

PropBank is the source of participant roles for incorporation of events into Wikidata. PropBank is a verb lexicon initially annotated over the Wall Street Journal sentences in the Penn Treebank (Taylor et al., 2003), which separates verbs (and event nominalizations) into coarse senses, each with a set of verb-specific semantic roles (the "roleset"). While the roles are verb-specific, they are numbered 0-6, with ARG0 typically corresponding to Proto-Agent and ARG1 corresponding to Proto-Patient, allowing for some amount of generalization across verbs. An example of a PropBank roleset can be found in Table 1. PropBank was chosen because of its wide coverage of verbs and eventive nouns that could be easily matched to Wikidata concepts and because its roles could be represented as both broad, general roles (e.g., ARG0, ARG1) and as more event-specific (e.g., attacker, victim).

[2] Although this average may be inflated due to some labels being the same in different languages—for example, when the label is a person's name.

2.3 Use of the Overlay in Systems

Since the DWD Overlay covers eventive Wikidata concepts, it has been adapted for use as an event ontology in GLEN (Li et al., 2023). In particular, the mapping between Wikidata entities and PropBank rolesets was utilized to create distantly supervised training data by re-purposing PropBank annotation. In addition, the event detection model computed event type representations from Wikidata Item labels and descriptions. The DWD Overlay greatly accelerated the development of this general-purpose event detection model that now supports over 3000 event types.

In terms of argument extraction, the event templates (e.g., {A0_pag_transporter} transported {A1_ppt_thing_transported} from {A2_dir_source} to {A3_gol_destination} {A4_loc_path} at {AM_loc}) can be utilized to guide the extraction of arguments by formulating the argument extraction task as conditional generation (Li et al., 2021). Besides text generation, the argument extraction task can also be transformed into structured code generation (Wang et al., 2023) which takes advantage of the argument names and their type constraints.

The Wikidata labels and descriptions also contribute greatly to the schema matching and prediction component in the RESIN event extraction, tracking, and prediction system (Du et al., 2022). RESIN defines the similarity between an event mention and a schema event type as the WordNet (Miller, 1995) similarity between their corresponding Wikidata labels. Then, the matching is conducted based on these pairwise item similarities.

3 The Events and Role Frames WikiProject

To address the issues discussed in Section 2.1, a WikiProject has been created to integrate several thousand PropBank-Wikidata mapping annotations into Wikidata itself. This project merges together

ID	combust.01
Description	cause to burn
Roles	ARG0-PAG: agent/cause of combustion ARG1-PPT: thing combusting ARG2-MNR: instrument
Example	[$_{ARG1}$ Most fossil fuels] are combusted [$_{ARG2}$ with ambient air].

Table 1: The PropBank roleset for `combust.01`.

several proposals for how to format the integration, including introducing a handful of new properties. While our event mappings are only for generic event types (e.g., "eating" in Table 2), the proposed properties *can* be extended to instances (e.g., "assassination of Abraham Lincoln" in Table 4).

3.1 Goals and Motivation

The coverage of event classes and their role structures in Wikidata is limited. For example, some of the most common verbs in most languages are the verbs of perception, e.g., "to see" ("I see a house", "je vois une maison", "minä näen taloa", "ich sehe ein Haus"), "to hear" ("I hear the rain", "minä kuulen sadetta"), "to smell" ("I smell the coffee", "je sens le café"), etc. Wikidata has related concepts: visual perception (`Q162668`) "ability to interpret the surrounding environment using light in the visible spectrum"; hearing (`Q160289`), sense of smell (`Q1541064`), etc. We argue that an ability to do something is distinct from actually doing it which is an event in space-time. We examined over 11,300 rolesets contained in PropBank that describe English predicates (mostly verbs) and identified over 7,500 potentially missing Wikidata items. Each of these "gaps" needs to be further examined to determine if it warrants a new item, but the list gives us a starting point.

All action events have core semantic roles: "eating" has the "eater" and the "eaten"; "throwing" has the "thrower", the "target" and the "projectile". These roles are not optional. Every act of "eating" has an "eater" and the "eaten" independently of how and in which language it is expressed. However, most of the existing Wikidata items for action classes do not mention these roles. For example, throwing (`Q12898216`) "launching of a ballistic projectile by hand" does not have any statements that indicate the existence of the thrower, the target, or the projectile, let alone the specifications of the kinds of entities these attributes are likely to be.

Thus, the goal of the "Wikidata Events and Role Frames" project[3] is to use PropBank and other sources to fill these gaps in Wikidata and provide event items with role frames.

3.2 Proposed Properties

Event Role The key to our proposal is the new Wikidata property which we propose to call "event role" whose value is a Wikidata Item that describes the role in greater detail. For example, eating (`Q213449`) would have two statements, exemplified in Table 2. In this example, we used the existing Wikidata items for the "eater" and "eaten" roles. However, in most cases, such items do not currently exist and must be created.

Based on the number of PropBank predicates and roles per predicate, we currently estimate adding up to 25,000 – 30,000 event role items. It might also be possible to cluster multiple event role items and create a "subclass of" event concept hierarchy, using other event lexicon resources such as FrameNet (Baker et al., 1998) and VerbNet (Schuler, 2005). We want to stress that although we start from the English PropBank, the proposed event role items are not lexical or grammatical constructs. The existence of a killer in a killing event is not tied to any language or grammar. It is a part of the "killing" concept.

Selectional Preference Whether existing or newly created, event role items serve as anchors for whatever information we want to associate with the role. For example, we may specify that the eater is expected to be an organism. The existing item, eater (`Q20984678`), does not specify that. We propose to introduce another property "selectional preference" which we show in Table 3.

Multiple statements with "selectional preference" should be interpreted as an "OR", i.e., the filler of the role slot is likely to "descend from" at least one of the selectional preference items. The meaning of

[3] `http://www.wikidata.org/wiki/Wikidata:WikiProject_Events_and_Role_Frames`

eating (Q213449)	
ingestion of food to provide for all organisms their nutritional or medicinal needs	
event role	eater (Q20984678)
	food (Q2095)

Table 2: Sample usage of the proposed property "event role," using already-existing Wikidata items. Proposed additions to Wikidata in *italics*.

eater (Q20984678)	
human or other live being who eats something	
selectional preference	organism (Q7239)

Table 3: Sample usage of the proposed property "selectional preference." Proposed additions to Wikidata in *italics*.

"descend" could be application-specific, but, generally, we mean a combination of "subclass of", "parent taxon" and "instance of" properties. Violations of selectional preferences often signal metaphoric use as in "the house ate the savings". Other information such as dietary restriction statements can also be attached to the event role items.

Event Argument and Argument Role The proposed "event role" property applies only to event classes, not instances. For example, the assassination of Abraham Lincoln (Q1025404) is an instance of assassination (Q3882219). Our proposal will create the "assassin" and the "assassinate" event role items. We propose to create a new property "event argument", together with a qualifier "argument role", in order to represent the roles in an event instance, as shown in Table 4.

One might object that in this example, these properties convey the same information as "target" (P533) and "perpetrator" (P8031). Unfortunately, many instances of "assassination" use different properties or none at all to indicate the assassin and the victim. We propose to use a uniform approach even if it causes some redundancy.

We are aware that creating new properties in Wikidata is a time-consuming and difficult process. Our proposal involves one new property for event classes, two for event roles, and one property and one qualifier for event instances.

3.3 Aligning with Wikidata Standards and Structure

The incorporation of events and relationships from PropBank into Wikidata has the potential to greatly benefit the Abstract Wikipedia project under development (Vrandečić, 2020, 2021). The "abstract content" this project will use requires that information be stored in a language-independent fashion, so that anyone can edit it regardless of their spoken language. Hence, for this information to be grounded in terms of Wikidata entities, individual statements about event occurrences must not only have Wikidata items for particular participants, but also items for events and the roles those participants play in those events. Moreover, for a particular concept (such as PropBank's see.01) and predicates representing this concept across different languages (such as English "see", Finnish "nähdä", and Turkish "görmek"), an important goal is the ability to align the same semantic roles across languages, regardless of their syntactic expression in that language. For example, the syntactic object of English "see" and the subject of the isiZulu passive verb "bonwa" should both refer to the thing being viewed.

There has been one attempt to (1) map items for events to individual senses on Wikidata lexemes using the property "predicate for" (P9970), (2) map event roles for a given event across different languages' predicates using the property 'has thematic relation' (P9971) and (3) use these mappings in generating natural language text (Morshed, 2023).

Project Discussion Wikidata provides spaces for the discussion of the proposed properties[4] and of the project as a whole[5]. On top of the linguistic challenges found in annotating the mapping between PropBank and Wikidata, an essential component is negotiating the incorporation of these properties into Wikidata with the Wikidata community. Many annotation projects in natural language processing are developed within a team of perhaps a dozen researchers with relatively similar goals and interests. Due to the public nature of Wikidata, we must "convince" Wikidata users who may not

[4] e.g. https://www.wikidata.org/wiki/Wikidata:Property_proposal/event_role

[5] http://www.wikidata.org/wiki/Wikidata:WikiProject_Events_and_Role_Frames

assassination of Abraham Lincoln (Q1025404)		
1865 murder of the 16th President of the United States		
event argument	Abraham Lincoln (Q91)	
	argument role	*assassinated*
	John Wilkes Booth (Q180914)	
	argument role	*assassin*

Table 4: Sample usage of the proposed property "event role" and the proposed qualifier "argument role", using already-existing Wikidata Items. Proposed additions to Wikidata in *italics*.

share the same goals and interests that our proposal will benefit the wider community. We invite anyone who is interested to join the discussion of the many challenges in this project, some of which are described in the next section.

4 Challenges to Incorporation of PropBank into Wikidata

An inherent challenge to integration is inevitable discrepancies between projects. As a database of coarse-grained, predicate-specific semantic frames, PropBank entries are eventive in nature. While entries may be nominal, adjectival, or verbal, the majority are verbal. By contrast, the Wikidata database includes comparatively fine-grained entries for events and entities, in addition to a number of other types of Items. The majority of Wikidata entities are nominal and often non-eventive in nature. Differences in Item granularity and Item type necessitate careful comparisons of the scope of a roleset to the scope of an Item. Sometimes Items have better mappings to predicate arguments than to the predicates themselves, raising the question of whether such argument mappings are useful. Managing these differences has been largely guided by the utility of the overlay in downstream tasks.

4.1 Identifying Events

Because all PropBank entries are predicates, it is crucial that they are mapped to eventive Items. In some cases, the line between an event and a non-event in Wikidata is clear. For example, bathe.01 is defined as "have a bath, giving or having a bath" and can be used for verbs (e.g., *The place was **bathed** in sunlight*), eventive nouns (e.g., ***bathing** of the infant*), or light verb constructions (e.g., *I **gave** the dog a **bath***). Wikidata has entries for bathing (Q327651) "washing of the body with a liquid" and bathtub (Q152095) "large container for holding water in which a person may bathe". "Bathing" clearly refers to the event while "bathtub" refers to the physical item in which the event takes place. The line is less clear in cases such as signature (Q1373131) "mark of the creator on a work to identify themselves as such (name, initials, monogram)". The definition of "signature" suggests that the Item refers to the artifact itself but does not seem to preclude a signing event during which one creates or affixes a signature. However, the statements about "signature" indicate that this Item refers to artifact only; its subclass membership precludes extension to the event. Thus, we recommended that a new Item be created for the act of signing. There are many such process-result "logical polysemies" (Pustejovsky, 1995) that can be handled in a similar fashion.

4.2 Item Extensibility

A second consideration comes from differing levels of Item granularity, particularly with respect to the scope and extensibility of an Item. For example, see.01 is defined as "view" but is used for literal instances of visual perception (e.g., *I see you*), instances of observation (e.g., *They saw the value of their stocks decline*), and figurative extensions (e.g., *I see your point*). Wikidata has entries for sighting (Q52266213) "occurrence where a region of land is spotted from a ship" and "visual perception" (Q162668) "ability to interpret the surrounding environment using light in the visible spectrum".

While "sighting" does involve an act of seeing, it is very narrowly defined. "Visual perception" seems like a better match for literal instances. However, it is defined as an ability, specifying a capability rather than the active use of the ability that occurs during a seeing event. Given the specificity of Wikidata entries and the inherent differences between capabilities and actions, "visual perception" should not be extended to cover literal instances of the act of "seeing". Instead, a new Item is needed. This new Item for "seeing" can be linked to "visual perception" using the "uses" (P2283) property. There are no existing Item options for figurative

instances of "see". Given the narrow construal of Items in general, the lack of figurative Items is challenging as many PropBank rolesets are typically extended to include figurative usages.

A major goal of the mapping is to create full correspondences between PropBank and Wikidata. In some cases, this can be accomplished with a one-to-one correspondence between a single roleset and a single Item. However, as long as the full scope of a roleset is accounted for, 1-many or many-1 mappings are also acceptable.

Where Wikidata has gaps, new items can be created. However, accounting for figurative language, which is often language specific, may be more appropriately handled through the use of Wikidata Lexemes (language-specific Wikidata entities whose IDs begin with 'L') rather than concept Items. Thus, we propose the alternative solution of adding figurative senses of terms like "see" to Wikidata Lexemes, rather than creating concept Items for these and other such multi-sense English lexemes.

4.3 Multiple Mappings

For full project compatibility, one should be able to do semantic role annotation on text by using PropBank practices but replacing the rolesets with their mapped Items, thus linking the events in the text to associated world knowledge. This is similar to the use case employed by Li et al. (2023) in their usage of the overlay (see Section 2.3). Similarly, a Wikidata user should be able to substitute a PropBank roleset for an Item, allowing for the incorporation of predicate-specific thematic role information into their Wikidata-based project. A prerequisite for this compatibility is very precise mappings that preserve the scope of the roleset without incidentally adding extraneous meaning, hence the 1-many or many-1 mappings discussed above. However, multiple mappings require careful consideration. Multiple mappings can be added for two reasons.

Two or more Items are indistinguishable First, if two or more Items are identical or would be indistinguishable in underspecified contexts, both are retained for consideration as mappings to rolesets. For example, death (`Q4`) "permanent cessation of vital functions" and dying (`Q267505`) "final process of life" are potential mappings to die.01 which represents the "dying or death" sense of the lexeme "die". "Dying" is a subtype of change of state and is, therefore, more applicable to verbal instances when it is clear that the dying event is actively occurring. However, die.01 can account for the verb "die" and the eventive nouns "die", "dying", and "death" where the ability to differentiate between applicability of "death" and "dying" is difficult. Rather than making an arbitrary decision, both are mapped to die.01.

An Item covers only a portion of the scope of a roleset The second reason for multiple mappings is when an Item covers a portion of the scope of a roleset and one or more Items cover the remaining scope of the sense. For example, physical contact (`Q38183514`) is defined as the "state of physical items and materials with no spatial separation, in which surface interactions may occur" and touch (`Q877674`) is defined as "physical contact involving one or more sentient agents (for contact between non-sentient objects, use (`Q38183514`)".

Both were evaluated as potential matches for touch.01, which refers to the "come into contact with" sense. "Physical contact" is a member of the subclass "intentional human activity" and specifies the involvement of one or more sentient agents. "Touch" is complementarily distributed relative to "physical contact" as it is used in cases of physical contact that do *not* involve sentient agents. Touch.01 does not differentiate between sentient and non-sentient agents, so, individually, these Items are more narrowly defined than the roleset. Together, their scope is similar to that of touch.01, so they are both mapped.

In the case of death (`Q4`) and dying (`Q267505`), the Items have a similar scope. However, in the case of physical contact (`Q38183514`) and touch (`Q877674`), touch covers a larger portion of the scope of touch.01 than physical contact. Weight is assigned to this Item via representation in a prioritized slot in the overlay, with additional Items added based on the degree of coverage they contribute. Additionally, future releases hope to include methods to add weights to these slots.

A mapping is considered "complete" when the collectively mapped Items can account for at least 90% of the scope of the roleset and the collectively mapped Items do not add additional information that the roleset is unable to account for. If no complete mapping (one-to-one or otherwise) can be identified, we recommend that a new Item be created.

Synchronizing argument structures for multi-roleset mappings Expanding to multiple-roleset-

to-one-Item mappings also required a modification to the DWD Overlay to accommodate them. Mappings between a single Item and multiple rolesets presented a more serious challenge as they have the potential to attach conflicting argument structures to the Item.

For example, attempt (Q12897867) "action whose success is not guaranteed" is highly related to try.01 ("attempt") and attempt.01 ("try"). These rolesets share an argument structure, enriching the Item with consistent argument structure information. By contrast, want.01 and wish.01, which can both be mapped to wish (Q241625) "desire for a specific item or event", have different argument structures. Specifically, want.01 has two additional roles missing from the wish.01 frame – ARG3-PPT *in-exchange-for* and ARG4-DIR: *from*. Mapping them both to a single Item presents conflicting argument information.

One solution could be to supply the maximal number of arguments and then specify verb-specific selectional preferences. However, this issue cannot be entirely solved by simply increasing the valence of all co-mapped items to that of the largest structure as numbered arguments may be present in both rolesets but used differently. For example, issue.02 and problematic.01 "constituting a problem" (mapped to problem (Q621184) "situation that invites resolution") both have an ARG1. For issue.02, the ARG1 is an experiencer. For problematic.01, the ARG1 is a further description of the nature of the problem and the ARG2 is used for the experiencer. This requires encoding a more detailed set of specifications for the mapping between the rolesets and the Item within Wikidata to ensure backwards compatibility.

4.4 Predicate-level versus Argument-level Applicability

To aid in the identification of likely sense matches, annotators were given a list of automatically identified candidate Items as a starting point for evaluation. For opine.01, this list included opinion (Q3962655) (see Table 5). "Opinion" is listed as a type of entity, precluding it from mapping to the predicate. This information is important as full compatibility between projects will require mapping coverage at the argument-level in addition to predicate-level coverage.

Thus, we have begun this effort by mapping Items such as "opinion" at the argument level (in this case, to the ARG1 of opine.01). Identification of argument-level versus predicate-level applicability is critical, as failure to do so would introduce annotation circularity. Thus, not only have we increased the number of Items evaluated as potential mappings to a roleset, we have begun argument-level annotation efforts that can be used for richer capabilities of inference in information extraction tasks, argument extraction tasks, and future annotation refinement.

4.5 Eventiveness

In some cases, Items that map at the argument level are non-eventive. With the exception of noting Items that map to an argument of a predicate, this research has focused on events, as non-events cannot be mapped at the predicate level. Research collaborators sharing curated data have a similar focus. As such, we removed non-eventive Items from the shared DWD overlay so that they would not be considered as a potential predicate mapping. It should be noted that we take a conservative approach to the removal of non-eventive Items. For example, Items such as bathtub (Q152095) can safely be removed. However, alliance (Q878249) "coalition made between two or more parties to secure common interests" was not removed, despite the fact that its statements indicate it represents organizations.

4.6 Constraints

Another useful aspect of argument mapping is the ability to add in automatic, more fine-grained sense disambiguation, particularly when training data is limited. This can be seen in the case of contain.02 "restrain, halt the spread of", which has an ARG0: entity restraining ARG1 and an ARG1: thing being restrained, halted. Contain.02 is used for all types of restraint, including things such as diseases, chemicals as in the case of an oil spill, non-toxic items such as water, or more abstract items such as news or information. Because PropBank does not differentiate between these subtypes of containment, mapping at the predicate level can supply an argument structure for a general restricting event but cannot further enrich the event description.

However, the selectional preferences discussed in Section 3.2 can be added to the arguments. Selectional preferences specify the kind of information commonly associated with our containment subtypes (e.g., disease, chemicals, information, etc.). For example, see the usage of the predicate "contain" in Sentence 1:

Opine.01 argument structure	Candidate Item
ARG0-PAG: speaker ARG1-PPT: opinion ARG2-GOL: hearer	Opinion (Q3962655) "judgment, viewpoint, or statement that is not conclusive; may deal with subjective matters in which there is no conclusive finding"

Table 5: Opine.01 and Opinion *Q3962655*

(1) Public health officials are working to **contain** the spread of COVID-19, in part by using contact tracing.

We know from the predicate that this is interpreted as a general "halt the spread of" event. To this, we add a selectional preference to the ARG1, indicating that it should be an illness or malady. Such information would stand in contrast to the selectional preferences associated with chemicals, or news. For chemicals, we add a selectional preference of "geographic area" to the ARG0 and "non-infectious contaminant" to the ARG1. For news, we add the selectional preference of "information, communicated item, statement, or belief" to the ARG1, which does not preclude selection of misinformation.

Crucially, selectional preferences are intended as optional tools. They do not restrict the kinds of information that can fulfill a numbered argument role. Instead, they attempt to characterize frequent scenarios to aid in more fine-grained sense disambiguation.

5 Improving DWD Utility

The overlay's development has been closely guided by feedback from users of the overlay. Currently, the overlay is used for event extraction on newstext, so many of the changes and additions made to the overlay have been informed by the effect the change might have on that task.

5.1 Sentence Templates

To improve the clarity of the given roles for an event and to supply additional means for matching text instances of events and their arguments to a DWD event type, we added sentence-like templates to every event type. These templates place every PropBank argument role assigned to an event in an English language sentence with the event type expressed as the main verb for the sentence. For example, the roles for *combust.01* (see Table 1), mapped to Q133235, are incorporated into the following sentence template:

(2) {A0_pag_agent/cause_of_combustion} combusted {A1_ppt_thing_combusting} with {A2_mnr_instrument}.

These templates were created automatically using simple heuristics based on typical syntactic realizations of combinations of PropBank roles. For those Wikidata items that were easily converted to English verbs, this method worked fairly well, although some hand-correction was needed. However, for event types best expressed in English with nouns, such as coup d'etat (Q45383) or earthquake (Q7944), the sentence template needed additional manual curation. In many cases, a light verb or other multiword construction expressed the concept in a way that best allowed the incorporation of all roles into a sentence template. For the coup d'etat type, for example, we used

(3) {A0_pag_agent_coup_stager} staged a coup against {A1_ppt_theme_overthrown_government} at {AM_loc}.

5.2 Aligning with GLEN

The GLEN event-detection dataset was created using 3,465 event types drawn from the DWD ontology (Li et al., 2023). The dataset takes advantage of the DWD mappings between the Wikidata event Items and PropBank rolesets, which allowed them to build upon the extensive existing PropBank annotation. The resulting dataset, with 205K event mentions, has the broadest event-type coverage, 20x larger than the MAVEN dataset, which is based on 168 event types.

Li et al. (2023) selected a subset of DWD event types for GLEN, eliminating cognitive types like BELIEF, very fine-grained types like STAPEDECTOMY, and very low-frequency types like SINTERING. We are partially aligning the DWD with GLEN by eliminating the same fine-grained and low-frequency types. However, we are keeping the cognitive types for those who might need an expanded set of general-purpose types.

5.3 Concept Hierarchy

While Wikidata is not an ontology, some of the properties that describe relations between items can be used to construct domain-specific ontologies. The most obvious of these are "subclass of" (P279) and "instance of" (P31). There are, however, issues with this in practice. For example, no "subclass of" chain relates tiger (Q19939) to animal (Q729). Tigers are instead "instances of" taxon (Q16521) that are related through "parent taxon" (P171) and the concept "animal" is not in the "parent taxon" taxonomy. Another challenge is the length of the "subclass of" chain and the intermediate concepts it includes. These can be esoteric and not necessarily helpful for drawing the types of inferences NLP systems rely on. For the DWD overlay, an ontology with multiple inheritance using a "parent" relation was created. A parent of concept A is concept B in DWD such that there is a Wikidata chain of concepts $< A, c_0, ...c_k, B >$, where the intermediate concepts c_i are not in DWD and are connected via either "subclass of" or "parent taxon" properties. The last connection in the chain can also be "instance of".

6 Conclusion

The project of merging PropBank into Wikidata presents unique annotation challenges. Wikidata is highly entity-centric, making it difficult to identify good candidates for mapping to PropBank: moreover, differing granularity between verb senses in PropBank and concepts in Wikidata has proved to be one of the foremost challenges of mapping between resources, even when an eventive Wikidata item exists. Our approach to managing many of these challenges has been to prioritize feedback from the users of the overlay, allowing the needs of the information extraction use case to guide us.

On top of the linguistic challenges, organizational and logistical challenges abound. PropBank is a relatively static resource that was annotated by linguistic experts. Wikidata is a constantly changing project maintained and compiled by anyone with an internet connection. While the overlay has a well-defined use case—event and argument extraction from natural language text—Wikidata has so many different users and stakeholders that it is difficult to claim that there is any one use case for the resource. Our current approaches for managing the challenges of PropBank-Wikidata linking, therefore, may need to be adjusted as we work with the Wikidata community to integrate our mappings into Wikidata itself.

Acknowledgements

We gratefully acknowledge the support of DARPA FA8750-18-2-0016-AIDA – RAMFIS: Representations of vectors and Abstract Meanings for Information Synthesis and a sub award from RPI on DARPA KAIROS Program No. FA8750-19-2-1004. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA or the U.S. government. We especially want to thank Zixuan Zhang for his contributions to Section 2.3.

References

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet Project. In *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1*, pages 86–90, Montreal, Quebec, Canada. Association for Computational Linguistics.

Xinya Du, Zixuan Zhang, Sha Li, Pengfei Yu, Hongwei Wang, Tuan Lai, Xudong Lin, Ziqi Wang, Iris Liu, Ben Zhou, Haoyang Wen, Manling Li, Darryl Hannan, Jie Lei, Hyounghun Kim, Rotem Dror, Haoyu Wang, Michael Regan, Qi Zeng, Qing Lyu, Charles Yu, Carl Edwards, Xiaomeng Jin, Yizhu Jiao, Ghazaleh Kazeminejad, Zhenhailong Wang, Chris Callison-Burch, Mohit Bansal, Carl Vondrick, Jiawei Han, Dan Roth, Shih-Fu Chang, Martha Palmer, and Heng Ji. 2022. RESIN-11: Schema-guided Event Prediction for 11 Newsworthy Scenarios. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations*, pages 54–63, Hybrid: Seattle, Washington + Online. Association for Computational Linguistics.

Daniel Gildea and Martha Palmer. 2002. The necessity of parsing for predicate argument recognition. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, page 239–246, USA. Association for Computational Linguistics.

Paul Kingsbury and Martha Palmer. 2002. From TreeBank to PropBank. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02)*, Las Palmas, Canary Islands - Spain. European Language Resources Association (ELRA).

Sha Li, Heng Ji, and Jiawei Han. 2021. Document-level event argument extraction by conditional generation. In *Proceedings of the 2021 Conference of the North*

American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 894–908, Online. Association for Computational Linguistics.

Sha Li, Qiusi Zhan, Kathryn Conger, Martha Palmer, Heng Ji, and Jiawei Han. 2023. GLEN: General-purpose event detection for thousands of types. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2823–2838, Singapore. Association for Computational Linguistics.

Kevin Meng, Arnab Sen Sharma, Alex J Andonian, Yonatan Belinkov, and David Bau. 2023. Mass-editing memory in a transformer. In *The Eleventh International Conference on Learning Representations*.

George A. Miller. 1995. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41.

Mahir Morshed. 2023. Using wikidata lexemes and items to generate text from abstract representations. *Semantic Web Journal p.(submitted)*.

Natasha Noy, Yuqing Gao, Anshu Jain, Anant Narayanan, Alan Patterson, and Jamie Taylor. 2019. Industry-scale knowledge graphs: Lessons and challenges. *Commun. ACM*, 62(8):36–43.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Comput. Linguist.*, 31(1):71–106.

Fabio Petroni, Tim Rocktäschel, Sebastian Riedel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, and Alexander Miller. 2019. Language models as knowledge bases? In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 2463–2473, Hong Kong, China. Association for Computational Linguistics.

Alessandro Piscopo and Elena Simperl. 2019. What we talk about when we talk about wikidata quality: A literature survey. In *Proceedings of the 15th International Symposium on Open Collaboration*, OpenSym '19, New York, NY, USA. Association for Computing Machinery.

James Pustejovsky. 1995. *The Generative Lexicon*. MIT Press.

Karin Kipper Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D., University of Pennsylvania, United States – Pennsylvania. ISBN: 9780542200496.

Kartik Shenoy, Filip Ilievski, Daniel Garijo, Daniel Schwabe, and Pedro Szekely. 2022. A study of the quality of Wikidata. *Journal of Web Semantics*, 72:100679.

Elizabeth Spaulding, Kathryn Conger, Anatole Gershman, Rosario Uceda-Sosa, Susan Windisch Brown, James Pustejovsky, Peter Anick, and Martha Palmer. 2023. The DARPA Wikidata Overlay: Wikidata as an ontology for natural language processing. In *Proceedings of the 19th Joint ACL - ISO Workshop on Interoperable Semantic Annotation (ISA-19)*, pages 1–10, Nancy, France.

Ann Taylor, Mitchell Marcus, and Beatrice Santorini. 2003. The Penn Treebank: An Overview. In Anne Abeillé, editor, *Treebanks: Building and Using Parsed Corpora*, Text, Speech and Language Technology, pages 5–22. Springer Netherlands, Dordrecht.

Denny Vrandečić. 2020. Architecture for a multilingual wikipedia. Technical report, Google.

Denny Vrandečić. 2021. Building a multilingual wikipedia. *Commun. ACM*, 64(4):38–41.

Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, 57(10):78–85.

Xingyao Wang, Sha Li, and Heng Ji. 2023. Code4Struct: Code generation for few-shot event structure prediction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3640–3663, Toronto, Canada. Association for Computational Linguistics.