LDK 2023

SLIWC, Morality, NarrOnt and Senpy Annotations: four vocabularies to fight radicalization

J. Fernando Sánchez-Rada and **Oscar Araque** and **Guillermo García-Grao** and **Carlos Á. Iglesias**
{jf.sanchez,o.araque,g.ggrao,carlosangel.iglesias}@upm.es
Department of Telematic Engineering Systems
Intelligent Systems Group (GSI)
ETSI de Telecomunicación, Universidad Politécnica de Madrid

Abstract

This paper describes the vocabularies used in PARTICIPATION, a Horizon2020-funded project aimed at preventing extremism, radicalization, and polarization. To fully take advantage of Linked Data, all data in the project need to be expressed in a semantic format, and all annotation services should be accessible through a semantic API. Most of the data can be expressed by extensively leveraging common vocabularies in the Linguistic Linked Data sphere. However, certain key concepts were not present in any of the popular vocabularies, such as ideologies, morality, and narratives. Some types of analysis also required the use of resources aligned with Linguistic Inquiry and Word Count (LIWC) software. As a result, four vocabularies were developed: Senpy Annotations, SLIWC, Morality, and NarrOnt. Senpy Annotations is a vocabulary designed to represent any kind of annotation in the context of NLP services and resources. SLIWC is a vocabulary and SKOS taxonomy that aims to represent LIWC dimensions. The NarrOnt (Narrative Ontology) vocabulary models the concepts of a narrative and an ideology linked to a piece of content. Lastly, morality is a vocabulary for expressing annotations that follow the Moral Foundation Theory (MFT). These vocabularies have been designed and published using Linked Data principles and best practices. Most importantly, they follow an orthogonal design, integrate well with existing vocabularies, and only describe specific parts of a domain. We believe that the usefulness of these vocabularies will extend beyond the scope of this specific project.

1 Introduction

This work stems from efforts to semantically annotate resources and services in the context of PARTICIPATION, a project aimed at detecting and preventing extremism, radicalization, and polarization. According to previous work, the definition of formats and schemas followed a Linked Data approach to take advantage of all efforts of the Natural Language Processing (NLP) community both in the definition of specific vocabularies and in the integration of different vocabularies for new types of analysis and domains. However, the radicalization domain requires the use of techniques and resources that have not been fully incorporated into the Linguistic Linked Data sphere yet. More specifically, we identified the need to express the domain of ideologies, morality, and narratives, as well as resources aligned with the Linguistic Inquiry and Word Count (LIWC) software.

As a result, we have developed four vocabularies: Senpy Annotations, SLIWC, Morality, and NarrOnt. These vocabularies have been designed and published using Linked Data principles and best practices. Therefore, they follow an orthogonal design, integrate well with existing vocabularies, and describe specific domains. As a consequence, we believe they will prove to be useful beyond the context of this specific project.

2 The Linked Data approach

Part of the work in the project involves several types of data processing and visualization of social media content. This includes several sources, such as microblogging platforms, news sites, and social news aggregators. The majority of the processing involves cleaning, filtering, and automatic annotation. However, the specific processes are varied and constantly evolving to deal with the dynamic nature of online social networks and the multidisciplinary nature of the work.

In order to seamlessly deal with multiple sources of information and provide different types of annotation, all data captured from social media is converted to a common semantic format. All other processes then enrich this data by adding semantic annotations to it. Using a common format allows each process to consume data from multiple sources, regardless of its origin. Modelling

each annotation process as an independent additive process allows future growth. Both of these features could be achieved without Linked Data by modelling data as documents and defining each document property separately. On the other hand, using a Linked Data approach is a better alternative for two main reasons. First, the use of existing work reduces the development and modelling effort. There is already a set of well-known formats, protocols, and libraries, most of which rely on web standards, as well as multiple quality vocabularies to express concepts in most domains. Secondly, properly reusing these works translates into interoperability and compatibility with other projects. Lastly, but most importantly, a Linked Data approach results in data that can be understood not only by humans but also by machines. As proof of the last two points, semantically-annotated data could be easily exposed from an endpoint capable of responding to meaningful queries, such as "where was #example hashtag twitted from on January 1st?".

One of the downsides of a Linked Data approach is that many vocabularies may be needed to model the different types of data in the platform. Although this increases interoperability, it requires understanding them well. The following is an overview of all of the existing vocabularies used to model the data in the project:

- Semantically-Interlinked Online Communities (SIOC) (Breslin et al., 2006) [1] The SIOC Core Ontology provides the main concepts and properties required to describe information from online communities (e.g., message boards, wikis, weblogs, etc.) on the Semantic Web.
- Schema.org (Guha et al., 2016) [2] Provides schemas for structured data on the Internet, on web pages, in email messages, and beyond.
- Dublin Core Metadata Initiative (DCMI) (Initiative et al., 2012)[3]. Provides a model for structured metadata to support resource discovery.
- Marl (Westerski et al., 2011)[4]. Marl is a standardized data schema designed to annotate and describe subjective opinions expressed on the Web or in particular Information Systems.
- DBpedia (Auer et al., 2007)[5]. DBpedia is a community project that extracts structured, multilingual knowledge from Wikipedia and makes it freely available on the Web using Semantic Web and Linked Data technologies. In the context of this project, DBpedia serves both as a vocabulary to express properties and, most importantly, as a source of URIs to attach to people, entities, and other encyclopedic knowledge.
- NLP Interchange Format (NIF) (Hellmann et al., 2013)[6]. NIF is an Resource Description Framework (RDF)/Web Ontology Language (OWL)-based format that aims to achieve interoperability between NLP tools, language resources and annotations.
- Onyx (Sánchez-Rada and Iglesias, 2016)[7]. Onyx is a standardized data schema designed to annotate and describe the emotions expressed by user-generated content on the Web or in particular Information Systems.

3 Vocabularies

The vocabularies in the previous section were insufficient to model all the types of annotation necessary for this project. Instead of creating a single vocabulary with all the missing elements, these missing pieces have been separated into smaller individual vocabularies to foster re-usability. To encourage the use of different vocabularies in real-life scenarios, the vocabularies have been grouped under a common umbrella of PARTICIPATION ontologies. They are accompanied by web documentation describing their usage[8].

When designing a vocabulary, it is often necessary to reach a balance between expressiveness and simplicity. More general vocabularies tend to make use of additional nodes, which translates into more nodes in the knowledge graph. This is usually not a problem, other than having the side effect of making queries slightly more complex

[1] http://rdfs.org/sioc/spec/
[2] https://schema.org/
[3] http://purl.org/dc/terms/
[4] https://www.gsi.upm.es/ontologies/marl/
[5] https://www.dbpedia.org/
[6] https://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core/nif-core.html
[7] https://www.gsi.upm.es/ontologies/onyx/
[8] https://www.gsi.upm.es/ontologies/participation

and verbose or nested. But this project imposes additional constraints that make such complexity more difficult. One of the main constraint is use of common formats such as JSON-LD. This makes it so that regular document stores can be used to save commonly accessed data, and annotation services can serve their results in a more developer-friendly format. When translating a knowledge graph to a JSON-LD document (a tree), there are certain degrees of freedom. This is done by design to allow for the same data to be represented using different schemas. Nonetheless, a deep graph structure will translate into a deeply nested document. A common design principle for the vocabularies presented is that the complete annotations (see Section 4) remain reasonably shallow.

3.1 Senpy annotations

As explained in Section 2, the semantic model for text representation has been based on earlier work (Sánchez-Rada et al., 2020). Therefore, it heavily employs the NIF 1.0 (Hellmann et al., 2013) vocabulary and adds annotations through external vocabularies such as Marl and Onyx. Past experience has shown that some aspects of these vocabularies related to how were not limited to each specific domain (e.g., emotion annotation) and could be applied to other NLP tasks such as those involved in this paper. A better strategy would be to express these common parts in its own separate vocabulary.

Hence, a decision was made to design a very simple and modular vocabulary for the sole purpose of expressing annotations in text. This new vocabulary, called Senpy annotations, follows a structure similar to that of the newer versions of NIF. But, in contrast with NIF, this vocabulary can be easily adapted to provide a better mapping in the JSON-LD representation.

The vocabulary revolves around the concept of an annotation (`sa:Annotation`). The `sa:Annotation` class is designed to be used to annotate specific entries, as will be shown in Section 4. Any entity (e.g., a tweet, a lexicon entry) can be tagged with an annotation through the `sa:hasAnnotation` property. To differentiate between annotations to a single element (e.g., in a lexicon) and an annotation that applies to a larger piece of text (e.g., the count of words in a sentence), there is a special type of annotation, `sa:AggregatedAnnotation`. An `sa:AggregatedAnnotation` may specify both how many elements were used in the aggregation (`sa:count`), as well as the ratio of these elements to the total (`sa:ratio`). These classes can be specialized (subclassed) by specific vocabularies for annotation. As an example of this, another vocabulary in this project (which we will explain below) extends Senpy annotations to include the categories in Moral Foundation Theory. In documents, the actual annotations are an aggregate of the individual words/lemmas. Hence, corpora annotations should use the `sa:AggregatedAnnotation`, which also allows quantifying the frequency or ratio of appearance within the text.

3.2 SLIWC

The way in which the Linguistic Inquiry and Word Count (LIWC) (Pennebaker, 2011) program works is fairly simple. Basically, it reads a given text and counts the percentage of words that reflect different emotions, thinking styles, social concerns, and even parts of speech.

An important part of the LIWC project is the LIWC dictionaries. The importance and popularity of LIWC have led other researchers to adopt their annotation conventions and to use the same format to produce dictionaries that are compatible with LIWC programs.

In the Participation project, we have produced a semantic version of the LIWC annotation schema[9]. It reuses the Senpy Annotations ontology to represent the general concepts used in LIWC annotation (e.g., dimensions, categories, word-level dimensions, document-level dimensions, etc.). Then, it uses these concepts to provide elements specific to the LIWC dictionaries, such as specific categories and their hierarchical relation to one another. These categories have been modelled both as an ontology (i.e., classes) and as a SKOS taxonomy so that the hierarchical structure can be exploited independently of the ontological relations.

Using SLIWC to annotate is very simple. To add information about the LIWC category or dimension that is being annotated in a piece of text, an annotation uses the `sa:hasCategory` property, which links to a specific instance in the SKOS taxonomy. The same procedure works both for annotating lexical entries and word-

[9] https://www.gsi.upm.es/ontologies/participation/sliwc

level annotations (`Annotation`) as well as for annotating at a more general document-level (`AggregatedAnnotation`). A simplified example of SLIWC annotations is illustrated in Figure 1.

3.3 Morality

The popularity of LIWC has led to several LIWC-like dictionaries in the wild, such as the Moral Foundations Dictionary (Graham et al., 2009)[10], which includes new annotations on morality. The theory proposes that several innate and universally available psychological systems are the foundations of intuitive ethics (Graham et al., 2013). Each culture then constructs virtues, narratives, and institutions on top of these foundations, thereby creating the unique moralities we see around the world and conflicting within nations, too. The main foundations according to this theory are *care/harm*, *fairness/cheating*, *loyalty/betrayal*, *authority/subversion* and *sanctity/degradation*.

In order to use annotations for morality both in resources (dictionaries) and in the results of analyses, we have developed an extension of the Senpy Annotations ontology that includes the concepts defined in the Moral Foundations Dictionary. In particular, it provides a class for moral annotations and categories for each of the extremes in each of the dimensions/foundations. Moreover, each category is linked to its foundation (e.g., `Harm`, `InGroup`) and the relationship of the category to the foundation (Virtue, Vice). An example of a simple annotation of a tweet can be seen in Figure 2.

3.4 Narrative

The concept of narrative in the NLP community and in the humanities, social, and cognitive sciences is related but generally not synchronized (Piper et al., 2021). However, it is undeniable that recent work on detecting narrative (and counter-narrative) in texts is helping fight extremism and disinformation (Network, 2015; Upal, 2015).

Narrative Ontology (NarrOnt) is a pragmatic model of the ideologies and narratives present in user-generated content, especially on social media. The ontology provides the `Annotation` concept, which directly subclasses `sa:Annotation`. Narratives are represented with the `Narrative` class. Several narratives are included in the ontology, such as `ProReligion`, `CounterSeparatism`, etc. An example annotation of the narrative in a Tweet is illustrated in Figure 3.

[10] https://moralfoundations.org/other-materials/

4 Use case

The set of vocabularies in this work has been evaluated in two ways. First, we apply them in different scenarios using real excerpts of data extracted from social networks. The following two sections distill this process using placeholder data, with the main purpose of exemplifying the use of these vocabularies in a more realistic scenario where multiple annotations are needed. The examples will cover two distinct use cases separately: annotating corpora (i.e., set of Tweets with different labels) and annotating lexicons (i.e., dictionaries).

The second means of evaluation for these ontologies is their use in the project: to enable the creation of four different morality and narrative detection services; to automatically annotate more than 1,2 million tweets and 100,000 comments on Reddit using multiple services (including morality and narrative); and to power multiple dashboards for the exploration of radicalism in English, Italian, and Spanish, using the enriched data; to power advanced queries for advanced project partners, using SPARQL.

4.1 Annotation of a corpus of microblogging posts

The annotation of microblogging posts followed a model similar to TweetsKB (Fafalios et al., 2018), a public RDF corpus of anonymized data for a large collection of annotated tweets. As most of the annotated corpora in the Participation project and that of TweetsKB were limited to Twitter, we will refer to microblogging posts as tweets. Nevertheless, the model can be easily applied to any similar platform, such as Mastodon or BlueSky.

In TweetsKB, the information retrieved from a tweet is represented by the `sioc:Post` class. The SIOC Core Ontology, Schema.org and DCMI provide properties and attributes for most of the relevant fields in a tweet, such as the `soic:content` attribute for the text, `soic:has_creator` for the author user, `schema:inLanguage` for the language on which it is written, `schema:mentions` for its hashtags, `dc:created` for the creation date, and

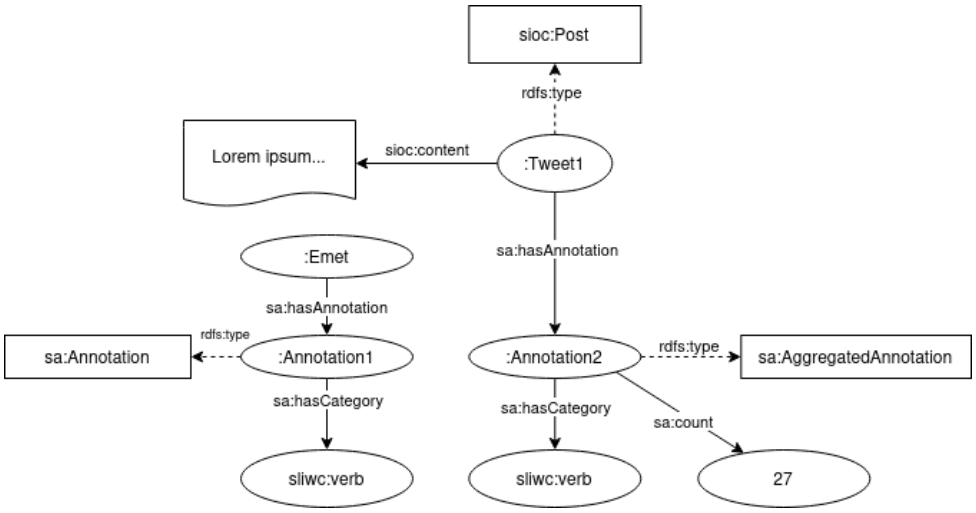

Figure 1: Example of LIWC-aligned annotations of a Tweet.

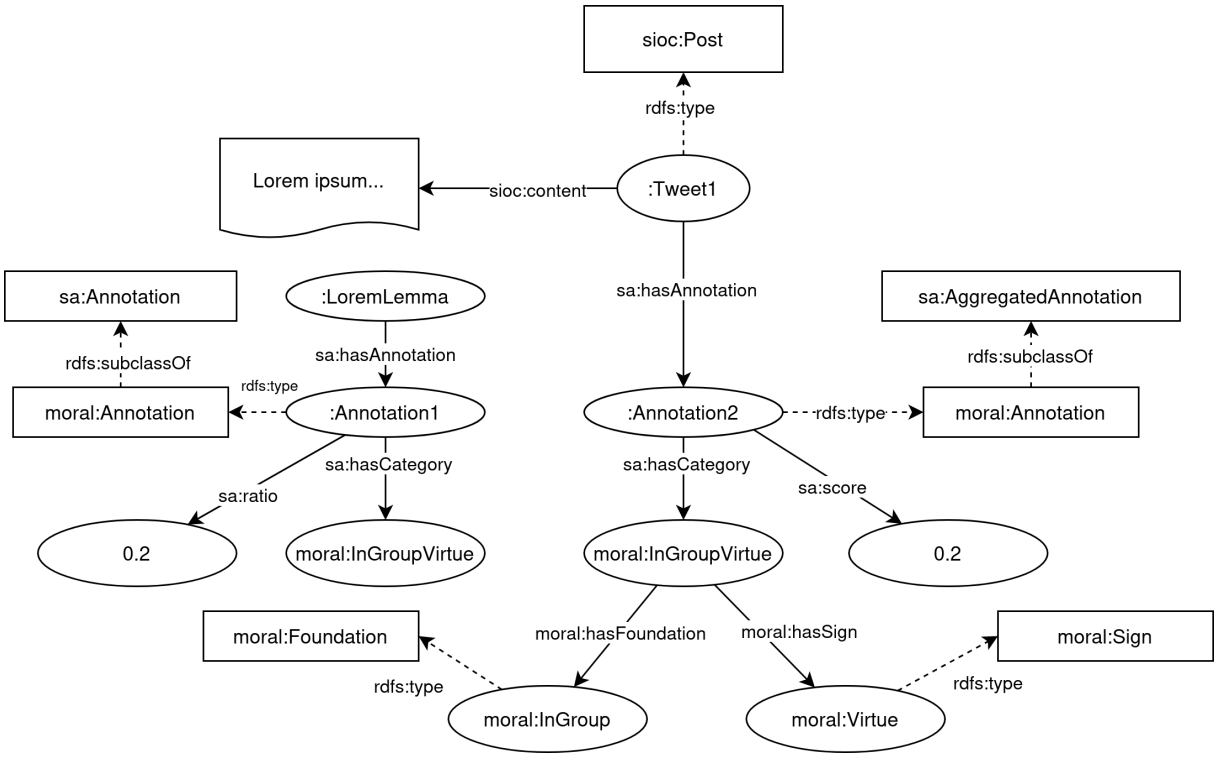

Figure 2: Example of annotation of morality (MFT) in a tweet and in an LIWC-aligned lexicon entry using the Morality ontology.

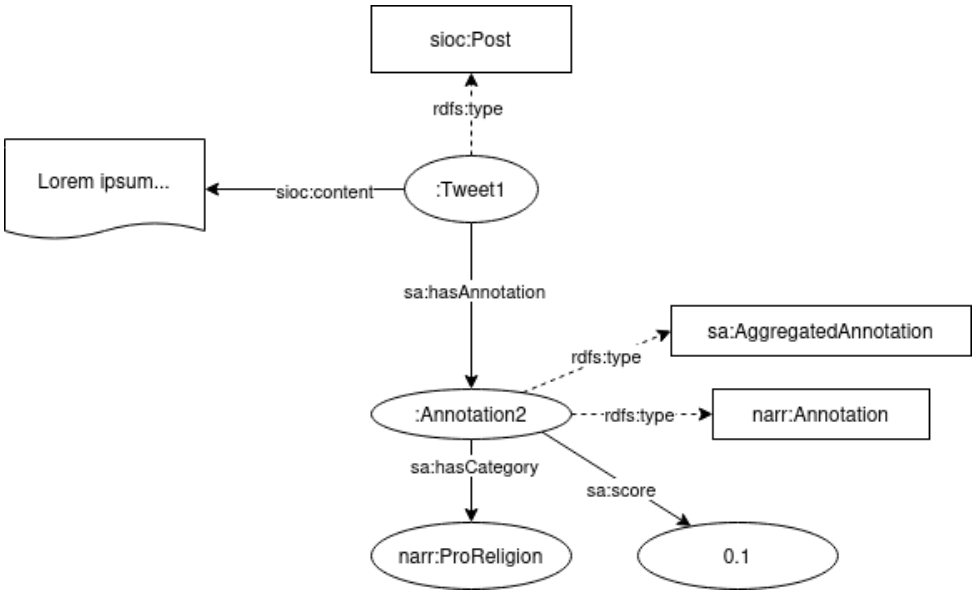

Figure 3: Example of annotation of the narrative in a tweet and a lexicon entry using the Narrative ontology.

`schema:locationCreated` for the location it was posted from.

Lastly, tweets can also be annotated with emotion labels, which are represented using the Onyx ontology. An element is annotated with emotions through the `onyx:EmotionSet` class and the `onyx:hasEmotionSet`.

An `onyx:EmotionSet` is comprised of one or more emotions, defined as `onyx:Emotion`, where the properties `onyx:hasEmotionCategory` and `onyx:hasEmotionIntensity` represent the type of emotion and value, respectively. Finally, the `nif:isString` property from the NIF ontology is used to provide compatibility with other NLP services.

A complete example of the annotation of a tweet can be observed in Figure 1.

It is important to note that the above example can be trivially translated into a mostly flat tree structure, making it ideal for representation as a JSON-LD document.

4.2 Annotating a lexicon

The annotation of a lexicon is very similar to that of a tweet. In this case, the difference is that lexical entries are represented using the lemon ontology. An example annotation of a lexicon can be observed in Figure 2.

4.3 Semantic queries

The data in the project is available to experts through an instance of Fuseki, allowing them to perform semantic queries through SPARQL Protocol and RDF Query Language (SPARQL).

For instance, it is possible to write a query that returns the narrative of every tweet that contains words from a specific Linguistic Inquiry and Word Count (LIWC) category, as well as the ratio at which that category appears. Figure 4 shows such a query, with the LIWC category of Death. An excerpt of the results returned by Fuseki can be observed in Figure 5.

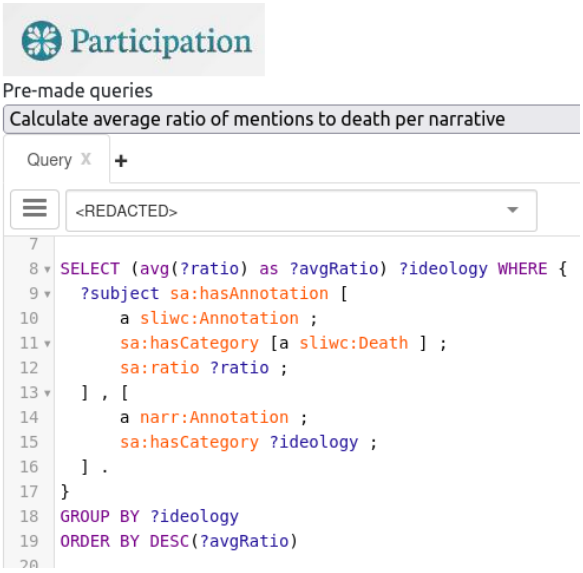

Figure 4: Example SPARQL query that fetches the ratio of the LIWC Death annotation for each narrative.

Listing 1: Example of annotation of a corpus entry

```
@prefix sa: <http://www.gsi.upm.es/ontologies/participation/senpy/ns
   ↪ #> .
@prefix sliwc: <http://www.gsi.upm.es/ontologies/participation/sliwc/
   ↪ ns#> .
@prefix narr: <http://www.gsi.upm.es/ontologies/participation/
   ↪ narrative/ns#> .
@prefix moral: <http://www.gsi.upm.es/ontologies/participation/moral/
   ↪ ns#> .

:Tweet1 a sioc:Post ;
    sa:hasAnnotation [
         a sa:AggregatedAnnotation ;
         a narr:Annotation ;
         sa:hasCategory narr:ProReligion ;
         sa:ratio 0.1 .
   ] ;
    sa:hasAnnotation [
          a sa:AggregatedAnnotation ;
          sa:hasCategory moral:IngroupVirtue ;
          sa:ratio 0.1 ;
    ] ;
    sa:hasAnnotation [
          a sa:AggregatedAnnotation ;
          sa:hasCategory sliwc:Filler ;
          sa:ratio 0.34 ;
          sa:count 23 .
    ] ;
    sa:hasAnnotation [
          a sa:AggregatedAnnotation ;
          sa:hasCategory sliwc:Adverb ;
          sa:ratio 0.15 ;
          sa:count 11 .
       ] .
```

Listing 2: Example of lexicon annotation

```
@prefix sa: <http://www.gsi.upm.es/ontologies/participation/senpy/ns
   ↪ #> .
@prefix sliwc: <http://www.gsi.upm.es/ontologies/participation/sliwc/
   ↪ ns#> .
@prefix moral: <http://www.gsi.upm.es/ontologies/participation/
   ↪ morality/ns#> .

_:compassion a lemon:Lexicalentry;
     lemon:sense [
         lemon:reference wn:synset-fear-noun-1;
         sa:hasAnnotation [
            a sa:Annotation, moral:Annotation ;
            sliwc:hasCategory moral:IngroupVirtue .
         ] .
    ] ;
    sliwc:hasAnnotation [
         a sa:Annotation, moral:Annotation ;
         sliwc:hasCategory moral:IngroupVirtue .
    ] ;
    lexinfo:partDfSpeech lexinfo:noun .
```

	ratio	ideology
1	"4.166667e-02"^^xsd:double	narr:Far_right
2	"7.142857e-02"^^xsd:double	narr:Far_right
3	"5.405405e-02"^^xsd:double	narr:Far_right
4	"4.761905e-02"^^xsd:double	narr:Far_right
5	"2.222222e-02"^^xsd:double	narr:Far_right
6	"2.564103e-02"^^xsd:double	narr:Far_right

Figure 5: Part of the triples returned by the query from Figure 4.

```
SELECT ?text ?category ?ratio
WHERE {
  ?subject sa:hasAnnotation [
        a narr:Annotation ;
        sa:hasCategory [ a narr:ProFar_right ] ;
      ] ;
      dc:created ?date ;
      nif:isString ?text ;
      sa:hasAnnotation [
        a moral:Annotation ;
        sa:hasCategory ?category ;
        sa:ratio ?ratio
    ]
}
ORDER BY ASC(?date)
```

Figure 6: Example SPARQL query that fetches the text of every Pro far-right tweet and their moral values.

Another example, displayed in Figure 6, demonstrates how to get the text of all tweets from a specified narrative, specifically pro far-right. This query also requests the moral categories present in the text and their ratios. It also orders the results by ascending date. Figure 7 shows a fragment of the result from that query.

5 Conclusions and future work

This work shows a successful use case of semantically annotating resources using a mixture of existing vocabularies and ad-hoc vocabularies for niche or otherwise unexplored domains. In particular, four vocabularies have been presented, which can be used independently or in conjunction. When analyzed in isolation, these vocabularies are rather simple by design. But their true power lies in their composition and orthogonal design, which is a testament to the power of the Linked Data approach. Although these vocabularies have been conceived with the main use case of fighting radicalism in the PARTICIPATION project, they have also been designed with extensibility, composability, and reusability in mind. We hope that this work will inspire other researchers to use these vocabularies, extend them, and share their results with the community.

	text	category	ratio
1	"<user> <user> <hastag> happynewyear<number> to my christ believing brother mike pompeo, as indeed to vp <allcaps> <user> and to the greatest president in living history <hashtag> maga <allcaps> <allcaps> <user> also to all americans and my fellow english patriots and believers in the truth! god bless you <url>"	moral:IngroupVirtue	"4.347826e-02"^^xsd:double
2	"make the start-up nation a vacci-nation! <hastag> vaccination <hastag> better <hastag> maga"	moral:IngroupVirtue	"2.222222e-01"^^xsd:double
3	""remember, comrades, your resolution must never falter. no argument must lead you astray" george <allcaps> orwell <allcaps> - animal farm. <hastag> lockhimup <hastag> trumptapes <hastag> trumptreason <hashtag> maga <allcaps> <allcaps> <hashtag> gop <allcaps> <allcaps> <hastag> trumpmarch <hastag> trumpsupporters <url>"	moral:IngroupVirtue	"3.448276e-02"^^xsd:double
4	"in <hastag> hannover sprach <hastag> javidkistel am wochenende von nanopartikeln in <hastag> corona-impfstoffen: „mit denen können sie euch steuern.“ unter den teilnehmern bei <hastag> h<number> waren zudem sympathisanten der „<hastag> querdenken“-bewegung und „<hastag> qa <allcaps>non“ <hastag> dud <allcaps><number> <hastag> impfgegner <url>"	moral:HarmVirtue	"2.777778e-02"^^xsd:double

Figure 7: Part of the response from the query in Figure 6.

Acknowledgements

This work is part of the PARTICIPATION project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 962547.

References

Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. Dbpedia: A nucleus for a web of open data. In *The Semantic Web*, pages 722–735. Springer.

John G Breslin, Stefan Decker, Andreas Harth, and Uldis Bojars. 2006. Sioc: an approach to connect web-based communities. *International Journal of Web Based Communities*, 2(2):133–142.

Pavlos Fafalios, Vasileios Iosifidis, Eirini Ntoutsi, and Stefan Dietze. 2018. Tweetskb: A public and large-scale rdf corpus of annotated tweets. In *European Semantic Web Conference*, pages 177–190. 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 *Advances in experimental social psychology*, volume 47, pages 55–130. Elsevier.

Jesse Graham, Jonathan Haidt, and Brian A Nosek. 2009. Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5):1029.

Ramanathan V Guha, Dan Brickley, and Steve Macbeth. 2016. Schema. org: Evolution of structured data on the web. *Communications of the ACM*, 59(2):44–51.

Sebastian Hellmann, Jens Lehmann, Sören Auer, and Martin Brümmer. 2013. Integrating NLP using linked data. In *International Semantic Web Conference*, pages 98–113. Springer.

Dublin Core Metadata Initiative et al. 2012. Dublin core metadata element set, version 1.1.

Radicalisation Awareness Network. 2015. Counter narratives and alternative narratives. *Ran Issue Paper*.

James W Pennebaker. 2011. Using computer analyses to identify language style and aggressive intent: The secret life of function words. *Dynamics of Asymmetric Conflict*, 4(2):92–102.

Andrew Piper, Richard Jean So, and David Bamman. 2021. Narrative theory for computational narrative understanding. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 298–311.

J. Fernando Sánchez-Rada, Oscar Araque, and Carlos A. Iglesias. 2020. Senpy: A framework for semantic sentiment and emotion analysis services. *Knowledge-Based Systems*, 190:105193.

J Fernando Sánchez-Rada and Carlos A Iglesias. 2016. Onyx: A linked data approach to emotion representation. *Information Processing & Management*, 52(1):99–114.

Afzal Upal. 2015. Alternative narratives for preventing the radicalization of Muslim youth. *Journal for Deradicalization*, (2):138–162.

Adam Westerski, Carlos Angel Iglesias, and Fernando Tapia Rico. 2011. Linked opinions: Describing sentiments on the structured web of data. In *SDoW@ ISWC*.