

Victim or Assailant? Exploring Narratives Through Knowledge Graph Queries

Beatrice Fiumanò^{*1}, Nicolas Lazzari^{*}, Simone P. Ponzetto[†], Valentina Presutti^{*}

University of Bologna, Italy^{*}, University of Mannheim, Germany[†]

{beatrice.fiumano,nicolas.lazzari3,valentina.presutti}@unibo.it, ponzetto@uni-mannheim.de

Corresponding author¹

Abstract

Our understanding of social reality is shaped by the specific ways in which that reality is framed by different sources. Analyzing framing means examining how these sources are able to convey particular worldviews by foregrounding or downplaying certain aspects of experience. Current computational approaches address this task by automatically identifying communicative patterns (e.g., topic selection or rhetorical strategies) that characterize individual artifacts. However, they often remain document-bound, overlooking the comparative dimension that enables the uncovering of convergent or conflicting narratives about the same actor, event, or issue. In this paper, we propose DORIS, an ontology that supports both document-level and cross-document framing analysis using SPARQL queries on automatically constructed Knowledge Graphs. We validate the proposed approach through a case study of historical news articles, exploring multiple framings of a real-world event using Fillmore’s Frame Semantics and the FrameNet resource. Code and data are available on GitHub at <https://github.com/beatrice-f/DORIS/>.

Keywords: Knowledge Graph, Ontology, Framing Analysis, Frame Semantics

1. Introduction

Decades of research across multiple disciplines, from linguistics to the social sciences, have highlighted that communication is never neutral, neither in the production nor in the interpretation of meaning (Hall, 1973; Foucault, 2013).

Any communicative artifact promotes a specific worldview that embeds social and political structures, personal and collective biases, ideology, and stance (Fairclough, 1995). This constitutes an act of *framing*, understood as the ad-hoc construction of meaning by selectively foregrounding or silencing certain aspects of experience (Entman, 1993). As a result, framing shapes how reality is perceived.

Consider, for instance, these two excerpts from different news sources covering the anti-Nazi protest that took place in New York on 20 February 1939¹, extracted from the NewsWire (NW) dataset (Silcock et al., 2024):

“Outside Madison Square Garden policemen had a six-hour struggle with throngs of anti-Nazis who repeatedly charged their lines trying to fight their way inside.”

NW5635

“..., a moving throng of anti-Nazis, theatergoers and the merely curious milled about in the streets. About 1,500 police reserves stood guard over the area, while violence spurted up inside the Garden and out.”

NW523

Excerpt from NW5635 emphasizes the physical confrontation between police officers and anti-Nazi protesters, who are presented as a disruptive force. In NW523, the emergence of violence does not have a clear agent: Protesters are portrayed as part of a bigger, heterogeneous crowd, whereas police officers are represented as a surveilling presence. We observe here that communication is, in itself, an act of perspectivization or interpretation of reality.

To analyze framing dynamics, scholars have adopted a variety of methodological approaches. Traditionally, research in discourse and narrative analysis has relied on qualitative, human-driven investigations (Entman, 1993; Fairclough, 1995; Wodak, 2001).

These approaches guarantee high-quality, nuanced insights, but they are both time- and effort-intensive, and susceptible to researcher bias (Parks and Peters, 2023). By contrast, computational approaches ensure reproducibility and scalability, enabling large-scale analyses that would otherwise be impractical (Hamborg, 2023; Parks and Peters, 2023).

A number of approaches have been proposed in the NLP field, allowing scholars to systematically analyze framing (Otmakhova et al., 2024). Nonetheless, they are generally bound to a single document, hampering a broader understanding of framing across multiple documents (Otmakhova et al., 2024; Ali and Hassan, 2022).

In this work, we propose overcoming this limitation by representing artifacts, their metadata, and automatically derived observations in a structured way using Knowledge Graphs (KGs) (Hogan et al., 2021). Our approach is based on the *encoding*-

¹https://en.wikipedia.org/wiki/1939_Nazi_rally_at_Madison_Square_Garden

decoding model (Hall, 1973), where a real-world situation (e.g., a protest) undergoes an *encoding* process that produces a semiotic artifact (e.g., a news article) that is finally *decoded* by a cognizer (e.g., a reader of the news article).

We propose organizing knowledge in the KG using a novel ontology, DORIS, based on the established Description and Situation Ontology Design Pattern (Gangemi and Mika, 2003; Gangemi and Presutti, 2009). DORIS allows dynamics of framing to be modeled within a broader relational system that supports cross-artifact and cross-language investigations and higher-level pattern detection. By leveraging KGs, semiotic items are linked through disambiguated actors, events, and topics, enabling systematic comparison across sources, languages, and contexts, and the exploration of how events, actors, and issues are constructed across artifacts.

We demonstrate this approach by leveraging Fillmore’s Frame Semantics (FS) (Fillmore, 1976), operationalized in FrameNet (Baker et al., 1998), as the guiding framework for analyzing discourse and semiotic representations, since it enables exploration of the discursive conceptualization of events and participants (Ziem, 2014). For example, FS shows that excerpt *NW5635* interprets the circumstance as a `HOSTILE_ENCOUNTER` between a `Side_1` (police officers) and a `Side_2` (anti-Nazi protesters), while text *NW523* employs a metaphor that casts violence as a `Fluid` within the `FLUIDIC_MOTION` frame, hence providing scholars with insights that “go beyond just words” (Ali and Hassan, 2022).

Nonetheless, our approach is not bound to a specific theory and can be adapted to different operational frameworks, such as the Moral Foundation Theory (Graham et al., 2013) or the Narrative Policy Framework (Shanahan et al., 2018), as well as a combination of them.

In summary, our contribution is threefold:

- C1)** we present DORIS (**D**omain-specific **O**ntology for **R**epresenting and **I**nterpreting **S**emiotic artifacts), a novel ontology modeling the relationship between a semiotic artifact, the real-world situation it represents, and the interpretation of that situation conveyed in the artifact;
- C2)** we show that DORIS supports the exploration of semiotic representations and framing strategies across different artifacts through the use of SPARQL queries, lowering the barrier to the use of complex computational methods;
- C3)** we present a use case that illustrates ontology-grounded KG construction and exploration, leveraging FS at the interpretative layer, using historical news articles.

The rest of the paper is organized as follows: in

Section 2, we provide an overview of the main approaches to framing analysis in NLP. In section 3, we present DORIS and describe its main modeling choices, which directly influence the construction of the KG and how it is queried. We demonstrate the feasibility of our approach through a case study based on historical documents, presented in Section 4. Finally, we summarize our contributions and highlight limitations and future extensions in Section 5.

2. Related Work

As introduced in Section 1, computer-driven techniques for framing analysis have gained traction in the past few decades, particularly within the NLP and computational social science fields. In this section, we provide an overview of the main computational approaches to framing analysis, specifically focusing on those that, like our proposal, rely on graph-based methods and FS.

Computational Approaches As noted by Ot-makhova et al. (2024), computational approaches to the study of framing dynamics are heterogeneous, varying across disciplines and the focus of analysis. In the NLP community, most research efforts focus on emphasis framing, i.e., issue dimensions explored through topic modeling or predefined codebooks (Ali and Hassan, 2022). Although popular, this line of research tends to simplify frames as topics. Another line of research adopts a more fine-grained approach, focusing on lexical (Wicke and Bolognesi, 2025), syntactic, and discursive choices (Reinig et al., 2024). Among these, several studies have relied on (frame-)semantic parsing to explore semantic relations in text. Adopting an approach similar to the one proposed in this paper, both Postma et al. (2020) and Minnema et al. (2022) use FrameNet frames to analyze variations in the framing of events and social issues.

Furthermore, Postma et al. (2020) underscore the importance of anchoring narratives to real-world entities and events, and of resolving them across documents. In line with this intuition, some studies have shown that a more effective approach for inter-document exploration of narratives is to structure them as graphs.

Graph-based Approaches Following this argument, Baden (2018) proposes to model news discourse as a network of entities and applicability relations among them, enabling the exploration of entity-centric framing patterns across news items. Similarly, Pournaki and Willaert (2025) rely on semantic dependencies to extract narrative signals from a corpus of political speeches and represent

them as graphs. As in our approach, framing can be analyzed through graph queries. Focusing on event-centric representations, [Rovera et al. \(2021\)](#) employ a graph-based model that encodes both surface-level information (e.g., actors, locations, and explicit relations) and the deeper semantic argument structure of events. [Motta et al. \(2025\)](#) propose a formal, ontology-based typology for news classification that captures the topic or issues addressed in an article and the issue-specific claims and viewpoints it conveys, supporting fine-grained content analyses.

Our approach follows this stream of research with a specific focus on adopting LOD, as we will show in [Section 4.4](#).

3. The DORIS Ontology

The main objective of this work is to demonstrate how KGs can support query-driven narrative exploration and extract meaningful insights into the framing strategies encoded within a given artifact. To achieve this, knowledge must be organized in the KG to support answering relevant research questions, such as “How is the police framed in the document *NW5635*?”, and “How do role attribution and agency of anti-Nazi protesters vary across documents?”

In this section, we describe DORIS, a novel ontology designed to support answering similar research questions. As introduced in [Section 1](#), DORIS draws on Hall’s *encoding/ decoding* theory ([Hall, 1973](#)). This theory conceptualizes communication as a process in which a state of affairs (e.g., a protest) is encoded into a semiotic artifact (e.g., a news article) by an agent operating within a specific socio-cultural and ideological context. This encoding process involves selecting a preferred reading or interpretation of events, which results in framing by means of specific lexico-syntactic and narrative choices ([Entman, 1993](#)). The encoded artifact is subsequently decoded by some agent, whose interpretation is also mediated by their own beliefs, values, and background.

DORIS, shown in [Figure 1](#), models the encoding-decoding process by reusing the Description And Situation ([Gangemi and Mika, 2003](#), DnS) Ontology Design Pattern ([Gangemi and Presutti, 2009](#), ODP) and the DOLCE ontology ([Borgo et al., 2022](#)).

Conceptually, the ontology is organized into four layers modeling different steps of the process: the *grounding layer*, the *encoding layer*, the *decoding layer*, and the *theory layer*.

The *grounding layer* models the fact that a `DUL:ENTITY` (e.g., a protest) is encoded into multiple artifacts, collected in an `:ENCODINGCOLLECTION` (e.g., a set of news articles), all realized through the same `:MODE` (e.g., natural language). Note

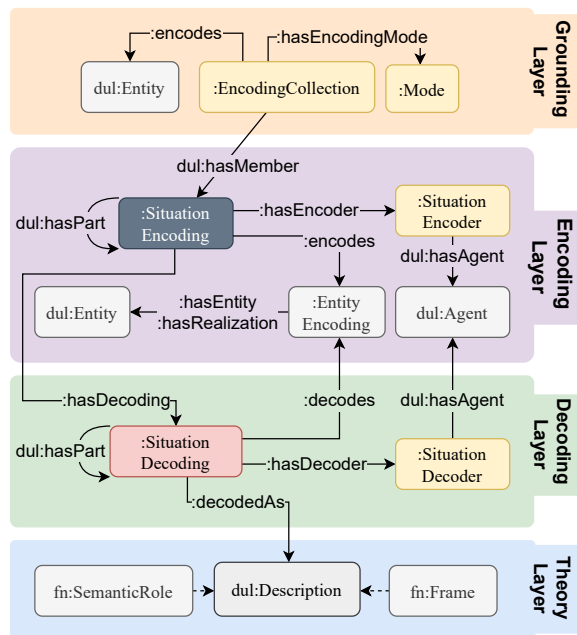


Figure 1: The DORIS ontology represented using [Graffoo](#). Grey boxes refer to reused ontology classes.

that due to the generality of `DUL:ENTITY` instances, DORIS allows modeling the framing of concrete entities (e.g., a politician) as well as abstract concepts (e.g., gender equality) or social entities (e.g., the government).

Each member of an `:ENCODINGCOLLECTION` is a semiotic artifact (e.g., a news item or a painting), formalized by the `:SITUATIONENCODING` class in the *encoding layer*. A `:SITUATIONENCODING` instance is a DnS Situation. Informally, a DnS Situation represents a relation over multiple entities, where each entity takes a specific role.

In a `:SITUATIONENCODING`, these entities are: (i) the agent (or possibly a group of agents) that produced the artifact, represented by the `:SITUATIONENCODER` class, (ii) the `:ENTITYENCODING`, which is itself a DnS Situation involving the entity encoded and how it has been realized in the artifact (e.g., its textual mentions) and (iii) the decoding(s) of an artifact, represented by the `:SITUATIONDECODING` class. Finally, a `:SITUATIONENCODING` might be decomposed into smaller parts. For instance, in the case of a news article, each of its paragraphs is again `:SITUATIONENCODING` instances.

The decoding output is represented by the `:SITUATIONDECODING` class, which is also a DnS Situation. Similar to a `:SITUATIONENCODING`, a `:SITUATIONDECODING` instance involves the agent performing the decoding and the entity being decoded. Unlike a `:SITUATIONENCODING`, a `:SITUATIONDECODING` instance involves a DnS Description that is used to decode it. A DnS Description instance represents a class of situations that share some aspect between

them (e.g., the set of all individuals decoded as victims).

The DnS Descriptions and the relations holding between them are formalized in the *theory layer*. In Figure 1, we show a possible instantiation of a theory layer using FS as formalized in Framester (Gangemi et al., 2016) (as further illustrated in Section 4). More generally, this layer is theory-agnostic, and any suitable theoretical framework can be adopted, provided that it is formalized as an ontology that re-uses the DnS ODP. In Section 4, we demonstrate that this approach enables the integration of multiple theories, yielding deeper insights into the framing of an entity. Finally, a :SITUATION-DECODING can also be decomposed into smaller parts, similarly to a :SITUATIONENCODING instance.

Figure 2 shows an example of how DORIS is used to represent the encoding and decoding of the social entity `police` in excerpts *NW5635* and *NW523* from Section 1. Although the two documents use different surface forms (“policemen” vs “police”), they both refer to the same conceptual entity (e.g., a general police entity represented by node Q35535 on Wikidata). DORIS allows comparing how the same entity is decoded across documents by explicitly decoupling a superficial mention from the actual entity it describes, overcoming the limitation of analyses bound to a single article. Figure 2 also shows how the choice of a theory layer does not influence the analysis of an artifact.

4. Case Study: Analyzing Media Framing in Historical Newspapers

For our use case, we focus on entity framing (Mahmoud et al., 2025), and in particular on how media sources represent an event and its participant entities. As previously illustrated, highlighting certain aspects of an event, such as confrontations between police officers and protesters, while downplaying or omitting others, such as the actors responsible for violence escalation, promotes different worldviews. Consequently, entities may acquire different connotations depending on the actions they are described as performing or undergoing. Within this use case, our objective is to examine how the constituent happenings of an event are portrayed across news articles and, based on these portrayals, how participant entities are assigned different semantic roles.

In this section, we outline a methodology for constructing a KG, structured according to DORIS, from a news article. Nonetheless, it is possible to apply a similar pipeline to other modalities beyond text, as also discussed in Section 5.

4.1. KG Construction

Constructing a KG from text has been widely explored as part of the Knowledge Extraction subfield of NLP, where one is interested in identifying a set of entities and the relations that hold between them starting from a natural language sentence (Zhong et al., 2024).

Given a set of documents describing the same event (e.g., excerpts *NW5635* and *NW523*), we rely on the set of named entities annotated in the documents and perform document-level coreference resolution, by relying on ChatGPT 5.1² and manually refining predictions. This step normalizes all textual mentions to refer to the same entity, ensuring consistency in the KG. Despite manual effort, we highlight that this approach can be automated using recent advances in NLP, e.g., Martinelli et al. (2025).

We then identify semantic relations between entities in the documents using the LOME frame semantic parser (Xia et al., 2021), which has been trained on FrameNet. Finally, we convert LOME’s predictions to a KG in RDF representation compliant with DORIS. To perform this step, we use SPARQL Anything (Asprino et al., 2023), which enables KG construction from heterogeneous document types. The result of this step produces a KG that follows the structure of the ones shown in Figure 2.

4.2. Dataset

In order to demonstrate how one can query the KG to analyze the framing in an article, we base our case study on a selection of five news pieces extracted from the NewsWire dataset³ (Silcock et al., 2024), which contains 2.7 million historical news wire articles in English. Since we are interested in identifying texts covering the same historical event, we filter the dataset by year and topic, selecting articles published in 1939 and classified under the topic PROTEST. Among this subset, we select five articles reporting on the Nazi rally held in New York on February 20, all issued by U.S. news wire services. Excerpts from *NW5635* and *NW523* reported in Section 1 have been extrapolated from these articles.

4.3. Qualitative Analysis

Before delving into the query-based framing analysis, we report a qualitative analysis conducted by one of the authors with a background in linguistics. This analysis serves a dual purpose. On one hand, it provides a reference for evaluating whether query

²<https://openai.com/gpt-5/>

³<https://huggingface.co/datasets/dell-research-harvard/newswire>

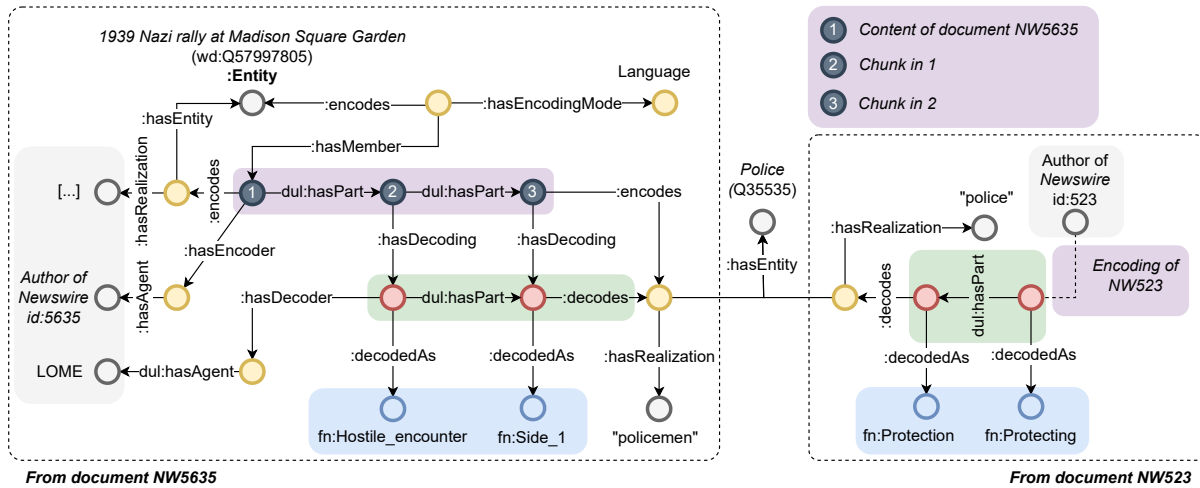


Figure 2: Example usage of DORIS to represent the decoding of the entity *police* in the excerpts *NW5635* and *NW523*. The block on the left shows how the metadata of the article is encoded in a KG (grey box) as well as how the event is encoded (purple box) in a news article (node 1), which contains a sentence from it (node 2) mentioning the entity *police* as “policemen” (node 3). Using FS, the sentence evokes the frame *HOSTILE_ENCOUNTER* with *police* taking the role of *SIDE_1* (i.e., one of the forces of the hostile encounter). Similarly, the block on the right shows that a sentence mentioning “*police*” evokes the frame *PROTECTING*, with *police* taking the role of *PROTECTION* (i.e., the entity that prevents harm to another entity). For the sake of compactness, we omit metadata and encoding/decoding instances related to this node.

results yield meaningful insights for framing analysis. On the other hand, it allows to assess whether the knowledge derived from the KG is consistent with human observations⁴.

In the following paragraphs, we examine the semiotic portrayal of the main actors involved in the news coverage of the rally, including the German American Bund (GAB), its paramilitary forces, and its leader, Fritz Kuhn; U.S. police; and U.S. citizen Isadore Greenbaum. Specifically, each article reports on two key episodes of the rally: the confrontations between police officers and counter-protesters, discussed above, and the episode involving Kuhn and Greenbaum, who leapt on the stage while the former was delivering an antisemitic speech. For each article, we analyze the stance adopted toward the actors and the discursive construction of both episodes.

NW523 This article attributes different evaluative weights to the key social actors. The Bund is construed as an ideologically marked and conflict-generating presence. Ideological non-alignment is encoded, for instance, in the structure *Bund* → *denounces* → *American alliance with European democracies*, where “European democracies” func-

tions as a positively loaded term whose alliance is, however, positioned as the target of denunciation. Storm troopers are framed as a source of violence in the Kuhn-Greenbaum episode, in which they are described as “setting upon” and “knocking down” Greenbaum. Conversely, the police are portrayed as a rescuing force that intervenes to remove Greenbaum from the troopers’ attack. Both the anti-Nazi protesters and the crowd are framed through limited agency, as violence is described as “spurting up” and the outside crowd is grouped into a heterogeneous group of people who “mill about”.

NW3763 A similar framing of storm troopers and police forces emerges in *NW3763*. Here, Greenbaum is explicitly construed as a victim of the trooper’s violence, saved by police intervention. While his actions are presented with little evaluative loading (he “advanced toward” Kuhn), Kuhn is framed as an antagonist through the depiction of his speech as a bitter attack on the Jews. Although no anti-Nazi group is explicitly mentioned, the outside crowd is here portrayed as a source of public disorder, characterized as quarrelsome and as disrupting the order maintained by the police.

NW5635 In this item, greater focus is given to the order-maintaining role of the police, as illustrated in Section 1, and on the disruptive conduct of counter-protesters. The GAB is again ideologically marked through a detailed account of the rally cere-

⁴We note that, given the limited scope of the study and the potential influence of annotator bias, this analysis is intended solely as a proof of concept. Further research is required for a systematic evaluation, as discussed in Section 5.

mony, which included “denunciations of Jews” and “salutes to swastikas”. The Greenbaum episode is given limited space and is explicitly framed as an “attempt to attack” the Bund leader, with omission of the ongoing antisemitic speech. A narrative reorientation occurs only at the last sentence of the article, where Bund members are construed as patients of an `attack` event with unspecified agency.

NW6072 As in the previous item, the core of this article lies in the positive construal of the police, here depicted as “a wall of invulnerable blue”. With limited attention devoted to the rally itself, the police role is emphasized by narrative minimization of conflict (“general rioting had not developed”, “inside the Garden itself only one fight occurred”). In contrast, the crowd is positioned as an antagonistic force, being politically identified as anti-Fascist and depicted as actively attempting to breach a “forbidden area”.

NW19221 In this article, police forces are again foregrounded as an order-containing force (“a solid ring of shoulder-to-shoulder foot and mounted police”) against antagonist anti-Nazi protesters. The GAB is strongly ideologically marked through explicit labeling of the rally as “anti-Jewish” and “anti-communist”. The Greenbaum episode is narrated in detail and framed as the central moment of violence. While his advance toward Kuhn is described with limited evaluative loading, storm troopers are assigned explicit and graphic violent agency (“felled”, “seized”, “hurled”), positioning Greenbaum as the patient of violence, who is later rescued by police. Greenbaum is subsequently also framed through institutional culpability, as he is “arraigned” and charged with disorderly conduct. At the end, his request for a doctor after the arraignment reintroduces vulnerability.

4.4. Framing Analysis by Querying a Knowledge Graph

Although we rely on a limited number of articles, the analysis in the previous section highlights several differences in how the same event is framed by different authors (in our setting, different newswires).

In this section, we demonstrate how SPARQL queries over a DORIS-based KG can support the exploration of relevant research questions in discourse and framing analysis. For each query, we report a representative subset of results, selected based on their relevance to the qualitative analysis conducted in Section 4.3⁵. However, we note

⁵We report the complete set of query outputs in the GitHub repository, including less relevant and noisy results

that, due to the number of less relevant or incorrect frames detected by the frame semantic parser, a considerable amount of noise emerges from each query. We further discuss these limitations in Section 5.

In this section, we focus on three main operations that support framing analysis: document-level analysis, cross-document analysis, and Linked Open Data (LOD)-enriched browsing. For each scenario, we present one or more queries related to our use case, and discuss the results in relation to the human-driven analysis outlined in Section 4.3.

Document Level: How is entity X framed in a given document? We examine here the framing of the social actor Isadore Greenbaum within the news item *NW3763* using the following Query 1:

```
SELECT ?role WHERE {
  ?articleSituation a :SituationEncoding;
    :encodes [ :hasRealization [
      provo:wasGeneratedBy <NW3763>
    ] ];
    :hasDecoding ?articleDecoding .
  ?articleDecoding dul:hasPart [
    :decodes [
      :hasEntity <Isadore Greenbaum>
    ];
    :decodedAs ?role] .
}
```

Query 1: What are the roles assigned to Isadore Greenbaum in document NW3763?

Semantic Frame	Role
Experience_bodily_harm	Experiencer
Cause_impact	Impactee
Cause_harm	Victim
Attack	Assailant

Table 1: Roles associated with Isadore Greenbaum in document NW3763 in the context of a semantic frame. Wrong parser predictions are highlighted in gray.

A relevant subset of the results, consistent with human observations, is reported in Table 1. The VICTIM role in the CAUSE_HARM frame, the EXPERIENCER role in the EXPERIENCE_BODILY_HARM frame, and the IMPACTEE role in the CAUSE_IMPACT frame align with the fact that the article largely construes Greenbaum as a victim. Nonetheless, the ASSAILANT role is incorrectly assigned to Greenbaum instead of Kuhn in the sentence “Kuhn had been bitterly attacking the Jews”, due to an erroneous prediction of LOME.

Cross-Document Level: What is the framing of Entity X across different documents? Query 2

investigates the construal of a given entity across all available documents. In this paragraph, we retrieve information relevant to analyzing the framing of the police and stormtrooper entities.

```
SELECT ?source ?description WHERE {
  ?articleSituation a :SituationEncoding;
  :encodes [ :hasRealization [
    provo:wasGeneratedBy ?source
  ] ];
  :hasDecoding ?articleDecoding .
  ?articleDecoding dul:hasPart [
    :decodes [
      :hasEntity <police/storm troops>
    ];
    :decodedAs ?description ] .
}
```

Query 2: What are the roles assigned to police across documents?

Article	Semantic Frame	Role
NW5635	Hostile_Encounter	Side_1
NW5635	Hostile_Encounter	Depictive
NW5635	Hostile_Encounter	Side_2
NW19221	Lvl_of_Force_Res.	Resisting_Entity
NW523	Protecting	Protection

Table 2: Roles associated with the police across documents. Wrong parser predictions are highlighted in gray.

Table 2 shows the results of Query 2 for the entity “police”. The framing emerging from the query results corroborates human observations. Across documents, the police are construed as active agents in physical confrontations, who oppose an antagonistic crowd for the maintenance of social order. This pattern is reflected in the police’s realization of the SIDE_1 role within the HOSTILE_ENCOUNTER frame and the RESISTING_ENTITY role, defined in FrameNet as an entity that “is capable of resisting or does resist a physical OPPOSING_FORCE”. It is further instantiated through the PROTECTOR role in the PROTECTING frame, where “Protection prevents a Danger from harming an Asset”.

Similarly, the query results for the storm troopers (Table 3) align with their qualitative portrayal as a violent actor. Across nearly all documents, they are instantiated in the AGENT role within the CAUSE_HARM and CAUSE_IMPACT frames, as well as in the ASSAILANT role in the ATTACK frame, consistently encoding them as a source of physical aggression.

Refinement and Expansions using External Knowledge Bases In this paragraph, we highlight some advantages of leveraging KGs to represent discourse and discourse dynamics, namely

Article	Semantic Frame	Role
NW3763	Cause_Harm	Agent
NW3763	Cause_Impact	Agent
NW5635	Cause_Harm	Agent
NW523	Cause_Impact	Agent
NW523	Attack	Assailant
NW523	Arrest	Charges
NW19221	Cause_Harm	Agent

Table 3: Roles associated with the storm troopers across documents. Wrong parser predictions are highlighted in gray.

the use of external Linked Open Data (LOD). The integration of LOD enables access to external knowledge bases (KB), which can be exploited to refine or expand an initial query with additional semantic information. In Query 3, we show how the Framester KB (Gangemi et al., 2016) can be used to retrieve only those instances in which Isadore Greenbaum instantiates a Frame Element of the ATTACK frame.

```
SELECT ?source ?role ?domain WHERE {
  ?articleSituation a :SituationEncoding;
  :encodes [ :hasRealization [
    provo:wasGeneratedBy ?source
  ] ];
  :hasDecoding ?articleDecoding .
  ?articleDecoding dul:hasPart [
    :decodes [
      :hasEntity <Isadore Greenbaum>
    ];
    :decodedAs ?role ] .
  SERVICE <Framester SPARQL endpoint> {
    ?role rdfs:domain fsframe:Attack .
  }
}
```

Query 3: In which documents does Isadore Greenbaum instantiate a role belonging to the ATTACK frame?

Article	Semantic Frame	Role
NW3763	Attack	Assailant
NW5635	Attack	Assailant
NW523	Attack	Victim

Table 4: Roles associated to Isadore Greenbaum and belonging to the ATTACK frame. Wrong parser predictions are highlighted in gray.

By filtering the query, we observe in Table 4 how Greenbaum instantiates both the ASSAILANT and the VICTIM roles, depending on the article. This variation reflects a difference in construal, whereby item NW5635 frames Greenbaum’s leap onto the stage as an “attempt to attack” Fritz Kuhn, while item NW523 foregrounds his patient role in the storm trooper’s assault.

Finally, in Query 4, we exploit supplementary knowledge available in Framester where FrameNet frames are aligned to MFT [Graham et al. \(2013\)](#), as formalized in the ValueNet ontology ([Giorgis et al., 2022](#)).

```

SELECT ?source ?value (COUNT(*) AS
  ?count) WHERE {
  ?articleSituation a :SituationEncoding;
  :encodes [ :hasRealization
    [ provo:wasGeneratedBy ?source
    ]];
  :hasDecoding ?articleDecoding .
  ?articleDecoding :decodedAs
    ?description .
  SERVICE <Framester SPARQL endpoint> {
    ?description vcvf:triggers ?value.
  }
  FILTER CONTAINS (STR(?value),
    "HaidtValues")
}
GROUP BY ?source ?value

```

Query 4: What is the number of occurrences of Moral Foundations in each document?

Article	Care	Oppression	Harm	Loyalty
NW523	3	2	1	1
NW5635		1	7	1
NW3763			6	
NW19221		1	2	1
NW6072				1

Table 5: Occurrences of Moral Foundations per article.

Overall, the distribution of MF (Table 5) endorses our qualitative analysis.

As many articles foreground physical confrontation, the *Harm* moral foundation predominates, typically activated by frames such as *АТТАК*. References to Greenbaum’s arrest also activate *Oppression*. By contrast, the order-maintaining role of the police evokes *Care*, which in article *NW523* is instantiated through the *PROTECTING* frame. The *Loyalty* value is recurrent across documents, mostly triggered by the *MEMBERSHIP* frame associated with Bund members and the *COME_TOGETHER* frame describing the rally crowd.

5. Limitations and Future Work

In this paper, we presented DORIS, an ontology based on Hall’s encoding-decoding theory that supports cross-document framing analysis by representing an artifact and its theory-grounded interpretation in a KG that can be easily queried via SPARQL. Using a case study on historical news articles, we have shown that the retrieved results align with human-derived findings. Section 4.4 shows

that, although we experimented with Fillmore’s FS as our theoretical lens, it is possible to expand the analysis to other relevant theories (e.g., the Moral Foundation Theory). This approach paves the way to the development of a human-in-the-loop paradigm that leverages different NLP tools to support framing analysis.

Nonetheless, our results also highlight a key limitation of the proposed approach, namely the frame semantic parser’s incorrect predictions. Further research is required to investigate the impact of such errors on the overall quality of the analysis, and to identify strategies for minimizing expert effort, for instance by automatically refining LOME’s outputs or by developing more accurate frame semantic parsers, especially for multilingual settings. Along similar lines, we plan to perform a more systematic evaluation of the devised method using established framing and discourse analysis benchmarks. Additional future work includes enhancing the KG construction pipeline described in Section 4.1 by integrating Entity Linking systems, which would greatly enhance the potential of this approach (akin to state-of-the-art machine readers such as Text2AMR2FRED ([Gangemi et al., 2026](#))), as well as using LLMs to elicit implicit knowledge, as illustrated in [De Giorgis et al. \(2025\)](#).

Finally, we highlight that although our case study focuses on the linguistic domain, DORIS has been designed to be medium- and domain-agnostic. Following a similar approach to the one in Section 4, it is possible to formalize the representation and interpretation process across different domains. For instance, [Ciroku et al. \(2024\)](#) relies on FS to analyze the implicit meaning of images by leveraging the relations among the entities they depict. By relying on DORIS, it is possible to harmonize such analyses across multiple images, allowing scholars to explore image collections through theoretically grounded interpretations.

Moreover, DORIS bridges the gap between large-scale *distant reading* techniques and more qualitative *close reading* approaches by preserving an explicit link between semiotic or discursive features and their analytical interpretations. Unlike many automated methods, the use of KGs allows scholars to derive more general, cross-artifact observations while retaining the possibility of progressively refining the analysis down to the surface (e.g., textual) level to examine the specific instantiations of a given phenomenon ([Rovera et al., 2021](#)).

6. Bibliographical References

Mohammad Ali and Naeemul Hassan. 2022. [A survey of computational framing analysis approaches](#). In *Proceedings of the 2022 Con-*

- ference on Empirical Methods in Natural Language Processing, pages 9335–9348, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
- Luigi Asprino, Enrico Daga, Aldo Gangemi, and Paul Mulholland. 2023. [Knowledge graph construction with a façade: A unified method to access heterogeneous data sources on the web](#). *ACM Transactions on Internet Technology*, 23(1):1–31.
- Christian Baden. 2018. Reconstructing frames from intertextual news discourse: A semantic network approach to news framing analysis. In *Doing news framing analysis ii*, pages 3–26. Routledge.
- Stefano Borgo, Roberta Ferrario, Aldo Gangemi, Nicola Guarino, Claudio Masolo, Daniele Porello, Emilio M. Sanfilippo, and Laure Vieu. 2022. [DOLCE: A descriptive ontology for linguistic and cognitive engineering](#). *Appl. Ontology*, 17(1):45–69.
- Fiorela Ciroku, Stefano De Giorgis, Aldo Gangemi, Delfina Sol Martinez Pandiani, and Valentina Presutti. 2024. [Automated multimodal sense-making: Ontology-based integration of linguistic frames and visual data](#). *Comput. Hum. Behav.*, 150:107997.
- Stefano De Giorgis, Aldo Gangemi, and Alessandro Russo. 2025. [Neurosymbolic graph enrichment for grounded world models](#). *Information Processing & Management*, 62(4):104127.
- Robert M. Entman. 1993. [Framing: Toward clarification of a fractured paradigm](#). *Journal of Communication*, 43(4):51–58.
- Norman Fairclough. 1995. *Critical Discourse Analysis: The Critical Study of Language*. Longman, London.
- Charles J. Fillmore. 1976. [Frame semantics and the nature of language](#). *Annals of the New York Academy of Sciences*, 280(1):20–32.
- Michel Foucault. 2013. *Archaeology of knowledge*. Routledge.
- Aldo Gangemi, Arianna Graciotti, Antonello Meloni, Andrea Giovanni Nuzzolese, Valentina Presutti, Diego Reforgiato Recupero, and Alessandro Russo. 2026. [Text2amr2fred, converting text into rdf/owl knowledge graphs via abstract meaning representation](#). *Knowledge and Information Systems*, 68(1):47.
- Aldo Gangemi and Peter Mika. 2003. [Understanding the semantic web through descriptions and situations](#). In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE - OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003*, volume 2888 of *Lecture Notes in Computer Science*, pages 689–706. Springer.
- Aldo Gangemi and Valentina Presutti. 2009. [Ontology design patterns](#). In Steffen Staab and Rudi Studer, editors, *Handbook on Ontologies*, International Handbooks on Information Systems, pages 221–243. Springer.
- Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto. 2013. [Moral foundations theory: The pragmatic validity of moral pluralism](#). In Patricia Devine and Ashby Plant, editors, *Advances in Experimental Social Psychology*, volume 47, pages 55–130. Academic Press.
- Stuart Hall. 1973. [Encoding and decoding in the television discourse](#). Discussion paper, AR-RAY(0x56177e4471c0), Birmingham. This paper has been published in: CCCS selected working papers. Vol. 2 / edited by Ann Gray et al (Abingdon, 2007) pp. 386-398.
- Felix Hamborg. 2023. *Revealing Media Bias in News Articles: NLP Techniques for Automated Frame Analysis*. Springer Nature, Cham, Switzerland.
- Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d’Amato, Gerard de Melo, Claudio Gutiérrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, Axel-Cyrille Ngonga Ngomo, Axel Polleres, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan F. Sequeda, Steffen Staab, and Antoine Zimmermann. 2021. [Knowledge Graphs](#). Number 22 in *Synthesis Lectures on Data, Semantics, and Knowledge*. Springer.
- Tarek Mahmoud, Zhuohan Xie, Dimitar Iliyanov Dimitrov, Nikolaos Nikolaidis, Purificação Silvano, Roman Yangarber, Shivam Sharma, Elisa Sartori, Nicolas Stefanovitch, Giovanni Da San Martino, Jakub Piskorski, and Preslav Nakov. 2025. [Entity framing and role portrayal in the news](#). In *Findings of the Association for Computational Linguistics, ACL 2025, Vienna, Austria, July 27 - August 1, 2025*, volume ACL 2025 of *Findings of ACL*, pages 302–326. Association for Computational Linguistics.
- Giuliano Martinelli, Bruno Gatti, and Roberto Navigli. 2025. [xcore: Cross-context coreference resolution](#). In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, EMNLP 2025, Suzhou, China, November*

- 4-9, 2025, pages 34264–34278. Association for Computational Linguistics.
- Gosse Minnema, Sara Gemelli, Chiara Zanchi, Tommaso Caselli, and Malvina Nissim. 2022. [SocioFillmore: A tool for discovering perspectives](#). In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 240–250, Dublin, Ireland. Association for Computational Linguistics.
- Enrico Motta, Enrico Daga, Aldo Gangemi, Maia Lunde Gjelsvik, Francesco Osborne, and Angelo A. Salatino. 2025. [The epistemology of fine-grained news classification](#). *Semantic Web*, 16(3).
- Yulia Otmakhova, Shima Khanehzar, and Lea Frermann. 2024. [Media framing: A typology and survey of computational approaches across disciplines](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15407–15428, Bangkok, Thailand. Association for Computational Linguistics.
- Louisa Parks and Wim Peters. 2023. [Natural language processing in mixed-methods text analysis: A workflow approach](#). *International Journal of Social Research Methodology: Theory & Practice*, 26(4):377–389.
- Marten Postma, Levi Remijnse, Filip Ilievski, Antske Fokkens, Sam Titarsolej, and Piek Vossen. 2020. [Combining conceptual and referential annotation to study variation in framing](#). In *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet*, pages 31–40, Marseille, France. European Language Resources Association.
- Armin Pournaki and Tom Willaert. 2025. [Extracting narrative signals from public discourse: a network-based approach](#). *Humanities and Social Sciences Communications*, 12(1):1774.
- Ines Reinig, Ines Rehbein, and Simone Paolo Ponzetto. 2024. [How to do politics with words: Investigating speech acts in parliamentary debates](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 8287–8300, Torino, Italia. ELRA and ICCL.
- Marco Rovera, Federico Nanni, and Simone Paolo Ponzetto. 2021. [Event-based access to historical italian war memoirs](#). *J. Comput. Cult. Herit.*, 14(1).
- Elizabeth A Shanahan, Michael D Jones, Mark K McBeth, and Claudio M Radaelli. 2018. The narrative policy framework. In *Theories of the policy process*, pages 173–213. Routledge.
- Philipp Wicke and Marianna M. Bolognesi. 2025. [Red and blue language: Word choices in the trump and harris 2024 presidential debate](#). *PLOS ONE*, 20(6):1–30.
- Ruth Wodak. 2001. The discourse-historical approach. In Ruth Wodak and Michael Meyer, editors, *Methods of Critical Discourse Analysis*, pages 63–94. Sage, London.
- Patrick Xia, Guanghui Qin, Siddharth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White, and Benjamin Van Durme. 2021. [LOME: Large ontology multilingual extraction](#). In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 149–159, Online. Association for Computational Linguistics.
- Lingfeng Zhong, Jia Wu, Qian Li, Hao Peng, and Xindong Wu. 2024. [A comprehensive survey on automatic knowledge graph construction](#). *ACM Comput. Surv.*, 56(4):94:1–94:62.
- Alexander Ziem. 2014. *Frames of Understanding in Text and Discourse: Theoretical Foundations and Descriptive Applications*. John Benjamins, Philadelphia.

7. Language Resource References

- Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The berkeley framenet project. In *COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics*.
- Aldo Gangemi, Mehwish Alam, Luigi Asprino, Valentina Presutti, and Diego Reforgiato Recupero. 2016. [Framester: A wide coverage linguistic linked data hub](#). In *European knowledge acquisition workshop*, pages 239–254. Springer.
- Stefano De Giorgis, Aldo Gangemi, and Rossana Damiano. 2022. [Basic human values and moral foundations theory in valuenet ontology](#). In *Knowledge Engineering and Knowledge Management - 23rd International Conference, EKAW 2022, Bolzano, Italy, September 26-29, 2022, Proceedings*, volume 13514 of *Lecture Notes in Computer Science*, pages 3–18. Springer.
- Emily Silcock, Abhishek Arora, Luca D’Amico-Wong, and Melissa Dell. 2024. [Newswire: A large-scale structured database of a century of historical news](#).