IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 7, 2021, accepted December 29, 2021, date of publication December 31, 2021, date of current version January 11, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3140030

Extraction of Poetic and Non-Poetic Relations From of-Prepositions Using WordNet

CHRISTIANA PANAYIOTOU
Cyprus University of Technology, 3036 Limassol, Cyprus
e-mail: pcamm16@gmail.com

ABSTRACT The main goal of this paper is to extract the semantic relations underpinning the concepts of English prepositional of-constructions derived from poetic and non-poetic datasets, using Princeton WordNet. The problem is addressed by two different algorithms, which are evaluated for their ability to model the different types of resources from which the relations are derived, and for their ability to predict unseen relations. The first algorithm introduces the concept of subsumption hierarchy between relations in order to derive the most general relations associated to each type of data source and identify a set of relations specific to each dataset. The second algorithm investigates the use of a weighting scheme in order to establish the importance of each association extracted. Of particular importance are the notions of subsumption hierarchies between relations (expressed as synset pairs) and the Inverse Relation Frequency (IRF) measure, which is inspired by the Inverse Document Frequency measure used in Information Retrieval. The ontological prospects of using Princeton WordNet and the above algorithms for the creation of ontologies are also briefly discussed. Although the main interest of the proposed methods lies to the identification of conceptual relations particular to poetic resources, the methods followed can be applied and are evaluated on other domains too.

INDEX TERMS Semantic relations extraction, ontology learning, NLP, knowledge representation.

I. INTRODUCTION

Several classification algorithms for natural language resources are cited in literature with the most promising ones over the recent years focusing on the use of deep learning methods for the automatic classification of phrases and texts [1]–[3]. What is missing from these approaches is the extraction and representation of the semantic relations between concepts that play a prominent role in the representation of domains, and in the prediction of resources from which phrases are derived by reference to the semantic relations specific to each resource. A representation of the semantic relations underpinning different sets of resources will help to identify domain specific associations between concepts and draw inferences from the existing relations. It will also help to provide insight about the extend to which the sets of relations underpinning different datasets overlap.

The primary goal in this paper is twofold: (i) to study the use of Princeton WordNet (PWN) [4] in the extraction of semantic relationships underpinning poetic versus non-poetic datasets or domains of of-prepositions, and (ii) to use the relations extracted from a (training) subset of each type of resource, in order to predict the resource type of relations, not included in the training set. Although the experiments were contacted primarily on poetic and non-poetic resources, the proposed methods can also be used to model any type of resources. For this reason, they have been tested on other domains too. The ontological consequences of the proposed approach and the potential of using the extracted relationships to create ontologies are also discussed.

To this end, the current work proposes two algorithms, Algorithm 1 and Algorithm 2. Algorithm 1, was briefly introduced in [5] from a different perspective, and was tested on a very small training set; it was intended to illustrate the conceptual discrepancies between poetic and non-poetic of-prepositions. However, due to the contextualization of of-prepositions, the wide use of overlapping of-prepositions between poetic and non-poetic resources, and the scarcity of of-prepositions, an annotation of a sufficient volume of of-preposition figures of speech would be very difficult. Instead, the current article focuses on the automatic extraction of the actual concepts underpinning the datasets from which the phrases are extracted, aiming at identifying those relations between the lexical terms of of-prepositions that are

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

associated to the particular datasets from which the textual resources are derived. In this work, Algorithm 1 is evaluated on larger datasets, and for its performance as a Prediction algorithm using the measures of precision and recall.

Algorithm 2, proposes a weighting scheme that evaluates the importance of each semantic relation associated to each of-construction. The weighting scheme is inspired by Information Retrieval measures [6]. For example, the *Inverse Relation Frequency* (IRF) measure forming part of the proposed weighting scheme, takes into consideration the scarcity of English of-prepositions, but also the impact of the hierarchical relations between concepts relevant to the of-prepositions used.

Systems incorporating PWN in the tasks of text classification and disambiguation [7]–[11] have been cited in literature and address, in their majority, the problem of conceptual similarity in document classification via the extension of the representation of document features, with semantic features derived from PWN. These systems, have recorded an improvement in their performance resulting from the integration of PWN in the representation of documents and, or document categories.

The importance of the proposed approach lies to the fact that it focuses on the extraction of semantic *relations* (expressed via hypernym pairs) underpinning the particular data resources (gathered as datasets or domains of knowledge), rather than on the integration of PWN features into a vector representation of text, for classification. It proposes (subordinate - superordinate) hierarchical associations between *relations* which are amenable to ontological representation. Unlike clustering methods, which focus on the establishment of similarity measures between documents (via the clustering of words into concepts), the current work uses the relationships between the PWN entities extracted from data into the identification of the corpus, or dataset from which the data is derived. The term *relations*, in this case refers to the association of concepts involved in each of-preposition.

Example 1: The phrases 'book of God', and 'pipe of Hermes' constitute examples of of-prepositions extracted from poetic resources. They are represented by the relations (book, God) and (pipe, Hermes), where each of the words *book*, *God*, *pipe*, and *Hermes* are interpreted via the PWN concepts (*synsets*): book.n.01, god.n.01, pipe.n.01, and hermes.n.01 respectively, giving rise to the associations: (book.n.01, god.n.01), and (pipe.n.01, hermes.n.01). The hierarchical organization of concepts in PWN makes it possible to extract more generic associations by traversing through its lexical inheritance structure. For example, since the concept artifact.n.01 is more generic than book.n.01, and deity.n.01 is more generic than god.n.01 in PWN, then the association (book.n.01, god.n.01) entails the more generic association (artifact.n.01, deity.n.01). In the current example, by defining hierarchical relations between associations of concepts that can be mapped to upper-level ontologies (as discussed in Section VII), it is possible to make inferences like, for example, that 'a book of God' is an 'artifact of Deity'.

Example 2: Consider the of-prepositions: 'wheel of phoebus', 'daughter of jove', and 'friend of zeus' extracted from the poetic dataset. Each of these of-prepositions can be considered as an ordered pair of noun expressions: (wheel, phoebus), (daughter, jove), and (friend, zeus).
Each noun word of these ordered pairs (relations of lexical terms) has a corresponding concept entry in the PWN lexical database. The words: *wheel*, *phoebus*, *daughter*, *jove*, and *zeus* can be interpreted by the PWN concepts: wheel.n.01, apollo.n.01, daughter.n.01, jupiter.n.02, friend.n.01, and zeus.n.01, respectively. Then, each of the of-prepositions of the example, can be represented by the relations: (wheel.n.01, apollo.n.01), (daughter.n.01, jupiter.n.02), and (friend.n.01, zeus.n.01). Following the hierarchical structure of PWN, each of the senses: wheel.n.01, daughter.n.01, bird.n.01, and friend.n.01, is more specific (referred to as *subordinate* in PWN) than the entry object.n.01. In addition, all of: apollo.n.0, jupiter.n.02, angus_og.n.01, and zeus.n.01 are subordinates of the concepts: belief.n.01, deity.n.01, and psychological_feature.n.01.

The advocated subsumption hierarchy between associations PWN concepts, makes it possible to infer that all of the above associations between PWN concepts, are more specific than the associations: (object.n.01, deity.n.01), (object.n.01, belief.n.01), and (object.n.01, psychological_feature.n.01). Among these associations of synsets (*hyper-pairs*) only (object.n.01, deity.n.01) is associated to of-prepositions extracted purely from poetic resources. This is because (object.n.01, psychological_feature.n.01) also represents of-prepositions extracted from non-poetic resources, for example: 'area of harvest', 'fruit of experience'. Also, the phrases 'leader of opinion' and 'system of belief' are examples of (object.n.01, belief.n.01) in the non-poetic dataset. The objective is to extract the most generic associations of PWN concepts that are associated to a particular dataset.

The phrases explored in this work, are prepositional phrases of the form *N1-of-N2*, where *N1* and *N2* denote noun terms (in some cases multi-word noun terms), which enable us to benefit from the noun hierarchies in PWN. In Linguistics, they are widely known under the terms *of-genitives*, or *of-genitive constructions* [12], or simply *of-constructions*, and sometimes they take the form (Article)-N-of-NP, meaning that they consist of a nominal followed by a prepositional phrase with the preposition *of* generally playing the role of a modifier [12]. *Of*-genitive constructons attracted the interest of several researchers due to their different roles in the interpretation of the whole sentences [13]–[15]. The present work is not concerned with the inherent ambiguity, or typology of of-constructions [12], [16], which is a challenging task on its own.

In Poetry, natural language is used as a tool (via figures of speech [17]) to convey ideas extending beyond the common literal meaning of the words used in a phrase. For example, the preposition *'the roses of heaven'* appeals to the imaginary, and conveys a symbolic meaning beyond the literal meaning of the phrase, that can occur only in imaginary planes. Phrases of this form are strongly related to the author identity [18] and recognition of these phrases sets the foundations for further analysis and understanding of the symbolism used by authors. Prepositions appearing abundantly in poems may be common prepositions appearing in spoken, and written, non-literary text. To classify of-prepositions strictly as 'poetic' versus 'non-poetic' in literary text would be very difficult, since 'poetic' prepositions are scarce, and are highly contextualized. Especially in narrative poetry, non-poetic prepositions appear frequently in poem verses.

An important use of PWN is in the task of Word Sense Disambiguation (WSD) [19]–[22], which is used in text classification [102], text clustering [7], [8] and ontology learning [23], to improve performance. Despite the importance of WSD, the current work does not address WSD as this task constitutes a separate task on its own. PWN has also been employed extensively in the task of ontology alignment. The use of PWN in extracting the models of text is encouraged by the fact that there is substantial research in mapping PWN concepts to large-scale upper ontologies [24], [25], like SUMO [26].

The contribution of this paper is summarized as follows:

1) Algorithm 1 identifies those semantic relations that are specific to each one of two different types of resources (poetic versus non-poetic). In order to do that it introduces the notion of *subsumption hierarchy* between concept relations (*hypernym pairs*). By using a subsumption hierarchy between relations of concepts representing of-prepositions, it identifies the set of the most generic relations specific to each type of resource considered. Algorithm 1 is evaluated for its ability to extract the conceptual relations of the of-prepositions extracted from two different types of resources, and for its use as a resource prediction algorithm. The questions answered for this algorithm are whether it can be used to extract the semantic relations underpinning each resource of of-constructions considered, and, whether these relations can be used to predict the resources of unseen of-prepositions.
2) Algorithm 2, uses a weighting scheme in order to assess the importance of each semantic relation (*hypernym pair*) *associated to* each of-preposition. The IRF forms part of the calculation of the weight of each semantic relation. Its role in Algorithm 2 is to reduce the value of generic relations that inherit the weights of more specific relations. Algorithm 2 is evaluated for its ability to determine the resource from which each of-preposition is extracted and for its use as a prediction algorithm.

The paper is outlined as follows: Section II introduces PWN [4] and discusses some of the most important works in the areas of classification and clustering using PWN, linguistics research in of-genitives, and ontology learning from text. Section III, introduces some basic concepts and definitions necessary for the understanding of this work. Included in this section, are the definitions of the notions of *hyper-pair*, *Subsumption between hyper-pairs*, and *Association* between hyper-pairs, and of-constructs, which are vital for the understanding of the article. Section IV describes the proposed algorithms of this work. Section V describes the resources used and the experiments performed to evaluate the proposed algorithms. Section VI analyzes the results obtained from the experiments performed to evaluate the proposed algorithms. Section VII, discusses the ontological consequences that result from the representation of subsumption relations between associations of concepts and the potential of using the hyper-pair subsumption relations advocated, in order to enrich ontology entailment, via a toy kinship ontology example using PWN taxonomic relations. Section VIII, is a conclusion section summarizing the conclusions drawn from the current work and future work. The results of this work are influenced by the method used to map the lexical concepts in of-prepositions to PWN entities. Named Entity Recognition was used to map lexical entries not included in the vocabulary of PWN to types of entities. The natural language methods used to identify and represent relations of entities in PWN are discussed in Appendix A.

II. RELATED WORK

The current work makes extensive use of the *synonymy* and *hyperonymy* associations between lexical entities as defined in PWN [4], [27]. PWN is a lexical database developed at Princeton University by a group of Psychologists and Linguists. It is underpinned by a hierarchical organization of synonyms (different meanings) of word forms [28], based on the observations and psycholinguistic theories of human lexical memory advocated by Miller and Johnson-Laird [29], about the factors determining linguistic and lexical knowledge.

The basic concepts upon which PWN is based, are the concepts of *synonymy*, *hyperonymy*, *hyponymy*, *meronymy*, and *antonymy*. Each lexical entry (word) is linked to a set of possible word meanings (senses), called *synsets*. A classical example is the lexical term 'bank', which appears under 18 different *synsets*, derived from the different senses of the word 'bank', and from the different grammatical categories in which each sense can appear (for example, verbs and nouns). Each synset is associated to more generic terms (super-ordinate concepts referred to as *hypernyms*), which can be reached by traversing through the lexical inheritance structure of PWN [4], [28], [30]. More specific terms are included as *hyponyms* of each synset. The hierarchical structure in which lexical entries[1] are organized, is influenced by experimental evidence suggesting a hierarchical organization of human lexical memory [29].

Due to its hierarchical organization and definition of synonym senses, PWN has been cited and integrated in many applications as an Ontology [31]. Tom Gruber's most cited definition of an ontology, states that an ontology is *'an*

[1]Only nouns and verbs are organized in a hierarchical structure

explicit specification of a conceptualization' [32]. Following this definition, Borst [33] defined an ontology as a *'a formal specification of a conceptualization'*. Studer *et al.* [34] combined the two definitions, defining an ontology as a *'a formal, explicit specification of a shared conceptualization'*. In [35] the notion of formality is interpreted as the requirement that an ontology be represented via a knowledge representation language with formal semantics. Also, in [36] the authors emphasized the need for a formal specification of conceptualizations that are expressed by means of PWN's synsets. Due to the fact that the hierarchical relations between entities in PWN are not formally defined via a knowledge representation language, PWN may be described as an informal lightweight ontology [37].

However, the informal nature of lexical relations supported by PWN make it particularly useful in the extraction of information from textual resources due to the direct mappings of word forms to its lexical entries and the consequent ability to traverse its hierarchical structure for synonymous, similar, or more generic terms. In addition, many ontologies provide direct mappings to the PWN concepts, which makes it possible to use PWN to map word forms to ontology concepts.

The taxonomic relations supported by PWN, made it a desirable candidate for use in the alignment of ontologies (for example, via the mapping of synonymous terms and hyperonymy-hyponymy relations) [24]–[26], [37] and in the task of disambiguation [20], [22] of words in phrases. There are several methods for mapping, or merging ontologies using PWN as a background knowledge. Examples are, the HCONE [25], [39] approach for ontology merging, and the direct mappings between PWN entries and the Suggested Upper Merged Ontology (SUMO) [26].

The integration of PWN in text classification and clustering algorithms in several works by incorporating synsets and taxonomic features in the representation of terms and documents, has shown an improvement in their performance [8]–[10]. For example, Rodriquez *et al.* [9] employed PWN in supervised classification of text documents by integrating manually extracted synsets in the vectors of terms representing text categories, in the Rochio [111] and the Windrow-Hoff algorithm. In both cases, the use of PWN synsets improved the results of classification. In [8] synsets were also employed as features for document representation and clustering. The results showed improvement when weights related to synsets were concatenated to the term vectors of the documents, their disambiguation strategy involved information about sub-ordinate and superordinate synsets, and the information about more super-ordinate classes included information about their subordinates.

To date, most of the work employing PWN in a classification or clustering task (as shown in the above examples), extend the vector model for the representation of documents so as to include information about synonyms and, or hypernymy-hyponymy.

Ontology learning is concerned with the extraction of ontological knowledge from textual resources and the representation of this knowledge via an appropriate computational formalism. Ontology learning from text is a multi-phase process involving the application of many methods and tools. Wong *et al.* [40] classified the techniques and resources employed for ontology learning under the headings: Statistics-based techniques, Linguistics-based techniques and resources, and Logic-Based Techniques and Resources.

Buitelaar *et al.* [41] depicted the process of ontology learning as an Ontology Learning Layer Cake where each layer corresponds to a different step in the ontology learning process which can be accomplished via a number of tools and techniques. The different steps are: terms extraction from unstructured text, terms synonyms extraction, concepts formation, concept hierarchy extraction, non-taxonomic relations extraction, axiom schemata instantiation, and general axioms extraction. Asim *et al.* [42] suggested a similar methodology for ontology learning consisting of many sequential stages including: Preprocessing of unstructured text, terms/concept extraction, relation extraction, axiom induction, and evaluation.

Ontology learning from text may include an amalgam of Linguistic-Based and Statistical -Based techniques. Examples of Linguistic-based techniques are: tokenization, part-of-speech (pos) tagging, sentence parsing and structural analysis, and dependency parsing in order to identify dependencies between the terms. Semantic lexicons and PWN may be used to identify synonyms and relevance of terms.

Hearst [43] advocated the use of lexico-syntactic patterns in order to extract particular relations between terms. For example, the lexico-syntactic pattern: $\mathrm{NP_c}$ **such as** $\mathrm{NP_I}$ was used to extract phrases of the form: 'Presidents `such as` Obama', where the first noun phrase ($\mathrm{N}P_c$) denotes a concept and the second noun phrase ($\mathrm{N}P_I$) denotes an instance of that concept. Following Heirst, many researchers have applied lexico-syntactic patterns to extract relations from particular domains. For example, Mukherjea *et al.* [44] used an enriched set of lexico-syntactic patterns produced by Cimiano *et al.* [45] to extract biometical relations from the World-Wide Web.

Subcategorization frames were advocated by Agustini *et al.* [46]; they are based on the realization that each word constraints the words with which it can combine with any other word within a syntactic structure. The constraints imposed by each word are referred to as subcategorization restrictions [46]. These restrictions can be both syntactic (subcategorization frames) and semantic (called selection restrictions).

Statistical analysis techniques for the purpose of extracting and assessing the similarity of terms include, but are not limited to, the identification of collocations of terms [47], the evaluation of mutual information of terms, the assessment of the Semantic Similarity of terms, for example via latent semantic analysis [48], [49], Formal Concept Analysis [50], and Hierarchical Clustering [51] for the automatic acquisition of hierarchical relations from text; including also, conceptual clustering algorithms for word and verb classifications aiming at the acquisition of concepts [52].

Up to date, there are several tools and methods for the automatic or semi-automatic extraction of ontological knowledge from texts. Below, are only but a few of these approaches [51], [53], [54] below, with a brief description for each.

1) ASIUM [51] is a system for extracting semantic knowledge from parsed technical texts. Parsed texts are produced by the Sylex syntactic parser [96] which identifies instantiations of case frames of the form: ⟨verb⟩((⟨preposition⟩ | ⟨function⟩)⟨headword⟩)*. The learning method is based on the assumption that words occurring after the same preposition or function with the same verbs in the instantiated case frames, are instances of the same concept. A conceptual clustering method is applied in order to build associations between basic clusters and create a hierarchy of concepts. The system was applied on domain specific, technical texts with limited vocabulary involving the use of a particular terminology, and the use of mostly action and concrete verbs, which rendered PWN inappropriate to be employed in the acquisition of semantic relations from text.
2) TextToOnto [53] is based on the notion of semantic patterns (specific formats for natural language expressions declaring a specific meaning [53]) that have to be created via the use of a pattern editor manually. Semantic patterns match text fragments to ontological classes (obtained from top level ontologies) and verb groups linked via text constants, like prepositions. The purpose of semantic patterns is to ascribe specific meaning to natural text expressions. A number of possible sentences may be matched to a single semantic pattern where classes and verbs take on group names of synonymous words [53], for example the semantic pattern ⟨Plant Part⟩⟨Becomes.Verb⟩⟨Color⟩. Semantic lexicons including PWN may be used to determine groups of words, or classes of words matched to upper level ontology concepts.
3) OntoLT [54] is a Protégé plug-in, used for the extraction of concepts (Protégé classes) and relations (Protégé slots) from linguistically annotated text collections. The text collections are annotated at various levels providing information about for example, the part-of speech of tokens, morphological information, the clause structure etc. The user is able to define mapping rules through a precondition language, that maps linguistic entities in annotated text collections to concept and attribute candidates via XPath expressions. The predicates and terms of the condition language validate XPath expressions. Linguistic entities may be mapped to a EuroWordNet sense.
4) Another notable approach is demonstrated by the LExO (Learning EXpressive Ontologies) prototype [55]. In this case, definitions of classes expressed in NLP (constituting the corpus of LExO) are parsed by the minipar dependency parser [56], which produces a dependency tree containing information about the syntactic category and grammatical role of each token. A minimal set of rules for building a complete axiomatization of each class is applied, and the resulting list of axioms expressed in the KAON2 [57] internal syntax, is then transformed to the OWL DL representation of the class [58].
5) In [59], the authors advocate a method that enables the automatic extraction of semantic relations from text, and the representation of this knowledge in DL $\mathcal{ALC}$. The textual data, consisting of concept axioms and Wikipedia glossaries, is fed into the syntactic parsing module for tagging and dependency analysis (using Probabilistic Context Free Grammar). Tagged data and dependency trees are further processed by the semantic module. In the tagged text, anything else than nouns and adjectives is discarded. The relations between the extracted terms are recognized in the original phrases through the use of patterns, and rules applied in the relation extraction activity. The final stage is the construction of hierarchical and non-hierarchical OWL DL axioms from the instantiated rules of the previous stage. The method provides for the representation of negated concepts and properties.

A. NOVELTY VERSUS RELATED WORKS

In common with Ontology Learning, the current proposal is concerned with the extraction and representation of conceptual relations. In common with the lexicographical and linked data approaches [60] and Lemon [61], [62], the current proposal attempts to bridge the gap between the lexical form of each word participating into an of-relation and its description by an ontological concept. None of the existing approaches to the author's knowledge raises the issue of a representation of lexical associations of words by hierarchical association between concepts that can in turn relate phrases to phrases represented by more generic terms. The set-theoretic approach followed in the definition of Algorithm 1 enables the proof of important properties about the underlying domains of relations, and is in line with the formal representation of ontological concepts.

The second novelty of the proposed approach, is the suggestion for the prediction of the types of resources from which particular relations are extracted. This suggestion is implemented via an unsupervised approach to the identification of those relations (represented as hyper-pairs) that appear only in one type of resource. Low recall levels lead to the conclusion that the set of relations common to both resources is big. The prediction of the category of phrases, is different from the methods of classification of phrases via the use of supervised, or unsupervised methods, reviewed so far to the author's knowledge. The reason is that under the proposed method described by Algorithm 1, a phrase is predicted to have been extracted from a particular type of resource if it is represented by a PWN hyper-pair that is equally or more generic than relations that are extracted from the same resource and it is not associated to relations in another type of resource. Also, by Algorithm 2, the prediction of the category of resource of each phrase is dependent on the weight of each hyper-pair in each type of resource.

Algorithm 2 follows a statistical approach for the assignment of weight to each hyper-pair. The IRF measure used in the calculation of weights of hyper-pairs has important properties on its own. Together with the notion of subsumption hierarchy between the extracted relations representing of-prepositions, it can lead to the proof of important properties relating of-prepositions. For example, it can determine which hyper-pairs on the same hierarchy path are associated to exactly the same relations.

Pattern-based parsing techniques with conceptual clustering methods are more appropriate in cases where the data resources use a restricted vocabulary and a technical language. In these cases, all relevant patterns need to be stated explicitly. Manually edited semantic patterns can be cumbersome and voluminous. Instead of employing manually edited semantic patterns for the extraction of groups of words of a particular type, the proposed method for extracting relations out of of-prepositions in this work, is based on the automatic extraction of hypernym pairs (the notion of *hyper-pair* is described in Section III) associating pairs of PWN entities to the lexical terms of the resources' of-prepositions.

Although the automatic extraction of hyper-pairs is also constrained by the existence of lexical forms and noun expressions in PWN, the majority of lexical terms involved in of-prepositions can be mapped either to entities in PWN, or to Named entities recognized my NER.

The method advocated in [59] describes an automatic extraction of axioms directly derivable from phrases, in an expressive decidable fragment of description logics. However, the actual terms derived from noun and, or adjective relations may have associations to concepts not explicitly stated in the relations extracted from text. The mapping of lexical terms to PWN entities and the PWN relationships of synonymy, hypernymy-hyponymy, antonymy and meronymy could be used for the explicit representation of these associations that can be used to improve reasoning over conceptual relations.

III. BACKGROUND

The definitions of *hypernyms*, *synsets* and *hypernym_paths* in PWN as cited by George Miller [4], [30], form a fundamental building block in the definition of the algorithms introduced in this work.

Definition 3 (Synset [63]): A synonym set (synset) is a set of words that are interchangeable in some context without changing the truth value of the proposition in which they are embedded.

An alternative weakened version of definition 3 cited in literature, is the following.

TABLE 1. PWN synonym set for word 'library'.

Synset name	Definition
library.n.01	a room where books are kept
library.n.02	a collection of literary documents or records kept for reference or borrowing
library.n.03	a depository built to contain books and other materials for reading and study
library.n.04	(computing) a collection of standard programs and subroutines that are stored and available for immediate use
library.n.05	a building that houses a collection of books and other materials

Definition 4 (Synonymy Per [30]): Two expressions are synonymous in a linguistic context C if the substitution of one for the other in C does not alter the truth value.

Definition 5 (Synset Name): Each synset in the current article is designated with a three-part label that takes the form:$\langle$word$\rangle$.$\langle$pos$\rangle$.$\langle$number$\rangle$. The $\langle$word$\rangle$ part will be referred to as the name of the synset. The $\langle$pos$\rangle$ denotes the part-of-speech (grammatical category of the synset), and the $\langle$number$\rangle$ designates its position in the descending frequency order of sense occurrences. For example, the sense director.n.01 occurs more frequently than the sense director.n.02.

Example 6: The following table shows the synsets of the word 'library' as defined in PWN:

Definition 3 above, provides a truth-functional interpretation of meaning, whereas definition 4 leads to the conclusion that *if concepts are represented by synsets, and if synonyms must be interchangeable, then words in different syntactic categories cannot be synonyms (cannot form synsets)* [30]. Thus, noun lexical terms in of-prepositions should be mapped only to noun synsets in PWN.

Miller *et al.* [4], [30], [64] described synonymy as the fundamental semantic relation which is used to define similar lexicalized concepts. Synsets are related to each other via hyponymy, hypernymy, antonym and meronymy relations [4]. The hyponymy - hypernymy relations describe subordinate - superordinate relations between synsets and are also referred to as ISA relations or as subsumption relations. The semantic relations used in this work are the relations of synonymy, hyponymy, and hypernymy.

Definition 7 (Hyponymy [4], [30]): A concept represented by the synset $\{x, x', \ldots, \}$ is said to be a hyponym of a concept represented by a synset $\{y, y', \ldots, \}$ if native speakers of English accept sentences constructed from such frames as *'An x is a (kind of) y'* [30].

Hyponymy and hyperonymy are dual semantic relations between noun word meanings [4], [30]. The transitive, asymmetric *'is a kind of'* [28] relation is represented via the symbol @→ so that if W_h, W_s are noun synsets, then the expression W_h@→W_s means that W_h is a hyponym of W_s, or equivalently that W_s is a hypernym of W_h [28], [30].

Example 8: In PWN, the synset director.n.01 of the lexical entry 'director' is a hyponym of the synset administrator.n.01, meaning that a director is a kind-of administrator in this sense. Equivalently, the synset administrator.n.01 is a hypernym of the synset director.n.01. This is represented as: director.n.01 @→ administrator.n.01.

Hyponymy-hyperonymy relations form a hierarchical semantic organization between word senses of nouns [28] leading to the existence of paths from each particular sense of a word to each super-ordinate sense in a transitive way.

Definition 9 (Hypernym Path): A hypernym path is a sequence $s_1, \ldots, s_k$ where each $(s_i, s_{i+1}) \in$ @→.

Example 10: An example of a hypernym path of the synset director.n.01 to the most general sense entity.n.01 is:

director.n.01 @→ administrator.n.01 @→ head.n.04 @→ leader.n.01 @→ person.n.01 @→ causal_agent.n.01 @ → physical_entity.n.01 @→ entity.n.01

For each synonym sense (synset) of a lexical entry, there is a set of possibly more than one hypernym paths to the most generic sense entity.n.01. The following function returns the set of all hypernym paths for a synset.

Definition 11 (hypernym_paths(s)): . Let $\mathcal{P}$ be the set of all paths that can be obtained by all hyponymy-hyperonymy relations between word senses. Then, if s_k is a sense of some lexical entry w say, hypernym_paths(s_k) represents the set of paths $\{p_1, \ldots, p_n\}$ where $p_i = s_{k_1} @\rightarrow, \ldots, @\rightarrow s_{k_m}$, $s_k = s_{k_1}$, and s_{k_m} is the most general sense. Whenever necessary to avoid ambiguity, the path will be included underneath the @→ relation, as follows: $s_{k_1} \underset{p_i}{@\rightarrow} \ldots \underset{p_i}{@\rightarrow} s_{k_m}$.

We shall henceforth refer to each path p_i as the sequence $s_{i1}, \ldots, s_{ik}$ where each $(s_{ij}, s_{ij+1}) \in$ @→.

Example 12: hypernym_paths(director.n.01) = {
director.n.01 @→ administrator.n.01 @→
head.n.04 @→ leader.n.01 @→ person.n.01 @→
causal_agent.n.01 @→ physical_entity.n.01 @→
entity.n.01,
director.n.01 @→ administrator.n.01 @→
head.n.04 @→ leader.n.01 @→
person.n.01 @→ organism.n.01 @→
living_thing.n.01 @→ whole.n.02 @→
object.n.02 @ → physical_entity.n.01 @→
entity.n.01 }

Definition 13 (Nodes(p)): Let $p = s_1 @\rightarrow \ldots @\rightarrow s_k$. Then, Nodes($p$) $= \{s_1, \ldots, s_k\}$.

Due to the fact that @→ is a transitive relation, it follows that its transitive closure, @→* say, is identical to @→. The transitive closure of hyponym-hypernym relation adheres to the normal definition of a transitive closure relation:

Definition 14: [Closure @→*] The closure of the relation @→ between two synsets is defined as follows: @→*

$$\begin{aligned} &h_i @\rightarrow\ h_j \Rightarrow h_i @\rightarrow^*\ h_j \\ &h_i @\rightarrow^*\ h_j,\ h_j @\rightarrow\ h_k \Rightarrow h_i @\rightarrow^*\ h_k \end{aligned} \tag{1}$$

In many cases, instead of referring to the transitive closure of the hyponym-hypernym relation, we shall refer to the *closure of a synset* as the set of hypernyms of the synset which are pairwise related via the hyponym-hypernym relation. The closure of a synset is defined below:

Definition 15 (Clos(s)): Let $P_s = \bigcup$hypernym_paths(s). Then, the closure of s is defined as the set $\{h_j : \exists p \in P_s \text{ and } h_j \in \text{Nodes}(p)\}$

The above definitions were necessary in order to define the notions of a *hyper-pair*, the association (Assoc) between a hyper-pair and a *relation* representing an of-construction, and the subsumption relation between two relations representing of-constructions.

Of-prepositions are represented as binary relations, referred to as of-relations, or simply relations where there is no ambiguity. For example, the phrase 'tent of God' [66] is represented as the binary relation (tent, god).

Let $R = R_{\text{poetic}} \cup R_{\text{non\_poetic}}$ represent the set of all of-relations retrieved, where R_{poetic} and $R_{\text{non\_poetic}}$ represent the sets of relations retrieved from poetic versus non-poetic datasets, respectively (referred to as poetic vs. non-poetic relations, for brevity). Each lexical term (normalized word) in an of-relation is mapped to a set of hypernyms by using the method described in Section A. The Cartesian product of the sets of hypernyms derived from the words of an of-relation r say, gives rise to a set of hyper-pairs. Each hyper-pair is considered to be associated to r.

Definition 16 (Hyper-Pair): Let, $\mathcal{S}$ be the set of all synonyms. Then, by the term *hyper-pair* we mean any pair $(h_i, h_j) \in \mathcal{S} \times \mathcal{S}$.

Definition 17 (Assoc): A hyper-pair (h_i, h_j) is *associated to* a relation $r = (c1, c2) \in R$, denoted as: Assoc($(h_i, h_j), r$) if and only if there is a sense s_k of c_1, a sense s_l of c_2 and paths $\{p_i, p_j\} \subseteq \mathcal{P}$ where $p_i \in$ hypernym_paths(s_k) and $p_j \in$ hypernym_paths(s_l), and $h_i \in$ Nodes(p_i) and $h_j \in$ Nodes($p2$).

Hyper-pairs associated to particular relations may also form hierarchical relations. The following notion of 'subsumption' relation between two of-relations is used extensively in the proposed algorithms in Section IV.

Definition 18 (Subsuming Hyper-Pairs): Let $\text{hp}_1 = (h_i, h_j)$ and $\text{hp}_2 = (h_k, h_m)$ be two hyper-pairs. Then, hp_1 subsumes hp_2, denoted as $\text{hp}_1 \sqsubseteq \text{hp}_2$, if and only if $\exists p \in$ hypernym_paths(h_i) and $\exists q \in$ hypernym_paths(h_j) such that: $h_i \underset{p}{@\rightarrow}^* h_k$ and $h_j \underset{q}{@\rightarrow}^* h_m$.

IV. PROPOSAL DESCRIPTION

The questions addressed in this work are:

1) Are there sets of hyper-pairs (taking the most generic) associated to one type of of-constructions that can be used to model the datasets?
2) If the answer to the above question is yes, can these sets of hyper-pairs be used to predict the dataset of an unseen of-relation?

The first question aims at the extraction of associations of concepts specific to the poetic and non-poetic resources of the datasets involved in the experiments. The second question investigates the possibility of using the model extracted from smaller subsets of datasets to predict the source dataset of an unseen of-relation derived from a testing set. These questions are addressed by the development of two algorithms. Each algorithm is described at an abstract level, making use of the definitions provided in Section III in order to serve clarity (the actual implementation is done in Python).

The first algorithm identifies disjoint sets of hyper-pairs that uniquely identify the type of resource from which each relation is extracted. The set-theoretic approach followed in the definition of the algorithm leads to the proof that if a hyper-pair associated to an of-relation belongs to one of these sets then the type of resource from which the of-relation is extracted can be uniquely identified. Another important property that can be formally proved from the definition of the first algorithm is that the measure of recall used in the evaluation of the algorithm is inversely proportional to the size of the set of overlapping of-relations between the different types of resources used in the datasets. Thus, the measure of recall provides feedback about the degree of overlap of the types of resources being considered.

The second algorithm follows a statistical approach and evaluates the importance of each hyper-pair to the type of resource considered. Based on the realization that not all hyper-pairs have the same importance in identifying the resource of of-relations and that some hyper-pairs are more frequent in one type of resource than in another, it assigns a weight to each hyper-pair based on its frequency and the IRF measure. The IRF accounts for the fact that hyper-pairs consisting of the most generic terms have higher frequency than hyper-pairs consisting of more specific terms (for example the hyper-pair (entity.n.01, entity.n.01) is associated to almost every relation).

As shown in Section VI, the IRF measure plays an important role in modeling the underlying domain of a resource, as it can also be used to identify which hyper-pairs in the subsumption hierarchy of hyper-pairs are associated to the same set of of-relations. The notion of Inverse Relation Frequency (IRF) has never been defined before, to the author's knowledge.

A. ALGORITHM 1

Let S_{all} denote the set of all *hyper-pairs* (refer to definition in Section III) associated via the Assoc relation to each of-relation in the corpus. Then,

$$S_{\text{all}} = \bigcup_{r \in R} \{(h_i, h_j) : \text{Assoc}((h_i, h_j), r)\} \tag{2}$$

The set of hyper-pairs associated to relations in the non-poetic and poetic resources respectively, are also defined as follows:

$$S_{\text{np}} = \bigcup_{r \in R_{\text{non-poetic}}} \{(h_i, h_j) : \text{Assoc}((h_i, h_j), r)\} \tag{3}$$

$$S_p = \bigcup_{r \in R_{\text{poetic}}} \{(h_i, h_j) : \text{Assoc}((h_i, h_j), r)\} \tag{4}$$

Obviously, $S_{\text{np}} \cap S_p \neq \emptyset$ since the most generic hypernyms in the hierarchy of synsets in PWN (like abstraction.n.06, and entity.n.01) appear in the hierarchies of synsets of words in the relations of both poetic and non-poetic resources. Let S_c be the set of synsets that appear in the of-relations of both types of resources.

$$S_c = S_{\text{np}} \cap S_p \tag{5}$$

Example 19: The relation r = (choregraphy,nature) is extracted from the poetic resources. A number of hyper-pairs in the corpus are associated to r via the Assoc relation. From those hyper-pairs associated to r, some are associated only to of-relations extracted from the poetic resources, and some to of-relations extracted from both types of resources. For example, the hyper-pairs:

(stage_dancing.n.01, nature.n.01),
(performing_arts.n.01, quality.n.01),
(stage_dancing.n.01, attribute.n.02), and
(performing_arts.n.01, attribute.n.02)

are associated only to poetic resources. whereas the following hyper-pairs associated to r, are also associated to non-poetic relations and they are elements of the set S_c:

(event.n.01, nature.n.01),
(abstraction.n.06, nature.n.01),
(performing_arts.n.01, entity.n.01),
(psychological_feature.n.01, nature.n.01),
(entity.n.01, nature.n.01), and
(act.n.02, quality.n.01).

All the hyper-pairs associated to r in this example, including those in S_c, are currently in S_p.

As shown in Example 19, S_p may include some hyper-pairs associated to both poetic and non-poetic relations (those pairs that are included in S_c), and the same applies for S_{np}. As these hyper-pairs do not provide any valuable information regarding the type of the resources from which the of-relations were extracted, they are removed from both S_p and S_{np}.

$$S'_p = S_p - S_c \tag{6}$$

$$S'_{\text{np}} = S_{\text{np}} - S_c \tag{7}$$

The next step is to delete from S_c, those hyper-pairs that subsume hyper-pairs in S'_p and S'_{np} aiming to leave in S_c only the hyper-pairs of of-relations that are common to both poetic and non-poetic of resources. Removing from S_c those hyper-pairs which are more generic than hyper-pairs in S_p or S_{np} removes from S_c all hyper-pairs associated to relations that exist in only one time of resource (for example poetic) as opposed to the other.

$$\begin{aligned} S_c &:= S_c - \left\{(h_k, h_m) : \exists (h_i, h_j) \in S'_p \ \wedge \ (h_i, h_j) \sqsubseteq (h_k, h_m)\right\} \\ &\quad - \left\{(h_k, h_m) : \exists (h_i, h_j) \in S'_{\text{np}} \ st.\ (h_i, h_j) \sqsubseteq (h_k, h_m)\right\} \end{aligned} \tag{8}$$

Example 20 (Example 19 Revisited): Continuing with example 19 above: Since (stage_dancing.n.01, nature.n.01) $\sqsubseteq$ (event.n.01, nature.n.01), where (stage_dancing.n.01, nature.n.01) $\in S_p$ and (event.n.01, nature.n.01) $\in S_c$, then (event.n.01, nature.n.01) does not provide information specific to poetic resources (choreography, nature) and should be removed from S_c.

The sets S'_p, S'_{np} contain hyper-pairs associated entirely to poetic and non-poetic relations, respectively. The next step is to reduce the size of these sets by removing unnecessary specification.To do this, hyper-pairs subsumed by more generic hyper-pairs in each of S'_p and S'_{np} are removed from these sets.

$$S''_p = \{(h_i, h_j) \in S'_p : \nexists (h_k, h_m) \in S'_p \ st.\ (h_i, h_j) \sqsubseteq (h_k, h_m)\} \quad (9)$$

$$S''_{\text{np}} = \{(h_i, h_j) \in S'_{\text{np}} : \nexists (h_k, h_m) \in S'_{\text{np}} \ st.\ (h_i, h_j) \sqsubseteq (h_k, h_m)\} \quad (10)$$

Example 21 (Example 19 Revisited): Due to the fact that:
(stage_dancing.n.01, nature.n.01) $\in S'_p$,
(performing_arts.n.01, attribute.n.02) $\in S'_p$,
and the following subsumption relation:
(stage_dancing.n.01, nature.n.01) $\sqsubseteq$
(performing_arts.n.01, attribute.n.02)
it follows (by (9)) that:
(stage_dancing.n.01, nature.n.01) $\notin S''_p$.
It follows trivially from the above definitions, that since S''_{np} and S''_p are disjoint sets of hyper-pairs associated to only one type of resource, then if a hyper-pair belongs to one of these sets, it cannot be a member of the other set. Thus, if a hyper-pair of an of-relation is an element of S''_p (versus S''_{np}), then it can be inferred that the relation is extracted from the poetic set of resources (versus. non-poetic).

Proposition 22: For any relation $r \in R$:

$$r \in R_p \quad \text{if } \exists (h_i, h_j) \in S''_p \wedge \text{Assoc}((h_i, h_j), r)$$
$$r \in R_{\text{np}} \quad \text{if } \exists (h_i, h_j) \in S''_{\text{np}} \wedge \text{Assoc}((h_i, h_j), r)$$

The proof of proposition 22 can be traced in Appendix B.

Proposition 22 forms the decision rule for determining poetic versus non-poetic relations. Note, that, when the algorithm is used for the prediction of 'unseen' relations (not included in the training set of relations), only a subset of the available data is used to determine the sets S''_p, and S''_{np} and the set of hyper-pairs associated to the of-relations in the testing set may overlap with both S''_p, and S''_{np}.

The following proposition is also useful in interpreting experiment results (as will be discussed in Section VI).

Proposition 23: If $r \in R_{\text{poetic}}$ then $\forall (h_i, h_j)$ such that $\text{Assoc}((h_i, h_j), r)$ it follows that $(h_i, h_j) \notin S''_{\text{np}}$.
The proof follows trivially from definition of Algorithm 1.

Proposition 23 implies that hyper-pairs associated to a poetic (non-poetic) resource cannot be members of S''_{np}. For example, if $r \in R_{\text{poetic}}$ then $\forall (h_i, h_j) : \text{Assoc}(h_i, h_j, r) \implies (h_i, h_j) \notin S''_{\text{np}}$.

B. ALGORITHM 2

Hyper-pairs form the underpinning semantic relations (features) characterizing each of-relation. However, not all hyper-pairs have the same importance. The importance of each hyper-pair is determined by a weight, which is defined in this section. As before, the set of of-relations is represented with the set $R = \{r_1, \ldots r_k\}$, where each $r_i \in R$ is an ordered pair of words: (c_{i1}, c_{i2}).

Let cat denote the type of resource of a relation, i.e. cat $\in$ {poetic, non_poetic}. The following sequence of steps is followed in order to determine a weight for each hyper-pair, starting from the creation of two separate relation frequency matrices, one for each source of relations.

$$M_f^{\text{cat}}[j, k] = \|\{r_i : h_j \in \text{hypers\_of}(c_{i1}) \ and \ h_k \in \text{hypers\_of}(c_{i2})\}\| \quad (11)$$

The above equation may be restated as:

$$M_f^{\text{cat}}[j, k] = \|\{r : \text{Assoc}((h_j, h_k), r)\}\|$$

where j, k are the indices of h_j and h_k, respectively.

Of-relations are considered to be ordered pairs of word-forms. In order to take into account the order in which the hypernyms appear in of-relations and reduce the impact of any difference in the volumes of available resources each number in the above frequency matrix is divided by the corresponding row sum of the matrix. This gave rise to the following initial definition of weights:

$$w_{jk}^{\text{cat}} = \frac{M_f^{\text{cat}}[j, k]}{\sum_i M_f^{\text{cat}}[i, k]} \quad (12)$$

Next, the two matrices (poetic versus not poetic) are subtracted by performing an element-wise subtraction of their weights:

$$\forall i, j \in \{1, \ldots, \|\mathcal{S}\|\}, w_{ij}^d = w_{ij}^{\text{non\_poetic}} - w_{ij}^{\text{poetic}}. \quad (13)$$

Thus, $w_{ij}^d \geq 0$, means that h_j follows h_i more frequently in non-poetic relations than in poetic whilst $w_{ij}^d \leq 0$ means the opposite. Having derived the value of each hyper-pair relations from the available data and recorded the frequencies as above, it was obvious that:

1) Within the same path of hypernyms, hypernyms higher in the hierarchy get the weights of the hypernyms lower in the hierarchy,
2) Hyper-pairs hierarchically higher are more common in both poetic and non_poetic relations;
3) As a consequence of the above items, although the weight of hyper-pairs lower in a hierarchy may be ≤ 0 (for example, when hyper-pairs are more frequent in poetic relations), the sign of the weight may change in subsuming hypernyms.
4) Rare hyper-pairs associated to rare relations (for example, rare poetic relations) have very small weights (ref. to equation 12) since their frequency is small.

The IRF measure, inspired by the *Inverse Document Frequency* measure used in Information Retrieval [6] aimed to address these issues.

$$\text{IRF}_{ij} = \frac{\|R\|}{\|\{r : \text{Assoc}((h_i, h_j), r)\}\|} \quad (14)$$

Using the IRF measure, the definition of the weight of each ordered hyper-pair (h_i, h_j) becomes:

$$w_{ij} = M_f^d[i, j] \cdot \mathrm{IRF}_{i,j} \tag{15}$$

where w_{ij} denotes the weight of an ordered hyper-pair: (h_i, h_j).

The weight of each of-relation is then defined as the sum of the weights of all the hyper-pairs associated to it:

$$\mathrm{weight}(r_i) = \sum_{ij} w_{ij} \tag{16}$$

If the weight of an of-relation is positive, then the of-relation is considered to be derived from the non-poetic dataset; otherwise, it is considered to be derived from the poetic dataset.

TABLE 2. Dataset 1.

Dataset 1	
Poetic Resources	
Title	Type
The Complete Works of W. Shakespeare [65] by W. Shakespeare	text, GP[a]
Songs of Innocence and Experience [66] by W. Blake	text, GP
Paradise Lost [67] by J. Milton	text, GP
Byron's Poetical Works, Vol. 1. [68] by Lord Byron	text, GP
Cathay Translations [69] by E. Pound	text, GP
The Wind Among the Reeds [70] by W. B. Yeats	text, GP
The Waste Land [71] by T. S. Eliot	text, GP
Poems in Two Volumes, Vol. 1 [72] by William Wordsworth	text, GP
Sea Garden [73] by H. Doolittle	text, GP
Poems: Three Series, Complete [74] by E. Dickinson	text, GP
Non-poetic Resources	
Brown corpus [75]	pos tagged sentences

[a] GP is used as a shorthand for Gutenberg Project

V. EXPERIMENTS

The algorithms were evaluated using six experiments. The dataset used in each of the Experiments 1–4, was a subset of the dataset used in the next experiment. Experiments 5 and 6 were irrelevant to poetic resources and demonstrate the use of the algorithms in other domains. The datasets used for Experiments 1–4 are included in Tables 2–4.

Although the main interest of this work lies in the extraction of conceptual relations relevant to poetic versus non-poetic resources, the algorithms introduced in this article can be applied in other domains too. Experiment 5, applied the algorithms on the religion and hobbies categories of the Brown corpus. Experiment 6, was contacted on a much bigger sample, by using the first 10377 relations from an Art-history dataset created automatically via a free automatic tool, available at: https://ithaka-labs.s3.amazonaws.com/tdm/v2/datasets, and an equal amount of relations from Reuters relations from categories: gold, earn, and acq. The sample is much bigger than the one used in Experiment 5, and the domains of the datasets being modeled are more diverse. The methods used for the extraction of relations from the resources of each dataset are described in detail in Appendix A.

TABLE 3. Dataset 2.

Dataset 2 = Dataset 1 + Following Poetry resources	
Title	Type
Bay Book of Poems by D. H. Lawrence	text, GP
Aurora Leigh by E. Barrett	text, GP
The Poetical Works of Elizabeth Barrett Vol. 1 by E. Barrett	text, GP
The Poetical Works of Elizabeth Barrett Vol. 2 by E. Barrett	text, GP
The Man Against the Sky by E. A. Robinson	text, GP
Poems by W. Owen	text, GP
Collected Poems by D. Thomas	text, GP
The Collected Works in Verse and Prose Vol. 1 by W. B. Yeats	text, GP
The Collected Works in Verse and Prose Vol. 2 by W. B. Yeats	text, GP
Selected Poems by O. Wilde	text, GP

TABLE 4. Datasets 3, and 4.

Poetry resources of dataset 3 = Poetry resources of dataset 2	
Non-poetic Resources of dataset 3	
Title	Type
Prepositional Phrase Attachment Corpus by A. Ratnaparkhi [76]	text, NLTK [75]
Brown corpus (as for the previous experiments)	pos tagged sentences, NLTK [75]
Dataset 4 = Dataset 3 + Prep. Phrase Attch. Corpus [76]	

Poetic resources resort frequently to imagery [77] and symbolic language to transcend meaning from literal to figurative [17] through vivid sensual perceptions and links between images and ideas that aim to deepen understanding. The actual poetic of-constructions are scarce, since poems use frequently non-poetic of-constructs (for example, part of life) in a wider poetic context. Resorting to imagery and lyrical poetry genres, was made in an attempt to increase the frequency of appearance of of-relations that are specific to these genres (due to vivid contrasting imagery concepts). The overlap of relations between the different types of resources used, influences the results of the experiments since in the case of a large set of overlapping relations, the sets of hyper-pairs specific to each particular resource become smaller. This can be proved formally in the case of Algorithm 1 as will be shown in the next section.

Both algorithms were evaluated for their ability to model the datasets by extracting the associations of concepts underpinning the of-relations of each dataset, and for their ability to predict the dataset from which an of-relation not encountered before is derived. In particular, the experiments contacted for the evaluation of the proposed algorithms aimed to determine:

- Whether algorithm 1 could extract the semantic relations (expressed as hyper-pairs) specific to each type of data resource.

- Whether algorithm 2 could assign a weight to each hyper-pair so that the sum of weights associated to each of-relation can determine the type of resource from which it is extracted.
- Whether the algorithms can be used to predict each type of dataset given an of-relation not encountered before.

In order to determine how well the algorithms model different types of data resources, the algorithms were firstly run on the whole datasets (answer to the first two questions). In order to determine the answer to the third question, the data resources were split into training (80%) and testing sets (20%) for each type of resource. In this case, the hyper-pairs specific to each resource were computed by using the training datasets. The algorithms were then tested by classifying unseen instances derived from the testing data.

VI. RESULTS

The evaluation measures used in this work are the measures of precision, and recall, which are widely used in Information Retrieval (IR). Precision is used in IR to evaluate the prediction of positive results and is defined as the number of correct predictions out of all the predictions made by the system. Recall is the proportion of the number of instances of a positive class that are correctly predicted. Since the current work aims to model two different types of resources (poetic and non poetic in experiments 1-4) in each experiment, the precision and recall is calculated for both types of resources, separately.

Let $\text{CorrectlyPredicted}_t$ be the set of relations of type t whose resource type is correctly predicted. Also, let Predicted_t be the set of relations whose resource type is predicted to be t. Then,

$$\text{Precision}_t = \frac{\|\text{CorrectlyPredicted}_t\|}{\|\text{Predicted}_t\|}. \tag{17}$$

Recall is also calculated for each type of resource separately.

$$\text{Recall}_t = \frac{\|\text{Correctly Predicted}_t\|}{\|R_t\|} \tag{18}$$

where $\text{CorrectlyPredicted}_t = \text{Predicted}_t \cap R_t$.

A. EVALUATION OF ALGORITHM 1 ON THE WHOLE DATA

The set-theoretic approach followed in the definition of Algorithm 1 leads to the derivation of important properties related to the interpretations of results. It is already proved by Proposition 22 that if a hyper-pair associated to an of-relation r is in S''_p (the set of hyper-pairs used to determine poetic resources) then r is extracted from a poetic resource. The same applies for all hyper-pairs in S''_{np}. This property holds only when S''_p (S''_{np}) is determined by considering the whole dataset (that is, when all extracted relations are considered). Similarly, if a hyper-pair associated to an of-relation r is in S''_{np}, then r is extracted from a non-poetic resource. Thus,

$$\text{Predicted}_t = \{r : \exists \text{hp} \in S''_t \wedge \text{Assoc}(\text{hp}, r)\}$$

TABLE 5. Alg. 1, Evaluation results for experiments 1–6 based on the whole datasets.

	Exp. 1		Exp. 2	
	Poetic	Non-Poetic	Poetic	Non-poetic
Precision	1.0	1.0	1.0	1.0
Recall	89.7%	92.2%	89.8%	91.1%
	Exp. 3		Exp. 4	
	Poetic	Non-Poetic	Poetic	Non-poetic
Precision	1.0	1.0	1.0	1.0
Recall	89.5%	90.8%	89%	90%
	Exp. 5		Exp. 6	
	Brown (Religion)	Brown (Hobbies)	Brown (Religion)	Brown (Hobbies)
Precision	1	1	1	1
Recall	95.1%	95.7%	91.3%	85.7%

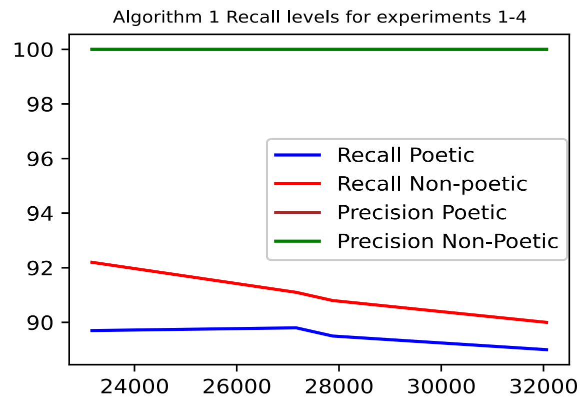


FIGURE 1. Precision and Recall levels for experiments 1-4.

where t denotes a type of a resource. Also, by Proposition 23, any hyper-pair associated to a poetic (non-poetic) relation, cannot be a member of S''_{np} (S''_p).

It follows that $\text{Predicted}_t = \text{CorrectlyPredicted}_t$ when Algorithm 1 is applied on the whole data, proving that in this case precision will always be equal to 1. This property is confirmed by the results shown in Table 5. Fig. 1 illustrates the recall and precision levels for Algorithm 1 on poetic data resources used in experiments 1-4. The fall in recall levels is partly due to an increase in the number of of-relations common to poetic and non-poetic resources and, or an increase in the number of of-relations whose associated hyper-pairs are subsumed by hyper-pairs in S''_p or S''_{np} or S_c (and are therefore removed during the construction of these sets by the definition of Algorithm 1). The latter include relations containing words which are not in the lexical entries of PWN (for the synset assigned to these words please refer to Appendix A), since each of these words is mapped to the most generic synset: entity.n.01 or to a NER entity type. The high recall levels in Experiment 5 where the Algorithm is tested on different domains of relations show that the set of (overlapping) relations that appear in both domains is small. Experiment 6 uses a larger sample, leading to slightly smaller recall values.

The relationship between recall levels and the degree of overlap between the relations extracted from each type of resource, can be formally proved for Algorithm 1 (whole dataset).

Proposition 24: Let $O = R_{\text{poetic}} \cap R_{\text{non-poetic}}$, denote the overlap of relations between the two domains. Then, as the size of O increases, the value of Recall decreases, i.e. recall $\sim \frac{1}{\|O\|}$.

The proof of Proposition 39 can be traced in Appendix B.

Proposition 39 formalizes the relationship between recall and overlapping sets of relations extracted from different types of resources (mainly poetic and non-poetic in this work). The next proposition formalizes the relation between different datasets. It is very important in that it points to the parameters that influence the differences in recall levels between different experiments and datasets, and help to focus interpretation on the causes of recall levels variations.

Proposition 25: Consider two datasets D_1 and D_2 say, such that $D_1 \subseteq D_2$. Let R_{D_1} and R_{D_2} be the sets of relations extracted from D_1 and D_2 respectively, R_{cat,D_i} be the set of relations extracted from the type cat resources in D_i and $O_{D_i} = R_{\text{cat},D_i} \cap R_{\overline{\text{cat}},D_i}$ where cat is one of two categories of resources (for example poetic) and $\overline{\text{cat}}$ is the alternative category (for example non-poetic). If $\text{recall}_{\text{cat},D_i}$ is the recall of a category of relations in $D_i \in \{D_1, D_2\}$, then $\text{recall}_{\text{cat},D_1} > \text{recall}_{\text{cat},D_2}$ implies $O_{D_i} \subset O_{D_{i+1}}$

The proof of Proposition 40 is provided in Appendix B.

Proposition 40 states that for any two datasets D_1 and D_2 where the of-relations in D_1 are a subset of the of-relations in D_2, if the Recall level of Algorithm 1 for D_1 is higher than the recall level for D_2, then this implies that the set of relations common to both types of resources (referred to as overlapping relations) in D_2 is a superset of the corresponding set of overlapping relations in D_1. This result applies to experiments 1–4 and shows a possible explanation of the variation in recall levels between the experiments 1–4. Another possible explanation, is the existence of an increased value of unknown words (words not in the PWN) in relations.

Considering the above formally proved properties, the results of each experiment were further analyzed in order to explain the (small) deviations of recall values between experiments 1–4 (although recall levels are still high in the case where the whole datasets are modeled). Further analysis of results for Experiments 1–4 are included in Tables 6 and 7, below. For clarity, the analysis of the remaining experiments is included in Appendix C.

TABLE 6. Algorithm 1 on datasets of experiments 1, and 2.

	Experiment 1		Experiment 2	
No. of Relations whose hpairs are a subset of:	Poetic	Non-Poetic	Poetic	Non-Poetic
S''_p	11624	0	15235	0
$S''_{\text{np}} \cap S''_p$	0	0	0	0
S''_{np}	0	9492	0	9375
S_c	175	157	222	191
Removed hpairs	1161	641	1517	724
Total	12960	10290	16974	10290

TABLE 7. Algorithm 1 on datasets of experiments 3, and 4.

	Experiment 3		Experiment 4	
No. of Relations whose hpairs are a subset of:	Poetic	Non-Poetic	Poetic	Non-Poetic
S''_p	15199	0	15070	0
$S''_{\text{np}} \cap S''_p$	0	0	0	0
S''_{np}	0	9998	0	13696
S_c	228	197	244	219
Removed hpairs	1547	810	1660	1283
Total	16974	11005	16974	15198

Tables 6, 7 show:

- The number of relations whose set of associated hyper-pairs are elements of S''_p. By proposition 22 these relations are extracted from poetic resources.
- The number of relations whose set of associated hyper-pairs are in S''_{np}. By proposition 22 these relations are extracted from non-poetic resources.
- The number of relations whose set of associated hyper-pairs is a subset of S_c. These relations cannot be classified.
- The number of relations whose set of associated hyper-pairs have been deleted during the construction of the sets S''_p, S''_{np} and S_c (8)–(10). These hyper-pairs are referred to as *Removed hpairs*.

Obviously, by the construction of the sets S''_p and S''_{np}, it follows that $S_c \cap S''_p = \emptyset$, $S_c \cap S''_{\text{np}} = \emptyset$ and $S''_p \cap S''_{\text{np}} = \emptyset$ in the case where the whole data is considered. Example 26 shows a situation where all hyper-pairs associated to an of-relation are removed before the extraction of the final sets of S''_p, S''_{np}, and S_c.

Example 26: Let $r =$ (love, detail) be the relation corresponding to the English preposition 'love of detail', extracted from the non-poetic resources. All hyper-pairs associated to r are a subset of *Removed hpairs*. The reason can be traced to the steps followed in the construction of sets S''_p and S''_{np} in Algorithm 1.

All hyper-pairs associated to (love, detail) were initially in $S_c = S_p \cap S_{\text{np}}$. Examples, are: [(love.n.01, detail.n.01), (state.n.02, detail.n.01), (feeling.n.01, detail.n.01)].

These hyper-pairs were firstly removed from both S_p and S_{np} as they belonged to S_c (5). But, since they were subsumed by more generic hyper-pairs in S_p, they were also subsequently removed from S_c (6). Thus, they are not included the sets: S''_p, S''_{np}, S_c. These hyper-pairs are placed in the set *Removed hpairs*.

B. EVALUATION OF ALGORITHM 1 FOR PREDICTION

The sets of relations extracted from each type of resource are split into a training set (80%) and a testing set (20%), each. Let $S''_{p,\text{tr}}$, and $S''_{\text{np, tr}}$ be the sets of hyper-pairs associated to the training relations extracted from the poetic and non-poetic

TABLE 8. Results of Algorithm 1 for Experiment 1 under Method 1 and Method 2.

	Exp. 1 Method 1		Exp. 1 Method 2	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	71.8%	66.3%	84.4%	83.2%
Recall	80%	78%	56.9%	52%

resources, respectively. Then the same properties apply as before for these sets, for example: $S''_{p,\text{tr}} \cap S''_{\text{np, tr}} = \emptyset$. The problem in this case, is that some hyper-pairs associated to an of-relation in the testing set may be included in $S''_{p,\text{tr}}$ and some in $S''_{\text{np, tr}}$

Example 27: Consider the poetic relation (touch, joy) corresponding to the preposition 'touch of joy' which is included in the testing set. Examples of hyper-pairs associated to this relation are: (contact.n.04, feeling.n.01), (happening.n.01, joy.n.01). The problem is that the hyper-pair (contact.n.04, feeling.n.01) is included in the training set of hyper pairs associated only to poetic resources, whereas the hyper-pair (happening.n.01, joy.n.01) is in the training set of hyper-pairs associated only to the non-poetic resources, leading to the conclusion that the relation is extracted from both a poetic and a non-poetic dataset. This problem appears only in the Prediction of 'unseen' relations that were not encountered in the training set.

In this case, using the same formulas for the calculation of precision and recall as before (when evaluating on the whole data) would inflate the results of Algorithm 1 considerably.

Let $S''_{p,tr}$ represent the set of hyper-pairs associated only to the poetic resources of the training set, and $S''_{\text{np,tr}}$ represent the set of hyper-pairs associated only to the non-poetic resources of the training set.

Possible methods of evaluation of the performance of Algorithm 1:

- **Method 1.** Predict the type of poetic (non-poetic) resources as in the case where the whole data is considered. This means that if there is a hyper-pair in $S''_{p,tr}$ (or $S''_{\text{np,tr}}$) associated to an of-relation in the testing poetic (or non-poetic) set, then the type of resource is predicted to by poetic (or non-poetic).
- **Method 2.** Only if all the hyper-pairs associated to a testing relation are in $S''_{p,\text{tr}}$ ($S''_{\text{np,tr}}$) the resource is predicted to be poetic.

Method 1 leads to overstated recall levels (relations from either type of resources are predicted to belong to both types of resources) and understated precision levels (many relations considered to belong to one type of resource actually belong to the other). Also, method 1 is ambiguous in the sense that it can lead to two different predictions for a single of-relation. Method 2 leads to higher precision levels but lower recall results. In order to avoid ambiguity, the adopted method for the evaluation of Algorithm 1 for prediction, is Method 2. Table 8 shows the difference in results under each method in Experiment 1. Comparisons for the remaining experiments can be traced in Appendix C. The evaluation results for Algorithm 1 under Method 2, are summarized in Table 9.

TABLE 9. Algorithm 1 metrics for Experiments 1 – 6 using the testing relations for the Prediction of their data resources.

	Exp. 1		Exp. 2	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	84.4%	83.2%	85.9%	81.9%
Recall	56.9%	52%	60.4%	47.6%
	Exp. 3		Exp. 4	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	84.8%	82.8%	84.2%	85.4%
Recall	59.5%	47.8%	54.6%	54.3%
	Exp. 5		Exp. 6	
	Brown Religion	Brown Hobbies	Art History	Reuters
Precision	78.2%	83%	93.1%	91.4%
Recall	39.7%	56%	69.9%	61.7%

The results of Algorithm 1 in prediction show a decline in the values of Recall in all experiments when the whole datasets are modeled. In Experiment 2 where more poetic resources are considered, the value of the Recall increases in the case of the Poetic type of resources. In Experiment 6, where the content of resources is more diverse, the recall levels are increased for both Poetic and Non-poetic resources.

C. EVALUATION OF ALGORITHM 2 ON THE WHOLE DATA

The algorithm is based on the assumption that each hyper-pair contributes differently to the identification of the type of resources from which an of-relation is extracted. For each type of resource, the weight of each hyper-pair is determined partly by a relative frequency measure (12), and partly by the proposed IRF measure.

Frequently occurring and very generic hyper-pairs that appear in poetic and non-poetic resources have smaller weights than to hyper-pairs associated to one particular type of resource and are less generic, due to the IRF measure.

Fig. 2, and Fig. 3 show that the IRF measure improves the performance of Algorithm 2 for both the poetic and non-poetic datasets.

Due to the method of calculation of the weight of each relation extracted from type of resources, the Predicted_t and $\text{CorrectlyPredicted}_t$ sets used in the calculation of precision and recall measures in this case, take the form:

$$\text{Predicted}_t = \{r : \text{weight}(r) \geq 0\}$$

where weight(r) as in (16).

$$\text{CorrectlyPredicted}_t = \text{Predicted}_t \cap R_t$$

where R_t is the set of relations of type t. A summary of the results for Algorithm 2 for the experiments 1–6 is displayed in Tables 10 and 11.

D. EVALUATION OF ALGORITHM 2 FOR PREDICTION

The same sets of training (80%) and testing (20%) relations as the ones used in the evaluation of Algorithm 1, are used for the

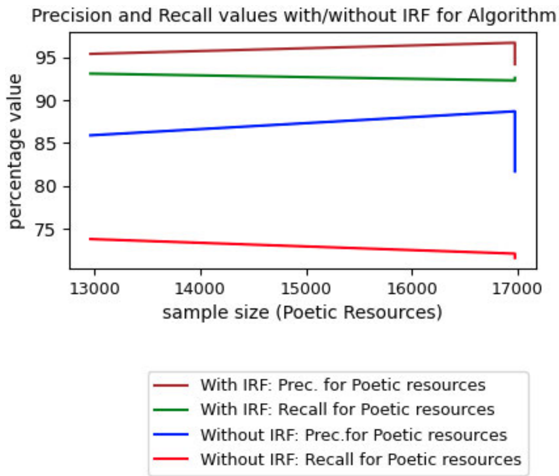


FIGURE 2. Precision-recall levels for experiments 1-4 with/with IRF for poetic resources.

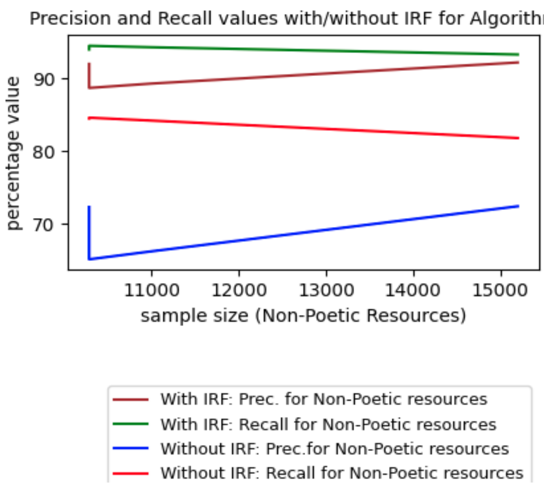


FIGURE 3. Precision-recall levels for experiments 1-4 with/with IRF for non-poetic resources.

TABLE 10. Algorithm 2 results on the whole datasets.

	Experiment 1		Experiment 2	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	95.4%	92%	96.7%	88.7%
Recall	93.1%	94%	92.3%	94.5%
	Experiment 3		Experiment 4	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	96.4%	89.3%	94.2%	92.2%
Recall	92.3%	94.3%	92.6%	93.3%

TABLE 11. Algorithm 2 results on the whole datasets.

	Experiment 5		Experiment 6	
	Brown Religion	Brown Hobbies	Art-History	Reuters
Precision	94.1%	98.8%	92.6%	92.5%
Recall	97.9%	96.1%	93.5%	90.6%

evaluation in Algorithm 2. In this case, the weights of hyper-pairs are determined by using only the training resources. Once the weights of hyper-pairs associated to the relations of the training sets are determined, they are used to predict the type of resources of the testing relations. The results are summarized in Table 12, below. A detailed analysis can be traced in Tables 20 – 23 in Appendix C.

TABLE 12. Algorithm 2 results for Experiments 1–4 using testing relations for the prediction of their resources.

	Experiment 1		Experiment 2	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	72.7%	64.1%	77.4%	59.1%
Recall	68.2%	67.4%	70.6%	65.8%
	Experiment 3		Experiment 4	
	Poetic	Non-Poetic	Poetic	Non-Poetic
Precision	76.7%	61%	70.8%	66.5%
Recall	68.2%	67.4%	67.3%	68.7%

E. COMPARISON OF ALGORTHMS 1 AND 2

Algorithm 1 follows a clearly defined set-theoretic approach that enables the proof of properties relevant to the interpretation of results. Algorithm 2 recognizes the fact that not all hyper-pairs are equally important to each type of resource and proposes a weighting scheme in order to assign a weight to each of-relation determining whether the particular relation is more relevant to a particular type of resource.

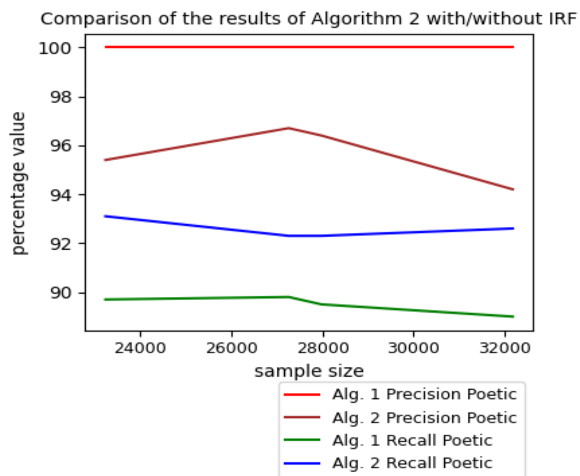


FIGURE 4. Precision and recall levels for experiments 1-4.

Both Algorithms perform well when modeling the whole datasets. Fig. 4 illustrates the performance of Algorithms 1 and 2 when considering the extraction of hyper-pairs from poetic resources. As discussed, the value of the precision of Algorithm 1 is one, in all experiments where the whole data is considered. This implies that the set of hyper-pairs extracted from poetic resources can predict correctly the type of resources from which the of-relations were extracted, and can be used to model their domain. However, Algorithm 1 produces the lowest recall levels of the two algorithms, which is primarily due to an increased number of hyper-pairs in 'Removed hpairs' and an increase in the size of overlapping relations between the different types of resources as the experiments' datasets are expanded. Algorithm 2 reaches its highest precision level in Experiment 2 when the largest set of poetic relations was considered. The overall recall levels for poetic resources are higher that 92% although with a declining slope. A further analysis has shown that about 11% of concepts participating in relations extracted from poetic resources were not mapped to any lexical entry in PWN. The performance of both algorithms declines when the algorithms are used for prediction. In this case too, Algorithm 1 has lower recall levels than Algorithm 2. As shown in Tables 20 – 23 (Appendix C) this is primarily due to the fact that hyper-pairs associated to each type of relations in the testing set, overlap with hyper-pairs associated to both types of relations, in the training set. However, the value of Precision is high showing that the training sets $S''_{p,\mathrm{tr}}$ and $S''_{\mathrm{np,tr}}$ can correctly classify the resource of each testing relation more than 80% of the time. Algorithm 2 demonstrates higher recall levels although the precision levels are significantly lower than the Precision levels of Algorithm 1. Due to the measure of frequency used in Algorithm 2, the weight of each relation in the testing set is dependent on whether the hyper-pairs associated to this relation, were encountered in the training set of resources. It is also dependent on whether or not the lexical terms participating in each relation are in the vocabulary of PWN.

F. IMPORTANT PROPERTIES OF IRF

The IRF measure possesses important properties which make it an important measure on its own that extends beyond its use in the determination of weights of hyper-pairs in Algorithm 2. The IRF measure can be generalized to apply to any representation of hierarchically organized relations.

One important property of the IRF measure is that two hyper-pairs have the same IRF value when they are associated to the same number of relations. When the hyper-pairs also appear in the same subsumption hierarchy, then it follows that they are associated to exactly the same set of relations.

Proposition 28: For any two hyper-pairs hp_1, and hp_2, such that $\mathrm{hp}_1 \sqsubseteq \mathrm{hp}_2$, if $\mathrm{IRF}_{\mathrm{hp}_1} = \mathrm{IRF}_{\mathrm{hp}_2}$, then: $\{r : \mathrm{Assoc}(\mathrm{hp}_1, r)\} = \{r : \mathrm{Assoc}(\mathrm{hp}_2, r)\}$.

Since hyper-pair relations are used to represent the conceptual relations underpinning a domain, then the IRF measure can provide information that is useful for inference.

Example 29: Let $r =$(sea, glory) be an of-relation representing the phrase 'sea of glory'. The hyper-pairs associated to this relation form a subsumption hierarchy. For example,the following sub-paths of hyper-pairs:

sea.n.01 @→ body_of_water.n.01 @→ thing.n.12 and glory.n.01 @→ honor.n.02 @→ standing.n.01 lead to the following subsumption relations between hyper-pairs associated to r:

(sea.n.01, glory.n.01) $\sqsubseteq$ (sea.n.01, honor.n.02) $\sqsubseteq$ (body_of_water.n.01, honor.n.02), and (sea.n.01,glory.n.01) $\sqsubseteq$ (sea.n.01, honor.n.02) $\sqsubseteq$ (sea.n.01, standing.n.01).

Now, the IRF value of each of the following hyper-pairs is the same:

(body_of_water.n.01, honor.n.02),
(sea.n.01, honor.n.02),
(sea.n.01, glory.n.01),
(sea.n.01, standing.n.01)

Since (sea.n.01, glory.n.01) and (sea.n.01, honor.n.02) appear in the same subsumption hierarchies with the rest, and their IRF value is the same, then all of the above hyper-pairs are associated to the same relations. It is then possible to make inferences of the form: a sea of glory is a sea of honor, and all examples of a sea of glory are examples of a sea of honor.

$$\begin{aligned}\{r &: \mathrm{Assoc}((\text{sea.n.01, glory.n.01}),\mathrm{r})\}\\ &= \{r : \mathrm{Assoc}((\text{sea.n.01, honor.n.02}),\mathrm{r})\}\\ &= \{r : \mathrm{Assoc}((\text{body\_of\_water.n.01, honor.n.02}),\mathrm{r})\}\\ &= \{r : \mathrm{Assoc}((\text{sea.n.01, standing.n.01}),\mathrm{r})\}\end{aligned}$$

Generally, it can be proved that if a hyper-pair hp_1 is more specific that a hyper pair hp_2, then the set of relations to which hp_1 is associated is a subset of the set of relations to which hp_2 is associated.

Proposition 30: Assume any two hyper-pairs (h_i, h_j) and (h_k, h_m) such that $(h_i, h_j) \sqsubseteq (h_k, h_m)$. Then, it follows that:

$\{r : \mathrm{Assoc}((h_i, h_j), r)\} \subseteq \{r : \mathrm{Assoc}((h_k, h_m), r)\}$

The proof follows easily from first principles. Intuitively, if (h_i, h_j) is associated to a relation r, then all its subsuming pairs are also associated to r, by definition.

Also, by the definition of the *Inverse Relation Frequency* (IRF) (see equation 14), it is straight forward to show that:

Proposition 31: If (h_i, h_j) and (h_k, h_m) are two hyper-pairs associated via the subsumption relation $(h_i, h_j) \sqsubseteq (h_k, h_m)$, then $\mathrm{IRF}_{i,j} \geqq \mathrm{IRF}_{k,m}$

The proof of proposition 31 follows easily from Proposition 30 and the definition of IRF. It is also trivial to show that:

Proposition 32: For any two hyper pairs (h_i, h_j) and (h_k, h_m) such that $(h_i, h_j) \sqsubseteq (h_k, h_m)$, if $\mathrm{IRF}_{i,j} = \mathrm{IRF}_{k,m}$, then $\{r : \mathrm{Assoc}((h_i, h_j), r)\} = \{r : \mathrm{Assoc}((h_k, h_m), r)\}$

The proof of proposition 30 follows easily from the definition of IRF and the fact that the set of relations associated to a hyper-pair are included in the set of relations associated to its subsuming hyper-pairs (proposition 30).

VII. ONTOLOGICAL CONSIDERATIONS AND PWN

Approaches of mapping PWN lexical entries to Ontologies, like, for example, the higher ontology SUMO [26] discussed in Section II, and KYOTO [78] for allowing language independent reasoning over multiple domain wordnets interlinked to a shared ontology, have been cited in literature. Complementary to these approaches, are the more recent advancements in web technologies that enable the representation of rich linguistic information, for example the LexInfo ontology-lexicon model [60] and Lemon [61], [62]. Of the most prominent of these approaches is Lemon, which enables the publication and linking of lexical networks/WordNets to ontologies and to the Linked open Data Cloud [79].

Lemon [61], [62], is a model of ontology-lexica, a standard for sharing lexical information on the semantic web [80] that enables the mapping of Ontologies to lexical terms, supporting conceptual interoperability.

Of-constructions can provide further knowledge about the semantic model of texts, either as single concepts in an ontology, or as semantic roles. For example the phrase *Director of Utilities* as a whole, may be represented as the subclass *DirectorOfUnitilies* of the class *Director*, or as the semantic relation *director_of* between the concepts *Human* and *Organization*, or as a sub_property of another ontological property. An example is the property *occupiesPosition* in the SUMO ontology [26], [81], where the classes involved are already mapped to the PWN entries:

```
(instance occupiesPosition TernaryPredicate)
(domain occupiesPosition 1 Human)
(domain occupiesPosition 2 Position)
(domain occupiesPosition 3 Organization)
(documentation occupiesPosition
(&%occupiesPosition, ?PERSON ?POSITION
?ORG)
```

Due to the availability of the above technologies and the fact that the semantic relations of the data can be expressed via hyper-pairs, it is possible to extract and represent ontological knowledge from the derived hyper-pair relations. In particular, once the main concepts are extracted from text, ontologies can be created via the mapping of hyper-pair relations extracted from text and the hierarchical relations derived from PWN taxonomic relations. However, creating ontologies by applying the entirety of hyper-pairs that can be extracted from text and PWN is costly.

1) CREATION OF A KINSHIP ONTOLOGY FROM THE CORPUS USING PWN HIERARCHICAL RELATIONSHIPS

The aim of this Section is to illustrate how the use of hyper-pair subsumption relations can potentially enrich Ontology representation and inference. The discussion is motivated by building a toy kinship ontology example, based purely on PWN taxonomic relations.

The toy ontology example maps synsets directly to ontology concepts; this is not provided as the suggested method for extracting ontologies, but only in order to demonstrate how the subsumption relations between concept pairs extracted from phrases, can enrich ontological reasoning.

For the kinship ontology being created, the hypernym relative.n.01 is considered as the most generic hypernym associated to the first lexical term of any kinship of-relation, since more generic hypernyms are not specific to the kinship relationship. Since the example refers to human kinship relations, the most generic hypernym for the second lexical term participating in an of-relation is considered to be person.n.01.

The uppermost property modeled, is the property *relative_of*. Sub-ordinate synsets of the synset relative.n.01 are used to determine sub-properties of *relative_of*.

Algorithm 1 Property and Class Hierarchies From PWN Hierarchy of Entities for $r = (N_1, N_2)$

Require: listOfClasses, listOfProperties

sense0 := sense extracted for N_1
sense1 := sense extracted for N_2
hpaths1 := hypernym_paths(sense0)
hpaths2 := hypernym_paths(sense1)
bigrams1, bigrams2 = [], []
for $p \in$ hpaths1 **do**
    bigrams1 := bigrams1 + get_bigrams(p, relative.n.01)
end for
for $q \in$ hpaths2 **do**
    bigrams2 := bigrams2 + get_bigrams(q, person.n.01)
end for
allBigrams ⇐ bigrams1 + bigrams2
for $(s1, s2) \in$ allBigrams **do**
    prop1_name ⇐ string(s1+_of)
    prop2_name ⇐ string(s2 + _of)
    newPropRel = (property(prop1_name), property(prop2_name))
    ObjectPropertyHier := ObjectPropertyHier + [newPropRel]
    newClassRel = (class(s1), class(s2))
    classHier := clasHier + (newClassRel)
end for

Function 2 get_bigrams

Require: path, upper_synset

biagrams:=[]
if upper_synset $\notin$ path **then**
    return ([])
end if
index := p2.position(upper_synset);
for (i, s) $\in$ enum(path[: index]) **do**
    bigrams := bigrams + [(path[i].name(), path[$i+1$].name())]
end for
return (bigrams)

The main steps for the creation of the basic classes and properties of the toy kinship ontology are outlined below via Algorithm 1. The algorithm simply makes use of the PWN concept hierarchies in order to create the properties and classes of the toy ontology. The notion of subsumption hierarchy between relations, is not considered yet in the construction of the toy ontology. It will be considered later in order to show the additional benefits of using subsumption hierarchies between concepts.

For each of-relation in the kinship ontology example being created, the above algorithm firstly extracts the synsets associated to each lexical term in the relation. Then, the algorithm searches for the paths from each synset to the hypernyms relative.n.01 and person.n.01, respectively. If they exist, the algorithm creates the property relative_of and all its relevant sub-properties, by reference to the sub-ordinate synsets of relative.n.01 in PWN. The following example illustrates the construction of object properties and classes extracted from one instance of a kinship relation between lexical terms, by using the steps outlined in Algorithm 1.

Example 33: The of-preposition being considered is: 'the father of Mary'. The senses selected for each of the words 'father' and 'Mary', are father.n.01, and person.n.01, respectively. 'Mary' is assigned the sense person.n.01 after being recognized as a type PERSON lexical entry (using spaCy [82], as described in Appendix A. Due to space limitations, only the first hypernym path of the synset father.n.01 is considered here: [father.n.01, parent.n.01, genitor.n.01, progenitor.n.01, ancestor.n.01, relative.n.01].

As the hypernym relative.n.01 is considered to be the most generic hypernym of the synset father.n.01 relevant to the kinship relation, more generic hypernyms are excluded from consideration. Similarly, since the most generic hypernym for the second lexical term participating in the of-relation is person.n.01, more generic terms are eliminated. Traversing the path from the most specific hypernym to the most generic hypernym and taking bigrams, the following list of bigrams is produced, where each bigram represents a subordinate-superordinate relation between the relevant hypernyms:

[(father.n.01, parent.n.01),
(parent.n.01, genitor.n.01),
(genitor.n.01, progenitor.n.01),
(progenitor.n.01, ancestor.n.01),
(ancestor.n.01, relative.n.01)].

Then, the algorithm converts bigrams to concept hierarchical relations and to object property subsumption relations. In this example, the relevant object property subsumption relationships are:

father_of $\sqsubseteq$ parent_of,
parent_of $\sqsubseteq$ genitor_of,
genitor_of $\sqsubseteq$ progenitor_of,
progenitor_of $\sqsubseteq$ ancestor_of,
ancestor_of $\sqsubseteq$ relative_of.

'Mary', is just an instance of the concept 'person', which is a class in the taxonomic relations created. Similarly, the concepts: father, parent, genitor, progenitor, ancestor, and relative are all added to the class hierarchy.

Diagrams 5, and 6 (created by the Protege editor) illustrate the concept and hierarchical object properties extracted from all of-relations in the non-poetic dataset by using the PWN hypernym hierarchical relations, as discussed above.

2) RELEVANCE OF HYPER-PAIR SUBSUMPTION TO THE KINSHIP ONTOLOGY

The taxonomic relations between concepts and properties created by Algorithm 1, can be used to infer information about objects in the universe of discourse. For example, it can be used to represent information like, 'Steve is a father' (by treating 'Steve' as an instance of the concept 'Father'), it can lead to the conclusion that 'Steve' is a parent (via the subordinate-superordinate relation: Father $\sqsubseteq$ Parent), and similarly it can be used to represent the information that

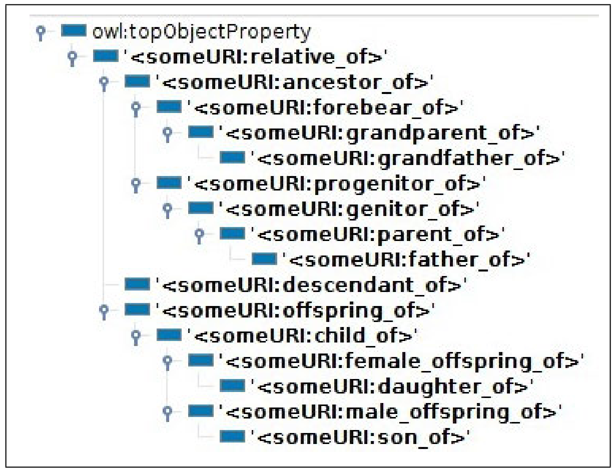


FIGURE 5. Example 33 Obj. Property hierarchy.

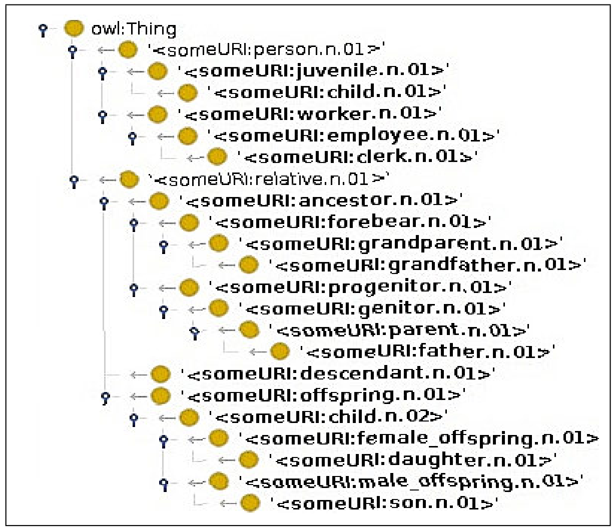


FIGURE 6. Example 33 class hierarchy.

'Steve is the father of Mary', from which it can be inferred that 'Steve is a parent of Mary'. To this extend, the subsumption hierarchies between hyper-pairs would have nothing more to offer in terms of the semantics.

Now consider the phrase 'father of a schoolboy'. This would be mapped to the of-relation: (father,schoolboy). As shown in Example 33, this of-relation can give rise to a number of object-relationships using both lexical terms, like for example, the object properties: 'father_of', and 'parent_of', and hierarchies of concepts like: schoolchild $\sqsubseteq$ young_child.

As each hyper-pair in the hyper-pair subsumption hierarchy is associated to a set of of-constructs, for example 'father of schoolboy', it can potentially lead to inferences like, for example: 'father of schoolboy', 'father of schoolchild', 'father of young_child', and 'parent of schoolboy', (the list is not exhaustive), by virtue of the fact that the relation (parent, schoolboy) is associated to the hyper-pair (parent.n.01, schoolboy.n.01) which is subsumed by the hyper-pair (parent.n.01, schoolchild.n.01).

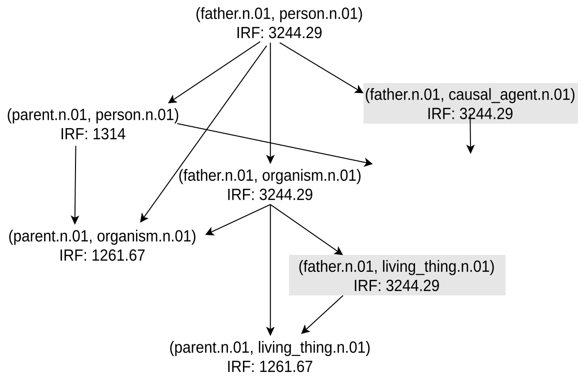


FIGURE 7. A fragment of the Kinship relation.

So, hyper-pairs are associated to sets of phrases (of-constructs) that are grouped together according to the interpretation of their lexical entries, and can be used to draw inferences from of-prepositional phrases via the subsumption hierarchies between hyper-pairs.

A fragment of the subsumption hierarchy of the hyper-pair: (father.n.01, person.n.01), with the relevant IRF values, is illustrated in Fig. 7. Some of the hyper-pairs in Fig. 7, for example the hyper-pair (genitor.n.01, living_thing.n.01), are not specific to the human kinship of-relation (father, person). As stated already, the diagram illustrates just a fragment of the subsumption hierarchy of the hyper-pair (father.n.01, person.n.01). The complete hierarchy includes more generic hyper-pairs which are not specific to a human kinship relation, fo example (causal_agent.n.01, object.n.01).

In this example, the most generic hyper-pair associated to a kinship of-relation is considered to be the hyper-pair (relative.n.01, person.n.01). For this reason, a hyper-pair subsumption hierarchy for kinship relations can exclude hyper-pairs which are more generic than the hyper-pair (relative.n.01, person.n.01). For example, the hyper-pair (entity.n.01, person) would be associated to of-constructs not necessarily expressing a kinship relation. Similar considerations apply to every hyper-pair hierarchy.

The IRF values on the diagram, play an important role too, in modeling the domain, as they show which hyper-pairs on the same hierarchy are associated to exactly the same of-relations in the dataset. By proposition 32 in Section VI, it follows that two hyper-pairs in the same path have the same IRF value if (and only if) they are associated to the same set of of-relations. For example, in diagram 7, the hyper-pairs: (father.n.01, person.n.01), (father.n.01, organism.n.01), and (father.n.01, living_thing.n.01) are associated to exactly the same relations. In an ontology modeling a domain where all objects in ‘living_thing’ and all objects in ‘organism’ are ‘human’, then the hyper-pairs (father.n.01, organism.n.01), and (father.n.01, living_thing.n.01) would provide redundant information.

VIII. CONCLUSION

This paper addressed the problem of extracting conceptual relations from different types of resources and of using the extracted associations between concepts to determine the type of resource from which they are derived.

Algorithm 1 performed well in identifying the conceptual relations specific to each resource. Algorithm 2 calculated successfully the weight of each hyper-pair depending on its presence in each type of resource.

The performance of the algorithms in identifying the associations between concepts that are specific to each type of resource provide an insight as to how diverse the domains of the different types of resources, are. For example, the weights of hyper-pairs when the resources being compared are identical, will be zero in Algorithm 2, and the set of hyper-pairs specific to each resource (for example S''_p) will be empty in Algorithm 1.

The lack of lexical entries in PWN of words (especially in the poetic corpus) participating in of-relations, influences the results of the proposed algorithms. To address the lack of entries in the PWN vocabulary, it is worth considering the definition of resource vocabularies, with mappings to synsets in PWN. For example, in poetry, a local vocabulary might include information about imaginary legends and mythical persons, which would then be linked to the more general PWN sense imaginary_being.n.01. Work aiming at publishing European Poetry Data, thus addressing the contextual interpretation of concepts via Linked Open Data, is cited in [83]. Recent research in the extraction of domain specific lexicons [84] focuses primarily on sentiment analysis.

During the processing and extraction of of-relations from datasets, a hierarchy of noun terms was extracted for each multi-noun expression in each of-relation (Appendix A). In addition, NER mappings were considered in those cases where none of the noun expressions could be mapped to a lexical entry in PWN database, for example proper nouns. However, in poetic resources, the adjectives preceding nouns play an important role in transforming the meaning of words. For example, the noun word ‘years’ cannot express adequately the phrase ‘sunny years’. Also, the morphological aspects of words are important in determining the type of resources from which phrases are extracted. The current work did not address the multinoun interpretation problem, neither the morphological characteristics of the noun expressions participating in of-relations. Future work will need to address the representation of multinoun expressions and the impact of the morphological characteristics of each noun expression in determining the type of resources from which relations are extracted.

APPENDIX A
PREPARATION OF DATA

The extraction of of-relations involved a multi-step process including a variety of tools due to the diversity of the formats in which data was made available. The main processing steps used for the extraction of of_constructs are described in this section.

Each of the poetic texts retrieved from the Gutenberg project for the extraction of poetic of-relations included a preface, bibliographical notes, remarks, contents,

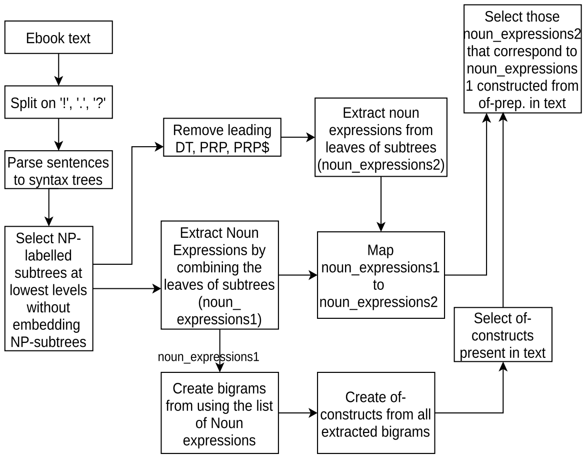


FIGURE 8. NLP for poetic texts.

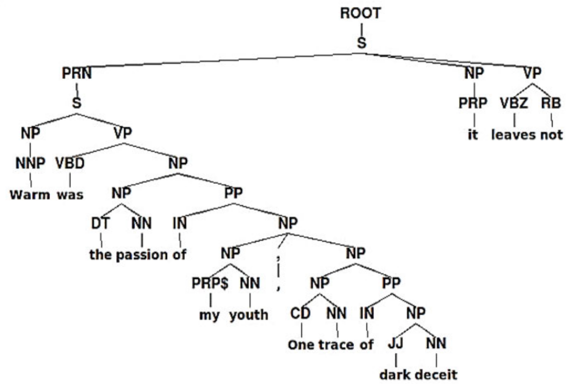


FIGURE 9. Example 34 syntax tree.

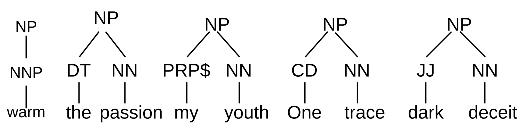


FIGURE 10. Example 34 lowest NP subtrees.

introductions to poem collections, postscripts, and letters that were removed manually. Regular expressions were used to search and remove footnotes, clean spaces and newlines. Since the formats lacked consistency, any footnotes not captured automatically, were removed manually.

Fig. 8 depicts the automatic natural language processing steps followed for the extraction of of-relations from texts.

1) Each poem is split into a list of sentences, using the characters {'!','.',';','?'}.
2) Each sentence is parsed into a tree structure by using the Stanford Parser [85].
3) The lowest Noun Phrase (NP) subtrees of the tree structures produced in step 2, i.e. the ones that do not have NP subtrees, are selected.
4) The NP-subtrees created in step 3 are further processed in order to:
    a) Remove any leading nodes that are not part of the noun expressions, like articles (for example, the word 'the'),
    b) Create noun expressions by combining the leaves of the NP-subtrees resulting from 4(a),
    c) Create also the noun expressions represented by the original NP-subtrees as in step 3. The resulting noun expressions constitute noun segments of the original sentences.
    d) Map the noun expressions obtained in steps 4(b), and 4(c), respectively (so that they can be mapped to the terms of of-relations at the end).
5) The noun expressions obtained in 4(c) are then used to create a list of bigrams. As the noun expressions of sentences, are linked either via a preposition or a verb, the bigrams of the noun expressions extracted in (3) contain the bigrams of the parts of 'of' prepositions included in the original sentences.
6) The bigrams collected in (5) are joined via the 'of' preposition and the resulting prepositional phrases are checked to determine whether they are substrings of the sentence being considered.

Example 34: Let us consider Byron's phrase:

Warm was the passion of my youth, One trace of dark deceit it leaves not. [68].

- The syntax tree of Byron's phrase is depicted in Fig. 9, below.
- The lowest NP subtrees are depicted in Fig. 10.
- The noun expressions created by combining the leaves of the subtrees (Fig. 10), are:
  warm, the passion, my youth, on trace, dark deceit.
- The bigrams of the above noun expressions are:
  [(warm, the passion), (the passion, my youth),
  (my youth, one trace), (one trace, dark deceit)]
- The mappings to the corresponding noun expressions after the removal of any articles:
  warm$\mapsto$warm,
  the passion $\mapsto$ passion,
  my youth $\mapsto$ youth,
  one trace $\mapsto$ one trace,
  dark deceit $\mapsto$ dark deceit.
- Join the bigrams via the preposition 'of'. The result is the production of the following phrases:
  warm of the passion,
  the passion of my youth,
  my youth of one trace,
  one trace of dark deceit.
- Determine which of the above phrases are substrings of the original phrase:
  the passion of my youth, one trace of dark deceit.
- Use the mappings derived earlier to determine the required relations:
  ((passion, youth), the passion of my youth),
  ((one trace, dark deceit), one trace of dark deceit)

The Brown corpus is already part-of-speech tagged. In this case of-prepositions were extracted via regular expression

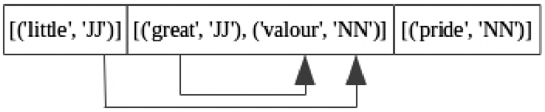


FIGURE 11. Splitting into 'and' | 'or' segments.

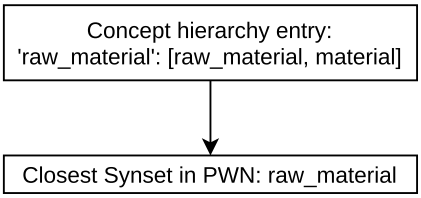


FIGURE 12. Concept hierarchy.

patterns of tagged words. An example of such a pattern is: $\langle NN.^*\rangle + \langle IO\rangle\langle JJ.^*\rangle^*\langle NN.^*\rangle +$. Finally, for the Prepositional Phrase Attachment Corpus [76], the method depicted in Fig. 8 was used.

Some of the noun expressions produced by the above process were conjuncts, or disjuncts of smaller noun expressions. Further steps were taken to split conjuncts and disjuncts in noun expressions into their constituent parts. Noun expressions with more than 5 words rarely occurred, meaning that the maximum of 2 conjuncts or disjuncts were considered (the majority of noun expressions consisted of 2-3 words). In addition, the derivation of concept hierarchies from multi-noun words and noun words preceded by adjectives helped to identify the most specific relevant lexical term in PWN. The actual processing is depicted in Fig. 13, and is described below.

For each noun expression on either site of an of-relation:

1) Get the tagged segments of the noun expressions between the words 'and, or' if present, otherwise get the whole tagged noun expression.
Example 35: if the noun expression is: *'little and great valour and pride'*, then the output of this processing step will be the following tagged segments:

[(little, JJ)], [(great, JJ), (valour, NN)], [(pride, NN)]

2) Using the tags of the segments, identify those tagged words that are adjective, and combine with the (last) noun of the next conjunct (or disjunct) to entail the noun expressions produced from the original expression, and create hierarchies of concepts. Also, create mappings between the original phrase and the produced noun expressions.
Example 36: Continuing with the above example, given the tagged segments produced at the previous stage, the adjectives of the segments produced earlier, are combined as follows:
   a) Concept (noun expression) mappings created:
   little and great valour and pride $\mapsto$
   [little valour, great valour, pride]
   b) Concept hierarchical mappings:
   little valour : [little_valour, valour],

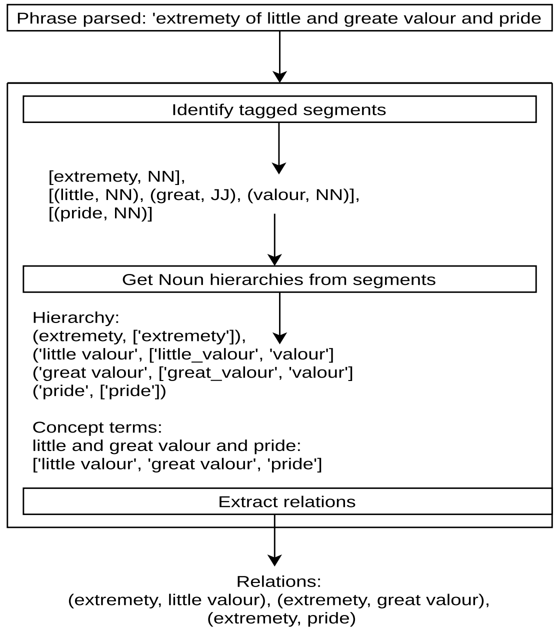


FIGURE 13. Further processing.

   great valour : [great_valour, valour],
   pride : [pride]

3) Use the concept mappings to update the set of relations and the concept hierarchy to determine the most specific matching with a lexical term in PWN.
Continuing with the above example, if $r =$ (extremity, little and great valour and pride) then, by using the concept mappings derived in the previous step, the following relations are produced: (extremity, little valour), (extremity, great valour), (extremity, pride).
The hierarchies of concepts (for example, the hierarchy: [great_valour, valour]), aim to determine the most specific matching of a concept with a lexical term in PWN. It will be used in the next step to determine a mapping with a lexical term in PWN.

The previous paragraphs of this section described the steps followed in order to extract of-relations from the collected resources. The steps followed in order to extract the hypernyms relevant to each concept participating in an of-relation, are outlined below.

1) Extract the details of each noun expression. This involves using the hierarchies of concepts extracted earlier in order to determine the most specific PWN lexical term matching each concept of an r-relation. If there is no entry in PWN for a term, then determine whether the noun expression is recognized as a NER entity, e.g. person. If there is no NER mapping too, then the concept is recorded as unknown so that words not in PWN can traced later on. The set of mappings between the Spacy NER entities [82] and the PWN lexical terms is illustrated in Table 13.
Example 37: Consider the noun expression: 'judicial approval'. The extracted noun hierarchy is

TABLE 13. NER mappings.

NER Type	PWN lexical mapping
ORG	organization
DATE	date
PERSON	person
LOC	location
GPE	geographical_area
PERCENT	percentage
CARDINAL	cardinal_number
QUANTITY	quantity
MONEY	money

[judicial_approval, approval]. The closest matching lexicon entry in PWN is approval. In the case of the noun expression raw_material however, the noun hierarchy extracted is: [raw_material, material] and the closest PWN entry in the lexicon is raw_material. However, the number 1.8% is recognized as the Named Entity PERCENT by Spacy NER, which is then mapped to the word 'percentage', which exists in the vocabulary of PWN.

2) Extract the hypernyms relevant to each noun expression. This step involves using the information derived above, to determine the hypernyms relevant to each noun expression: if there is an entry in PWN for the most specific lexical term matching each concept, then find all hypernym paths of the first sense of this entry. The set of hypernyms relevant to each concept of an of-relation is the union of all hypernyms of all hypernym paths. If there is no entry in PWN for the noun expression under consideration, then the NER mapping is used instead, in a similar way. If there is no NER mapping either, then the noun expression is mapped to the most generic synset entity.n.01 in PWN. This is required in order to be able to establish whether a relation is subsumed by another.
Named entities are processed though a Named Entity Recognition (NER) module (spaCy [82]) in order to determine the type of named entities not included in the lexical database. When proper nouns referring to names of persons, or locations etc. are recognized, they can be mapped to particular senses in PWN. For example, a person's name which is not included in the lexical database, when recognized by the NER module as a type 'PERSON' entity, it is mapped to the sense 'person.n.01' in order to capture the relevant semantic features of the entity being referred. The purpose of this transformation is to avoid loosing substantial information regarding the entities participating in relations where the lexical entries are not included in the vocabulary of PWN.
spaCy [82], is an open source software library for NLP written in Python and Cython. The language model used is the en-core-web-md core model of the English language which is general purpose pretrained model which can be used to predict named entities, part-of-speech tags and syntactic dependencies. The model is trained by using a Convolutional Neural Network (CNN) on annotated English resources including news, conversational telephone speech, weblogs, usenet newsgroups, broadcasts and talk shows, annotated as part of the OntoNotes project [95]. The reason of using a NER API in our case is simply to replace names (proper nouns) that do not exist in the vocabulary of PWN with the type of entity they represent.

The method outlined above for the extraction of hypernyms for each input word, selects the first synset in the list of synsets available for an English word (in most cases, the name of the first synset is identical to the input word lemmatized). The first synset is the most frequently occurring sense in the English SemCor corpus [86], which is a manually annotated subset of Brown corpus. The corpus contains about 700, 000 words, with 226, 036 of them sense-annotated at Princeton University [86]. Selecting the most-frequently occurring sense in a corpus is the default baseline in WSD systems [87]. Although the English SemCor corpus is shown to have 2, 5% incorrect tags [88], the most frequently occurring sense 'is used in supervised algorithms with insufficient training data' [87]. One may argue that WSD trained on the English Semcor corpus is probably inadequate for poetic data. Ambiguity however in poetry does not necessarily imply the use of a special vocabulary with uncommon (less frequent) interpretation. As stated in [89], poets overcome the limitations of language 'by using more or less ordinary words in special ways' [89].

In poetic text, ambiguity, whether it appears in the meaning of words or sentences, or at the syntactic or semantic level, is perceived as a creative tool used deliberately to explore the readers imagination and deepen understanding. For example, Brook [90] addressed ambiguity in proportion to the idea of 'close reading' which is a reading motivated by a skepticism toward apparent meanings' [90]. Empsom, placed ambiguity 'at the very roots of poetry', and as an 'intention to mean several things' [91].

In terms of the of-relations considered here, the purpose of the current work is to establish whether the concepts represented by the noun expressions in either site of an of-construct, can be used to determine whether the of-construct is specific to the poetic (vs non-poetic) dataset. This task is different from the task of predicting the senses of the words or interpreting a Multi-Word Expression (MWE) [92], although the importance of word sense disambiguation and /or Preposition Sense Disambiguation (PSD) should not be undervalued.

APPENDIX B
PROOFS

Proposition 38: For any relation $r \in R$:

$$r \in R_p \quad \text{if } \exists (h_i, h_j) \in S''_p \wedge \mathrm{Assoc}((h_i, h_j), r)$$
$$r \in R_{\mathrm{np}} \quad \text{if } \exists (h_i, h_j) \in S''_{\mathrm{np}} \wedge \mathrm{Assoc}((h_i, h_j), r)$$

Proof of Proposition 22: By contradiction.

Suppose $\exists(h_i, h_j) \in S''_p$ such that $\mathrm{Assoc}((h_i, h_j), r)$, and $r \notin R_{\text{poetic}}$. But, $(h_i, h_j) \in S''_p \Rightarrow (h_i, h_j) \in S_p$, and $(h_i, h_j) \notin S_c$ since $S''_p = S_p - S_c - \{(h_k, h_m) : (h_k, h_m) \sqsubseteq (h_i, h_j)\}$, by construction. But, by (4) it follows that $r \in R_{\text{poetic}}$, which contradicts our original assumption.

Proposition 39: Let $O = R_{\text{poetic}} \cap R_{\text{non-poetic}}$, denote the overlap of relations between the two domains. Then, as the size of O increases, the value of Recall decreases, i.e. recall $\sim \frac{1}{\|O\|}$.

Proof of Proposition 39: For the proof of the above proposition, the evaluation of Algorithm 1 is considered on the entire dataset. If $r \in O$, and hp is any hyper-pair, then $\forall \mathrm{hp} : \mathrm{Assoc}(\mathrm{hp}, r) \Rightarrow \mathrm{hp} \in S_c$, therefore $\mathrm{hp} \notin S''_p \cup S''_{\text{np}}$ (since $S'_p = S_p - S_c$, $S'_{\text{np}} = S_{\text{np}} - S_c$, and $S'_p \supseteq S''_p$, and $S'_{\text{np}} \supseteq S''_{\text{np}}$).

Without loss of generality, let us consider poetic recall (18).

$$\begin{aligned}\mathrm{recall}_p &= \frac{\|\mathrm{CorrectlyPredicted}_{\text{poetic}}\|}{\|R_{\text{poetic}}\|}\\ &= \frac{\|\{r : r \in R_{\text{poetic}} \cap \mathrm{Predicted}_{\text{poetic}}\}\|}{\|R_{\text{poetic}}\|}\end{aligned}$$

Now, a relation is classified as poetic if there is a hyper-pair $\mathrm{hp}_i \in S''_p$, st. $\mathrm{Assoc}(\mathrm{hp}_i, r)$ 22. But, according to what is stated earlier, it then follows that $r \notin O$. Therefore, $O \cap \{r : r \in R_{\text{poetic}}$ and r is classified as poetic$\} = \emptyset$. But, $R_{\text{poetic}} = \{r : r \in R_{\text{poetic}}$ and r is classified as poetic$\} \cup O$. Thus, given a particular set of relations in a poetic domain, as the size of O increases, the size of $\{r : r \in R_{\text{poetic}}$ and r is classified as poetic$\}$ (and therefore the value of recall_p) decreases.

Proposition 40: Consider two datasets D_1 and D_2 say, such that $D_1 \subseteq D_2$. Let R_{D_1} and R_{D_2} be the sets of relations extracted from D_1 and D_2 respectively, R_{cat,D_i} be the set of relations extracted from the type cat resources in D_i and $O_{D_i} = R_{\text{cat},D_i} \cap R_{\overline{\text{cat}},D_i}$ where cat is one of two categories of resources (for example poetic) and $\overline{\text{cat}}$ is the alternative category (for example non-poetic). If $\mathrm{recall}_{\text{cat},D_i}$ is the recall of a category of relations in $D_i \in \{D_1, D_2\}$, then $\mathrm{recall}_{\text{cat},D_1} > \mathrm{recall}_{\text{cat},D_2}$ implies $O_{D_i} \subset O_{D_{i+1}}$

Proof of Proposition 40: $D_1 \subseteq D_2$ if and only if $R_{D_1} \subseteq R_{D_2}$. For brevity, let $C_{\text{cat},D_i} = \mathrm{CorrectlyPredicted}_{\text{cat},D_i}$.

$$\mathrm{recall}_{\text{cat},D_i} = \frac{\|C_{\text{cat},D_i}\|}{\|R_{D_i}\|}$$

where $D_i \in \{D_1, D_2\}$.

Since $R_{\text{cat},D_2} \supseteq R_{\text{cat},D_1}$ then $\|R_{\text{cat},D_2}\| = \|R_{\text{cat},D_1}\| + k$, where $k = \|R_{\text{cat},D_2}\| - \|R_{\text{cat},D_1}\| \geq 0$ is the size of the (new) set of relations of type cat extracted from D_2 which are not included in D_1.

Then, by basic mathematical principles, since $C_{\text{cat},D_1} \leq R_{\text{cat},D_1}$, then:

$$\frac{\|C_{\text{cat},D_1}\|}{\|R_{\text{cat},D_1}\|} \leq \frac{\|C_{\text{cat},D_1}\| + k}{\|R_{\text{cat},D_1}\| + k}$$

But,

$$\frac{\|C_{\text{cat},D_1}\| + k}{\|R_{\text{cat},D_1}\| + k} = \frac{\|C_{\text{cat},D_1}\| + k}{\|R_{\text{cat},D_2}\|}$$

therefore:

$$\frac{\|C_{\text{cat},D_1}\|}{\|R_{\text{cat},D_1}\|} \leq \frac{\|C_{\text{cat},D_1}\| + k}{\|R_{\text{cat},D_2}\|}$$

Now, $\mathrm{recall}_{\text{cat},D_2} < \mathrm{recall}_{\text{cat},D_1} \Leftrightarrow$

$$\frac{\|C_{\text{cat},D_2}\|}{\|R_{D_2}\|} < \frac{\|C_{\text{cat},D_1}\|}{\|R_{D_1}\|}$$

Since $\mathrm{recall}_{\text{cat},D_2} < \mathrm{recall}_{\text{cat},D_1}$, then:

$$\frac{\|C_{\text{cat},D_2}\|}{\|R_{D_2}\|} < \frac{\|C_{\text{cat},D_1}\|}{\|R_{D_1}\|} \leq \frac{\|C_{\text{cat},D_1}\| + k}{\|R_{D_2}\|}$$

which implies that:

$$\frac{\|C_{\text{cat},D_2}\|}{\|R_{D_2}\|} < \frac{\|C_{\text{cat},D_1}\| + k}{\|R_{D_2}\|}$$

But,

$$\frac{\|C_{\text{cat},D_2}\|}{\|R_{D_2}\|} < \frac{\|C_{\text{cat},D_1}\| + k}{\|R_{D_2}\|} \Leftrightarrow \|C_{\text{cat},D_2}\| < \|C_{\text{cat},D_1}\| + k.$$

This is possible if:

- $C_{\text{cat},D_1} \cap R_{\overline{\text{cat}},D_2} \neq \emptyset$, or
- $(R_{\text{cat},D_2} - R_{\text{cat},D_1}) \cap R_{\overline{\text{cat}},D_2} \neq \emptyset$
- the hyper-pairs associated to the new relations extracted from cat resources of D_2 are subsumed by hyper-pairs in S''_{cat} extracted for type cat resources in D_2.

Since $C_{\text{cat},D_1} \subseteq R_{\text{cat},D_2}$ and $(R_{\text{cat},D_2} - R_{\text{cat},D_1}) \subseteq R_{\text{cat},D_2}$ then:

- $C_{\text{cat},D_1} \cap R_{\overline{\text{cat}},D_2} \neq \emptyset \Rightarrow R_{\text{cat},D_2} \cap R_{\overline{\text{cat}},D_2} \neq \emptyset$.
- $(R_{\text{cat},D_2} - R_{\text{cat},D_1}) \cap R_{\overline{\text{cat}},D_2} \Rightarrow R_{\text{cat},D_2} \cap R_{\overline{\text{cat}},D_2} \neq \emptyset$
- the hyper-pairs associated to the new relations extracted from the cat resources of D_2 which are subsumed by hyper-pairs in other sets do not influence O_{D_1} or O_{D_2}

Therefore, it follows that: $O_{D_1} \subseteq O_{D_2}$, as required.

Proposition 41: For any two hyper-pairs hp_1, and hp_2, such that $\mathrm{hp}_1 \sqsubseteq \mathrm{hp}_2$, if $\mathrm{IRF}_{\mathrm{hp}_1} = \mathrm{IRF}_{\mathrm{hp}_2}$, then: $\{r : \mathrm{Assoc}(\mathrm{hp}_1, r)\} = \{r : \mathrm{Assoc}(\mathrm{hp}_2, r)\}$.

Proof of Proposition 41: Suppose $\mathrm{hp}_1 = (\mathrm{hp}_{11}, \mathrm{hp}_{12})$, and $\mathrm{hp}_2 = (\mathrm{hp}_{21}, \mathrm{hp}_{22})$. By definition 18, $\mathrm{hp}_1 \sqsubseteq \mathrm{hp}_2$ iff $\exists p \in \mathrm{hypernym\_paths}(h_{11})$ and $\exists q \in \mathrm{hypernym\_paths}(h_{12})$ such that: $h_{11} @\underset{p}{\rightarrow}{}^* h_{21}$ and $h_{12} @\underset{q}{\rightarrow}{}^* h_{22}$. Now, h_{11}, h_{12}, h_{21}, and h_{22} are just synsets in the taxonomic hierarchy of PWN, so that any lexical entry, w_1 say, with sense h_{11} is also a member of the set of lexical entries that belong to the superordinates of h_{11}. A similar argument applies for lexical entries with sense h_{12}. Thus if $(\mathrm{hp}_{11}, \mathrm{hp}_{12})$ is associated to any relation $r = (w_1, w_2)$, then all subsuming hyper-pairs are associated to r too. Thus, $\{r : \mathrm{Assoc}(\mathrm{hp}_1, r)\} \subseteq \{r : \mathrm{Assoc}(\mathrm{hp}_2, r)\}$

APPENDIX C
FURTHER RESULTS

TABLE 14. Alg. 1, Exp. 1–2 (Prediction).

PREDICTION	Exp. 1		Exp. 2	
No. of Relations whose hpairs are a subset of	Poetic	Non-Poetic	Poetic	Non-Poetic
S''_p	1476	273	2051	336
$S''_{\mathrm{np}} \cap S''_p$	599	542	707	554
S''_{np}	215	1063	216	980
S_c	19	3	18	4
Removed hpairs	283	177	403	184
Total	2592	2058	3395	2058

TABLE 15. Alg. 1, Exp. 3–4 (Prediction).

PREDICTION	Exp. 3		Exp. 4	
No. of Relations whose hpairs are $\subseteq$:	Poetic	Non-Poetic	Poetic	Non-Poetic
S''_p	2019	362	1854	348
$S''_{\mathrm{np}} \cap S''_p$	724	584	840	701
S''_{np}	218	1052	281	1650
S_c	15	3	17	7
Removed hpairs	419	200	403	334
Total	3395	2201	3395	3040

TABLE 16. Alg. 1, Exp. 5 (Prediction).

Experiment 5 (Prediction)		
No. of relations whose hpairs are $\subseteq$	Brown Religion	Brown Hobbies
S''_{religion}	93	26
$S''_{\mathrm{hobbies}} \cap S''_{\mathrm{religion}}$	72	122
S''_{hobbies}	44	215
S_c	0	0
Removed hpairs	25	21
Total	234	384

TABLE 17. Alg. 1, Exp. 6 (Prediction).

Experiment 6 (Prediction)		
No. of relations whose hpairs are $\subseteq$	Art-history Cultural	Reuters Earn, acq, gold
S''_{AH}	1465	108
$S''_{\mathrm{AH}} \cap S''_{\mathrm{Reuters}}$	330	303
S''_{Reuters}	105	1118
S_c	7	21
Removed hpairs	188	261
Total	2095	1811

S''_{AH} = Set of hyper-pairs that appear only in relations extracted from the Cultural and Art History dataset.

TABLE 18. Algorith 1, Experiment 5 (whole data).

Experiment 5 (whole data)		
No. of relations whose hpairs are $\subseteq$	Brown Religion	Brown Hobbies
S''_{religion}	1109	0
$S''_{\mathrm{hobbies}} \cap S''_{\mathrm{religion}}$	0	0
S''_{hobbies}	0	1835
S_c	15	14
Removed hpairs	42	69
Total	1166	1918

TABLE 19. Alg. 1, Exper. 6.

Experiment 6 (whole data)		
No. of relations whose hpairs are $\subseteq$	Art-history Cultural	Reuters (Earn, Acq,gold)
S''_{AH}	9560	0
$S''_{\mathrm{AH}} \cap S''_{\mathrm{Reuters}}$	0	0
S''_{Reuters}	0	7755
S_c	61	100
Removed hpairs	851	1198
Total	10472	9053

S''_{AH} = Set of hyper-pairs that appear only in relations extracted from the Cultural and Art History dataset.

TABLE 20. Alg. 2, Exp.1 (Prediction).

Experiment 1				
No. of relations classified as:	Poetic	Non-Poetic	C^*	Total
Poetic	1767	667	-3	2431
Non-poetic	778	1387	-1	2164
Non-classified[a]	47	4	0	51
Total	2592	2058	-4	4646
Precision	72.7%	64.1%		
Recall	68.2%	67.4%		

[a] Non-classified = No. of relations with zero weights

TABLE 21. Alg. 2 Exp.2 (Prediction).

Experiment 2				
No. of relations classified as:	Poetic	Non-Poetic	C^*	Total
Poetic	2398	702	-1	3099
Non-poetic	938	1354	-2	2290
Non-classified[a]	59	2	0	61
Total	3395	2058	-3	5450
Precision	77.4%	59.1%		
Recall	70.6%	65.8%		

[a] Non-classified = No. of relations with zero weights

TABLE 22. Alg. 2 Exp. 3 (Prediction).

Experiment 3				
No. of relations classified as:	Poetic	Non-Poetic	C^*	Total
Poetic	2397	731	-4	3124
Non-poetic	942	1468	-2	2408
Non-Classified	56	2	0	58
Total	3395	2201	-6	5590
Precision	76.7%	61%		
Recall	68.2%	67.4%		

TABLE 23. Alg. 2 for Exp. 4 (Prediction).

Experiment 4				
No. of relations classified as:	Poetic	Non-Poetic	C^*	Total
Poetic	2286	947	-6	3227
Non-poetic	1052	2089	-1	3140
Non-Classified	57	4	-7	54
Total	3395	3040	-14	6421
Precision	70.8%	66.5%		
Recall	67.3%	68.7%		

TABLE 24. Alg. 2 for Exp. 5.

Experiment 5 (whole dataset)				
No. of relations classified as:	Religion	Hobbies	C^*	Total
Religion	1141	74	-3	1212
Hobbies	25	1844	-2	1867
Non-classified[a]	0	0	0	0
Total	1166	1918	-5	3079
Precision	94.1%	98.8%		
Recall	97.9%	96.1%		

TABLE 25. Alg. 2 for Exp. 6.

Experiment 6 (whole dataset)				
No. of relations classified as:	Art-history	Reuters	C^*	Total
Art-History	9795	848	-3	10640
Reuters	677	8205	-13	8869
Non-classified[a]	0	0	0	0
Total	10472	9053	-16	19509
Precision	92.6%	92.5%	-	-
Recall	93.5%	90.6%	-	-

[a] Non-classified = No. of relations with zero weights

TABLE 26. Alg. 2, Exp. 5.

Experiment 5 (Prediction)				
No. of relations classified as:	Religion	Hobbies	C^*	Total
Religion	138	104	0	242
Hobbies	92	276	0	368
Non-classified	4	4	0	8
Total	234	384	0	618
Precision	57%	75%		
Recall	59%	72%		

TABLE 27. Alg. 2, Exp. 6.

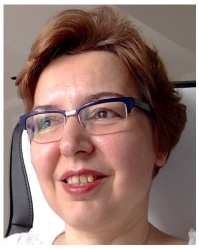
Experiment 6 (Prediction)				
No. of relations classified as:	Art-History	Reuters	C^*	Total
Art-History	1531	527	0	2058
Reuters	551	1277	0	1828
Non-classified	13	7	0	20
Total	2095	1811	0	3906
Precision	79.2%	75.8%	-	
Recall	78.5%	75.7%	-	

REFERENCES

[1] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks and multitask learning," in *Proc. ICML* 2008, pp. 160–167.
[2] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Comput. Intell. Mag.*, vol. 13, no. 3, pp. 55–75, Aug. 2018.
[3] C. D. Manning, "Computational linguistics and deep learning," *Comput. Linguistics*, vol. 41, no. 4, pp. 701–707, Dec. 2015.
[4] G. A. Miller, "WordNet: A lexical database for English," *Commun. ACM*, vol. 38, no. 11, pp. 39–41, 1995.
[5] C. Panayiotou, "An ontological approach to the extraction of figures of speech," in *Proc. 5th Int. Conf. Comput. Sci. Inf. Technol. (CSTY)*, Dubai, India, Nov. 2019, pp. 1–8.
[6] D. Jurafsky and H. J. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Upper Saddle River, NJ, USA: Prentice-Hall, 2009.
[7] T. Wei, Y. Lu, H. Chang, Q. Zhou, and X. Bao, "A semantic approach for text clustering using WordNet and lexical chains," *Expert-Syst. Appl.*, vol. 42, pp. 2264–2275, May 2015.
[8] A. Hotho, S. Staab, and G. Stumme, "Wordnet improves text document clustering," in *Proc. SIGIR Semantic Web Workshop*, 2003, pp. 1–5.
[9] M. R. de Buenaga, J. M. H. Gomez, and B. D. Agudo, "Using WordNet to complement training information in text categorization," in *Proc. RANLP*, 1999, pp. 353–364.
[10] R. Bai, X. Wang, and J. Liao, "Extract semantic information from WordNet to improve text classification performance," in *Advances in Computer Science and Information Technology* (Lecture Notes in Computer Science), vol. 6059. Berlin, Germany: Springer, 2018, pp. 409–420.
[11] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *Proc. Int. World Wide Conf.*, 2003, pp. 519–528.
[12] A. Blodgett and N. Schneider, *Semantic Supersenses of English Possessives*. Road Hulbert, OK, USA: LREC, 2018.
[13] J. Sinclair and L. Sinclair, *Corpus, Concordance Collocation*. London, U.K.: Oxford Univ. Press, 1991.
[14] C. Owen, "Notes on the of ness of—Sinclair and grammar," *Int. J. Corpus Linguistics*, vol. 12, no. 2, pp. 201–221, 2007.
[15] C. Liyin and S. Chung, "Of-constructions in the predicate of demonstrate and show in academic discourse," in *Proc. 27th Pacific Asia Conf. Lang., Inf., Comput.*, 2013, pp. 105–144.
[16] P. Partee and V. Borschev, "Genitives, relational nouns, and argument-modifier ambiguity," in *Modifying Adjuncts*, vol. 4, E. Lang, C. Maienborn, C. Fabricius-Hansen, Eds. Berlin, Germany: Mouton de Gruyter, 2003.
[17] B. Dancygier and E. Sweeter, "Figurative Language," in *Cambridge Textbooks in Linguistics*, Cambridge, U.K.: Cambridge Univ. Press, 2014.
[18] B. Vickers, *Counterfeiting' Shakespeare: Evidence, Authorship and John Ford's Funerall Elegye*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[19] M. Lesk, "Automated sense disambiguation using machine-readable dictionaries: How to tell a pine cone from an ice cream cone," in *Proc. SIGDOC*, Toronto, CA, USA, 1986, pp. 24–26.
[20] Y. Wang, M. Wang, and H. Fujita, "Word sense disambiguation: A comprehensive knowledge exploitation framework," in *Knowledge-Based Systems*, vol. 190. Amsterdam, The Netherlands: Elsevier, 2020.
[21] A. Montoyo, M. Palomar, G. Rigau, and A. Suarez, "Combining knowledge and corpus-based word-sense-disambiguation methods," 2011, *arXiv:1109.2130*.
[22] R. Navigli, "Word sense disambiguation: A survey," *ACM Comput. Surv.*, vol. 41, no. 10, pp. 1–69, 2009.
[23] H. Wimmer and L. Zhou, *Word sense disambiguation for ontology learning*. Koblenz, Germany: Research-in-Progress, 2013.
[24] D. Schmidt, R. Basso, C. Trojahn, and R. Vieira, "Matching domain and top-level ontologies exploring word sense disambiguation and word embedding," in *Proc. 13th ISWC Workshop*, 2018, pp. 1–5.
[25] K. Kotis and G. A. Vouros, "The HCONE approach to ontology merging," in *The Semantic Web: Research and Applications* (Lecture Notes in Computer Science), vol. 3053, C. Bussler, J. Davies, and R. Studer, Eds. Berlin, Germany: Springer, 2004.
[26] I. Niles and A. Pease, *Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology*. Chennai, India: IKE, 2004.
[27] C. Fellbaum, "WordNet and wordnets," in *Encyclopedia of Language and Linguistics*, K. Brown Ed. Oxford, U.K.: Elsrvier, 2005, pp. 665–670.
[28] G. A. Miller, "Nouns in WordNet: A lexical inheritance system," *Int. J. Lexicography*, vol. 3, pp. 245–264, Oct. 1993.
[29] G. A. Miller and P. N. Johnson-Laird, *Language and Perception*. London, U.K.: Belknap Press, 1976.
[30] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller, "Introduction to WordNet: An on-line lexical database," *Int. J. Lexicography*, vol. 3, no. 4, pp. 235–244, 1990.
[31] V. Basile, "Wordnet as an ontology for generation," in *Proc. 1st Int. Workshop Natural Lang. Gener. Semantic Web*, 2015, pp. 1–3.
[32] R. T. Gruber, "A translation approach to portable ontology specifications," *Knowl. Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.
[33] W. Borst, "Construction of engineering ontologies," Ph.D. dissertation, Inst. Telematica Inf. Technol., Univ. Twente, Enschede, The Netherlands, 1997.

[34] R. Studer, R. Benjamins, and D. Fensel, "Knowledge and engineering: Principles and methods," *Data Knowl. Eng.*, vol. 25, nos. 1–2, pp. 161–198, 1998.
[35] S. Grimm, P. Hitzler, and A. Abecker, "Knowledge representation and ontologies," in *Semantic Web Services*, R. Studer, S. Grimm, A. Abecker, Eds. Berlin, Germany: Springer, 2007.
[36] A. Gangemi, R. Navigli, and P. Velardi, "The ontowordnet project: Extension and axiomatization of conceptual relations in wordnet," in *Proc. Meaningful Internet Syst.*, 2003, pp. 820–838.
[37] F. Giunchiglia and I. Zaihrayeu. (2007). *Lightweight Ontologies*. [Online]. Available: http://www.dit.unitn.it
[38] M. Yatskevich and F. Giunchiglia, "Element Leven semantic matching using wordnet," University of Trento, Trento, Italy, Tech. Rep. DIT-04-035, 2007.
[39] J. Euzenat and P. Shvaiko, *Ontology Matching*. Berlin, Germany: Springer, 2013.
[40] W. Y. Wong, W. Liu, and M. Bennamoun, "Ontology learning from text: A look back and into the future," *ACM Comput. Surv.*, vol. 44, no. 4, pp. 1–36, 2012.
[41] M. Berland and E. Charniak, "Finding parts in very large corpora," in *Proc. 37th Conf. Comput. Linguistics*, College Park, MD, USA, 1999, pp. 57–64.
[42] M. N. Asim, M. Wasim, M. Khan, W. Mahmood, and H. M. Abbasi, "A survey of ontology learning techniques and Applications," *J. Biol. Databases Curation*, vol. 2018, Oct. 2018, Art. no. bay101.
[43] M. Hearst, "Automated discovery of wordnet relations," in *WordNet: An Electronic Lexical Database*, C. Fellbaum, Ed. Cambridge, MA, USA: MIT Press, 1998, pp. 131–153.
[44] S. Mukherjea and S. Sahay, "Discovering biomedical relations utilizing the World-wide web," in *Proc. Pacific Symp. Biocomput.*, 2016, pp. 16–175.
[45] P. Cimiano, S. Handschuh, and S. Staab, "Towards the self-annotating web," in *Proc. Conf. WWW*, New York, NY, USA, 2004, pp. 462–471.
[46] A. Agustini, P. Gamallo, and G. Lopes, "Selection restrictions acquisition for parsing improvement," in *Web Knowledge Management and Decision Support* (Lecture Notes in Computer Science), vol. 2543. Berlin, Germany: Springer, 2003, pp. 129–143.
[47] F. Smadja, "Retrieving collocations from text: Xtract," *J. Comput. Linguistics*, vol. 19, no. 1, pp. 143–178, 1993.
[48] P. W. Foltz, "Latent semantic analysis for text-based research," *Behav. Res. Methods, Instrum. Comput.*, vol. 28, no. 2, pp. 197–202, 1996, doi: 10.3758/BF03204765.
[49] S. T. Dumais, "Latent semantic analysis," *Annu. Rev. Inf. Sci. Technol.*, vol. 38, no. 1, pp. 188–230, 2004.
[50] P. Cimiano, A. Hotho, and S. Staab, "Learning concept hierarchies from text corpora using formal concept analysis," *J. Artif. Intell. Res.*, vol. 24, pp. 305–339, Oct. 2005.
[51] D. Faure, C. Nédellec, and C. Rouveirol, "Acquisition of semantic knowledge using machine learning methods: The system ASIUM," Universite Paris Sud, Orsay, France, Tech. Rep. ICS-TR-88-16, 1998.
[52] R. Basili, P. Velardi, and M. T. Pazienza, "Integrating general-purpose and corpus-based verb classification," *Comput. Linguistics*, vol. 22, no. 4, pp. 559–568, 1996.
[53] M. Dahab and A. Rafea, "TextOntoEx: Automatic ontology construction from natural English text," *Expert Syst. Appl.*, vol. 34, pp. 1474–1480, Oct. 2008.
[54] P. Buitelaar, D. Olejnik, and M. Sintek, "A protégé plug-in for ontology extraction from text based on linguistic analysis," in *The Semantic Web: Research and Applications* (Lecture Notes in Computer Science), C. J. Bussler, J. Davies, and R. Studer, Eds. Berlin, Germany: Springer, 2004, pp. 31–34.
[55] J. Völker, P. Hitzler, and P. Cimiano, "Acquisition of OWL DL axioms from lexical resources," in *Proc. Res. Appl.*, 2007, pp. 670–685.
[56] D. Lin, "Dependency-based evaluation of MINIPAR," in *Proc. Workshop Eval. Parsing Syst.*, vol. 20, Granada, Spain, Jan. 2003, pp. 317–329, doi: 10.1007/978-94-010-0201-1_18.
[57] A. Maedche and S. Staab, "Ontology learning for the semantic web," *IEEE Intell. Syst.*, vol. 16, no. 2, pp. 72–79, Mar./Apr. 2001.
[58] J. Völker, P. Haase, and P. Hitzler, "Learning expressive ontologies," *Frontiers Artif. Intell. Appl.*, vol. 167, Jun. 2009.
[59] R. Azevedo, F. Freitas, R. Rocha, J. Menezes, C. Rodrigues, and M. Gomes, "Representing knowledge in DL ALC from text," *Proc. Comput. Sci.*, vol. 35, pp. 176–185, Oct. 2014.
[60] P. Cimiano, P. Buitelaar, J. McCrae, and M. Sintek, "LexInfo: A declarative model for the lexicon-ontology interface," *J. Web Semantics*, vol. 9, no. 1, pp. 29–51, 2011.
[61] J. McCrae, D. Spohr, and P. Cimiano, "Linking lexical resources and ontologies on the semantic web with lemon," in *The Semantic Web: Research and Applications* (Lecture Notes in Computer Science), vol. 6643, G. Antoniou, Ed. Berlin, Germany: Springer, 2011, pp. 245–259.
[62] J. McCrae, C. Fellbaum, and P. Cimiano, "Publishing and linking wordnet using lemon and RDF," *Proc. 3rd Workshop Linked Data Linguistics*, 2014, pp. 1–4.
[63] C. Fellbaum, *WordNet: An Electronic Lexical Database*. Cambridge, MA, USA: MIT Press, 1998.
[64] G. Miller and C. Fellbaum, "Semantic networks of English," *Cognition*, vol. 41, nos. 1–3, pp. 197–229, 1991.
[65] W. Shakespeare. (2015). *The Complete Works of William Shakespeare*. [Online]. Available: https://www.gutenberg.org/ebooks/100
[66] W. Blake. (2017). *Songs of Innocence and Songs of Experience*. [Online]. Available: https://www.gutenberg.org/ebooks/1934
[67] J. Milton. (1991). *Paradise Lost*. [Online]. Available: https://www.gutenberg.org/ebooks/20
[68] L. Byron. (2015). *Byron's Poetical Works*. [Online]. Available: https://www.gutenberg.org/ebooks/8861
[69] E. Pound (2015). *Cathay Translations*. [Online]. Available: https://www.gutenberg.org/ebooks/50155
[70] B. R. Yeats. (2010). *The Wind Among the Reeds*. [Online]. Available: https://www.gutenberg.org/ebooks/32233
[71] T. S. Eliot (1998). *The Waste Land*. [Online]. Available: https://www.gutenberg.org/ebooks/1321
[72] W. Wordsworth. (2005). *Poems in Two Volumes*. [Online]. Available: https://www.gutenberg.org/ebooks/8774
[73] H, Doolittle. (May 2, 2009). *Sea Garden*. [Online]. Available: https://www.gutenberg.org/ebooks/28665
[74] E, Dickinson. (Jun. 6, 1998). *Poems: Three Series, Complete*. [Online]. Available: https://www.gutenberg.org/ebooks/12242
[75] S. Bird, E. Klein, and E. Lopez, *Natural Language Processing in Python*. Newton, MA, USA: O'Reilly Media, 2009.
[76] A. Ratnaparkhi, J. Reymar, and S. Roukos, "A maximum entropy model for prepositional phrase attachment," in *Proc. ARPA*, 1994, pp. 1–6.
[77] C. F. Spurgeon, *Shakespeare's Imagery and What it Tells Us*. Cambridge, U.K.: Cambridge Univ. Press, 1935.
[78] P. Vossen and G. Rigau, "Division of semantic labor in the global wordnet grid," in *Proc. Global WordNet Conf.*, 2010, pp. 1–8.
[79] J. Gracia, D. Vila-Suero, J. P. McCrae, T. Flati, C. Baron, and M. Dojchinovski, "Language resources and linked data: A practical perspective," in *Proc. Int. Conf. Knowl. Eng. Knowl. Manage.*, 2014, pp. 1–17.
[80] McCrae. (2010). *The Lemon Cookbook*. [Online]. Available: http://lemon-model.net/lemon-cookbook.pdf
[81] K. Ahrens, S. F. Chung, and C. Huang, "From lexical semantics to conceptual metaphors: Mapping principle verification with WordNet and SUMO," in *Proc. CLSW*, 2003, pp. 1–5.
[82] M. Honnibal and I. Montani, "spaCy2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing," 2017. [Online]. Available: http://citebay.com/how-to-cite/spacy/
[83] H. Bermúdez-Sabel, M. Curado Malta, and E. Gonzalez-Blanco, "Towards interoperability in the European poetry community: The standardization of philological concepts," in *Int. Conf. Lang., Data Knowl.*, 2019, pp. 156–165.
[84] P. Brazdil, P. Silvano, F. Silva, S. Hassan, F. Oliveira, J. Cordeiro, and A. Leal, "Extending a general sentiment lexicon to specific domains in a (semi) automatic manner," in *Proc. Workshop Sentiment Anal. Linguistic Linked Data*, 2021, pp. 1–5.
[85] D. Klein and C. D. Manning, "Accurate unlexicalized parsing," in *Proc. 41st Meeting ACL*, 2003, pp. 423–430.
[86] S. Landes, C. Leacock, and C. Fellbaum, "Building semantic concordances," in *WordNet: An Electronic Lexical Database*, C. Fellbaum, Ed. Cambridge, MA, USA: MIT Press, 1998, pp. 199—216.
[87] D. Jurafsky and H. Martin, *Speech and Language Processing : An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 3rd ed. 2020. [Online]. Available: https://web.stanford.edu/~jurafsky/slp3/

[88] L. Bentivogli and E. Pianta, "Exploiting parallel texts in the creation of multilingual semantically annotated resources: The multisemcor corpus," *Natural Lang. Eng.*, vol. 11, no. 3, p. 247—780, 2005.
[89] J. Korg, *An Introduction to Poetry*. New York, NY, USA: Rinehard and Winston, 1965.
[90] D. Brooks, "Ambiguity, the literary, and close reading," *Comparative Literature Culture*, vol. 12, no. 4, pp. 1–9, 2010.
[91] W. Empson, *Seven Types of Ambiguity The foundations of English Criticism*. Moscow, Russia: Progress, 1997.
[92] I. A. Sag, T. Baldwin, and F. Bond, "Multiword expressions: A pain in the neck for NLP," in *Proc. Int. Conf. Intell. Text Process. Comput. Linguistics*, 2002, pp. 1–15.
[93] M. Davies (2008). *The Corpus of Contemporary American English (COCA)*. [Online]. Available: https://www.english-corpora.org/coca/
[94] M. Ganapathibhotla and B. Liu, "Mining opinions in comparative sentences," in *Proc. COLING*, 2008, pp. 241–248.
[95] R. Weischedel, M. Palmer, M. Marcus, E. Hovy, S. Pradhan, L. Ramshaw, N. Xue, A. Taylor, J. Kaufman, M. Franchini, M. El-Bachouti, R. Belvin, and A. Houston. (2013). *OntoNotes Release 5.0*. [Online]. Available: https://catalog.ldc.upenn.edu/LDC2013T19
[96] P. Constant, "L'analyseur linguistique SYLEX," in *Proc. Ecole CNET*, 1995, pp. 1–5.
[97] G. Bison, "Conceptual clustering," in *Proc. MLnet Summer Schoo Mach. Learn. Knowl. Acquisition*, Dourdan, France, 1994, pp. 1–20.
[98] P. F. Patel-Schneider, P. Hayes, and I. Horrocks, "OWL web ontology language overview," W3C, Cambridge, MA, USA, Tech. Rep. REC-owl-features-20040210, 2004.
[99] A. Stefanowitsch, "Possession and partition: The two genitives of English," in *Proc. Cogn. Linguistics, Explor., Appl., Res.*, 1998, pp. 1–5.
[100] H. Mesarić, "Periphrastic vs. inflected genitive in different registers of English: A corpus study," Ph.D. dissertation, Dept. English Lang. Literature, Univ. Osijek, Osijek, Croatia, 2014.
[101] R. Navigli and P. Velardi, "Ontology enrichment through automatic semantic annotation of on-line glossaries," in *Proc. Int. Conf. Knowl. Eng. Knowl. Manage.*, 2006, pp. 126–140.
[102] H. J. Gomez and J. C. C. Perez, "The role of word sense disambiguation in automated text categorization," in *Proc. NLDB*, 2005, pp. 298–309.
[103] W. A. Gale, K. W. Church, and D. Yarowsky, "Estimating upper and lower bounds on the performance of word-sense disambiguation programs," in *Proc. ACL*, 1992, pp. 249–256.
[104] H. Gong, J. Mu, S. Bhat, and P. Viswanath, "Preposition sense disambiguation and representation," in *Proc. EMNLP*, 2018, pp. 1510–1521.
[105] A. A. Ibrahim, "Ambiguity in poetry: Definition, function and elements," *Alustath J. Hum. Social Sci.*, vol. 72, pp. 801–816, 2008.
[106] A. Kehagias, V. Petridis, V. Kaburlasos, and P. Frangou, "A comparison of word- and sense-based text classification using several classification algorithms," *Journal of Intelligent Information Systems*, vol. 21, no. 3, pp. 227–247, 2003.
[107] H. Sundblad, "Automatic acquisition of hyponyms and meronyms from question corpora," in *Proc. 15th ECAI*, Lyon, France, 2002, pp. 1–5.
[108] P. Buitelaar, P. Cimiano, and Bernardo, *Ontology Learning from Text: Methods, Evaluation and Applications*, Vol. 123. Amsterdam, The Netherlands: IOS Press, 2005.
[109] P. Bouquet, F. Giunchiglia, V. F. Harmelen, and L. H. Serafini, "C-OWL: Contextualizing Ontologies," in *Proc. Int. Semantic Web Conf.*, vol. 2870, 2003, pp. 164–179.
[110] C. Ghidini and F. Giunchiglia, "Local models semantics, or contextual reasoning=locality+ compatibility," *J. Artif. Intell.*, vol. 127, no. 2, pp. 221–259, Oct. 2010.
[111] J. Rocchio, "Relevance feedback in information retrieval," in *SMART Retrieval System: Experiments in Automatic Document Processing*, G. Salton, Ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2013, pp. 313–323.



CHRISTIANA PANAYIOTOU was born in Famagusta, Cyprus, in June 1964. She received the Bachelor of Science (B.Sc.) degree in statistics with computer science from the Queen Mary and Westfield College, University of London, in 1987, the Advanced Master of Science (M.Sc.) degree in database and information systems from the Birkbeck College, University of London, in 1993, the Master of Philosophy (M.Phil.) degree in argumentation and decision making, in 1999, and the Ph.D. degree in viewpoint discrepancies in ontological learning resources. During 1988-1992, she worked as a Systems Designer and Programmer, and she became part-qualified with ACCA. During 1996-1999, she worked as a GTA with the Department of Computer Science, Queen Mary and Westfield College. From 2001 to 2008, she worked as a Lecturer with the Department of Computer Science, Higher Technical Institute, Cyprus. During 2006-2010, she attended full-time studies at the University of Leeds. In 2008, she was transferred to the Cyprus University of Technology. She is currently an Assistant Professor with the Cyprus University of Technology.