On Aligning OpenIE Extractions with Knowledge Bases: A Case Study

Kiril Gashteovski[1,2], **Rainer Gemulla**[1], **Bhushan Kotnis**[2], **Sven Hertling**[1], **Christian Meilicke**[1]
[1]University of Mannheim, Germany, [2]NEC Labs Europe GmbH, Germany
{k.gashteovski,rgemulla}@uni-mannheim.de, bhushan.kotnis@neclab.eu
{sven,christian}@informatik.uni-mannheim.de

Abstract

Open information extraction (OIE) is the task of extracting relations and their corresponding arguments from natural language text in unsupervised manner. Outputs of such systems are used for downstream tasks such as question answering and automatic knowledge base (KB) construction. Many of these downstream tasks rely on aligning OIE triples with reference KBs. Such alignments are usually evaluated w.r.t. a specific downstream task and, to date, no direct manual evaluation of such alignments has been performed. In this paper, we directly evaluate how OIE triples from the OPIEC corpus are related to the DBpedia KB w.r.t. information content. First, we investigate OPIEC triples and DBpedia facts having the same arguments by comparing the information on the OIE surface relation with the KB relation. Second, we evaluate the expressibility of general OPIEC triples in DBpedia. We investigate whether—and, if so, how—a given OIE triple can be mapped to a single KB fact. We found that such mappings are not always possible because the information in the OIE triples tends to be more specific. Our evaluation suggests, however, that significant part of OIE triples can be expressed by means of KB formulas instead of individual facts.

1 Introduction

Open Information Extraction (OIE) systems extract relations and their corresponding arguments from natural language text in unsupervised manner (Banko et al., 2007). Consider the sentence *"Bell, which is a telecommunication company, is located in L. A."*; an OIE system may extract the triples: *("Bell"; "is"; "telecommunication company")* and *("Bell"; "is located in"; "L. A.")*. Such triples are used in downstream tasks, such as word embeddings generation (Stanovsky et al., 2015), information retrieval (Kadry and Dietz, 2017) and entity aspect linking (Nanni et al., 2019).

OIE triples contain surface relations, which often makes their semantics ambiguous (Gashteovski et al., 2019). This poses difficulties for OIE triples to be used in downstream tasks (Broscheit et al., 2017). By contrast, KB relations have precise semantics and are machine-readable (Banko and Etzioni, 2008). To bridge this gap between OIE and KBs, many methods were proposed for aligning OIE triples with reference KBs. In such work, the goal is to associate an OIE triple with an *existing* KB fact (assuming they have the same disambiguated arguments), such that both triples have the same semantics; e.g., the OIE triple *(Jeff Bezos; "be CEO of"; Amazon.com)* and the KB fact (Jeff Bezos; dbo:ceo; Amazon.com). These methods are primarily used for bootstrapping OIE systems (Lockard et al., 2019), but also for other tasks such as link prediction (Gupta et al., 2019). Other methods map any OIE triple to a KB schema (Zhang et al., 2019); e.g. the OIE triple *(Emmanuel Macron; "be president of"; France)* could be mapped to (E. M.; dbo:president; France) even if this fact is not present in the reference KB. Such methods are used for downstream tasks such as automatic KB construction (Dong et al., 2014).

To date, alignments between OIE triples and KBs are evaluated automatically w.r.t. downstream task. Such automatic evaluations, however, do not provide insights about the information content between the alignments, which require expert manual evaluation. In this paper, we manually compare the information content of an OIE corpus—OPIEC (Gashteovski et al., 2019)—and reference KB—DBpedia (Auer et al., 2007)—under optimal alignments.[1] Both resources are automatically generated from Wikipedia, making them comparable resources for evaluation.

[1] All resources of the study are available on `https://www.uni-mannheim.de/dws/research/resources/opiec/`

Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems (Eval4NLP), pages 143–154,
November 20, 2020. ©2020 Association for Computational Linguistics

First, we study the properties of the alignments between OPIEC triples and DBpedia facts which have the same argument pair. Consider the OIE triple *t: (Jeff Bezos; "is CEO of"; Amazon.com)* and two possible KB alignments f_1: (J. B.; dbo:ceo; Amazon.com) and f_2: (J. B.; dbo:employer; Amazon.com). The fact f_1 has same semantics as the OIE triple t. However, t is semantically more specific than f_2, since it provides additional information about J. B. being employed as CEO. Therefore, f_2 expresses *some* information in t, but not all information. In our evaluation, we consider the best possible alignment (e.g., f_1 is considered to be the best alignment) and we investigate its semantics. Note that our goal is *not* to compare different alignment strategies. Rather, *we consider the best possible alignment* and the goal is to *investigate the limits* of such alignments. We found that these alignments are usually semantically related, but quite often the open relation is more specific, thus carrying more information than the KB fact.

Second, we evaluate the expressibility of any OPIEC triple w.r.t. DBpedia by studying whether a given OIE triple can be mapped to a KB fact. In this case, there might not be a known relation in DBpedia between the arguments of the OPIEC triple. We evaluate whether an OPIEC triple can be expressed with a DBpedia fact. Consider the OIE triple *(Emmanuel Macron; "be president of"; France)*. DBpedia does not contain this fact, nevertheless, it can be fully expressed with (E. M.; dbo:president; France) and partially expressed with (E. M.; dbo:nationality; France). We found that most OPIEC triples can be expressed with DBpedia facts, but many of them only partially. Moreover, large fraction of the partially expressible triples can be fully expressed with KB formulas; e.g. OIE triple *(J. F. Kennedy; "be grandchild of"; P. J. Kennedy)* can be partially expressed with the KB fact (J. F. K.; dbo:relative; P. J. K.) and fully expressed with KB formula $\exists x$: (J. F. K.; dbo:parent; x) $\wedge$ (x; dbo:parent; P. J. K.).

Our evaluations focus on the OPIEC corpus, which was extracted with the OIE system MinIE. This makes the evaluation focused on one particular OIE system. To gain insight into transferability, we studied how the findings of our evaluations transfer to OIE triples produced by other OIE systems. We found that the results generally transfer over, though some OIE systems tend to produce more specific output.

2 Analysis of OPIEC Triples and DBpedia Facts with Same Arguments

In this section, we evaluate the semantics of alignments between OPIEC triples and DBpedia facts with the same arguments. Such alignments are inspired by the Distant Supervision Assumption (DSA), which is originally used for traditional information extraction tasks (Mintz et al., 2009). The DSA asserts that whenever there is a KB fact and a sentence mentioning the entity pair of the KB fact, then that sentence expresses the information contained in the KB fact. Similarly, the DSA within OIE context asserts that whenever there is an OIE triple for which there is a KB fact having the same arguments, then the OIE triple expresses the information of the KB fact.

DSA is key assumption used for bootstrapping OIE extractors (Pal and Mausam, 2016). By using the DSA, some methods bootstrap a training set that is used either for learning OIE extraction rules (Gotti and Langlais, 2019) or learning a neural model for extracting OIE triples (Cui et al., 2018; Kolluru et al., 2020). Wu and Weld (2010) used Wikipedia infoboxes (via DBpedia) as source for distant supervision: if there is Wikipedia sentence that contains an entity pair and a corresponding DBpedia entry with the same entity pair, then they store the syntactic patterns (e.g., shortest path in the dependency parse tree) between the two entities. These syntactic patterns are used for learning OIE extraction rules. The assumption is that the KB relation and the syntactic pattern instance (i.e., the *open* relation) express same information. Other OIE methods exploit the DSA similarly, including OLLIE (Mausam et al., 2012), ReNoun, (Yahya et al., 2014), NestIE (Bhutani et al., 2016), BONIE (Saha et al., 2017) and OpenCeres (Lockard et al., 2019).

2.1 KB Hits

For some OIE triples with disambiguated arguments, there are corresponding KB facts with the same argument pairs. Gashteovski et al. (2019) used this principle—which they called *KB hit*—to roughly measure the amount of OIE triples for which there is information in a reference KB. In particular, consider an OIE triple *(s, r_{open}, o)* where s, o are disambiguated and r_{open} is the open relation of the triple. Then, if there is at least one KB fact such that (s, r_{KB}, o) or (o, r_{KB}, s)—where r_{KB} is a KB relation—we say that the OIE triple

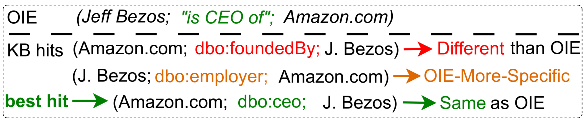

OIE (Jeff Bezos; "is CEO of"; Amazon.com)
KB hits (Amazon.com; dbo:foundedBy; J. Bezos) → Different than OIE
(J. Bezos; dbo:employer; Amazon.com) → OIE-More-Specific
best hit → (Amazon.com; dbo:ceo; J. Bezos) → Same as OIE

Figure 1: Hit categories indicate semantic relatedness between OIE triple and its KB hits.

has a KB hit. Note that one OIE triple may have more than one KB hit.

A single KB hit indicates an OIE triple for which a KB fact exists. However, it says nothing about *how* the OIE triple and the KB fact are semantically related. The DSA goes a step further and indicates semantic relatedness: if there is an OIE triple with a KB hit, then the OIE triple expresses the information of the KB triple. We study the semantic relatedness between an OIE triple and its KB hit using four *hit categories: Same, OIE-More-Specific, KB-More-Specific* and *Different*.

Same: OIE triple and KB fact are semantically equivalent, i.e., they express the same information. In Fig. 1, the OIE triple *(Jeff Bezos; "is CEO of"; Amazon.com)* expresses the same information as the KB fact (Amazon.com; dbo:ceo; Jeff Bezos).

OIE-More-Specific: OIE triple is semantically more specific than the KB fact, i.e., it expresses the KB fact along with additional information not present in the KB fact. In Fig. 1, the OIE triple is more specific than the KB hit (Jeff Bezos; dbo:employer; Amazon.com), because the OIE triple implies the KB fact and additionally expresses that Jeff Bezos is a CEO.

KB-More-Specific: KB fact is semantically more specific than the OIE triple, i.e., it expresses the OIE triple along with additional information not present in the OIE triple; e.g., OIE triple *(Angela Merkel; "is politician from"; Germany)* and its KB hit (A. M.; dbo:chancellor; Germany). Contrary to *OIE-More-Specific*, KB relations in such cases cannot be inferred from the OIE triple.

Different: OIE triple is semantically different than the KB fact, i.e., it expresses conceptually different information than the KB fact. Such KB hits cannot be compared in terms of more-general or more-specific relatedness. In Fig. 1, the KB hit (Amazon.com; dbo:foundedBy; Jeff Bezos) expresses different information w.r.t. the OIE triple, because *CEO* and *founder* are different concepts.

In case there are several KB hits for one OIE triple, each KB hit is assigned a separate category (Fig. 1). We assign only one label—*best hit*—describing the best possible semantic relatedness of the OIE triple w.r.t. all KB hits. Particularly, the best hit label is the first hit-category label appearing in the following order: *Same, OIE-More-Specific, KB-More-Specific* and *Different* (e.g., in Fig. 1, the best hit is *Same*). In our evaluation, we consider the best hits only, because we are interested in the best possible alignment of OIE triples with KB facts.

2.2 Study Design

The goal of the study is to *investigate the limits* of semantic relatedness between OPIEC triples and DBpedia facts having the same arguments. We study this by investigating what can be achieved if: (1) arguments are correctly disambiguated, (2) OIE triples are correctly extracted, (3) OIE relations are disambiguated. To this end, we used a subset of OPIEC-Linked: the largest OIE corpus to date, having 6M OIE triples with disambiguated arguments. Since we focus only on OIE triples that are correctly extracted, we filtered out triples from OPIEC-Linked that we found to be noisy, which left us with around 3M triples.[2] Details about the data are discussed in Appendix A.

Next, we constructed a random sample of 100 correctly extracted OIE triples from OPIEC which also have KB hits in DBpedia. We show to human annotator the OIE triple and the relevant KB hit information: 1) KB hits: every possible KB hit; 2) KB types: to assure the labeler that the types of the OIE triple's arguments match the domain/range constraints of the KB relation counterpart; 3) KB relation information: domain, range, description, etc., to help the labeler understand the exact semantics of the KB relation. Each KB hit of the OIE triple was labeled with one of the four hit categories. For each OIE triple, we keep the label of the best hit.

We split the OPIEC data into two subsets—*All relations* and *Is-a relation*—which are studied separately. The reason is that we have a substantial amount of triples having *Is-a relation* form *(subject; "be"; object)*, which express types; e.g., *(Berlin; "be"; City)*. We treat such triples differently to evaluate how the type information extracted from OIE compares with current KB information. The subset *All relations* are all OPIEC triples except the triples with *Is-a relation*. Both sub-studies follow the procedure explained in the previous two paragraphs.

[2] In the remainder of the paper, we refer to this dataset as OPIEC for simplicity

2.3 Experimental Results and Discussion

All-relations. We observed that in 88% of the cases, the OIE triple from OPIEC is able to semantically express its best hit KB fact from DBpedia (Fig. 2a). However, in almost half of these cases (40% of all triples) the OIE triple is more specific, meaning that it expresses the information of the KB fact along with additional information. Consider the OIE triple: *(All We Grow; "be debut album of"; S. Carey)* and its KB hit (All We Grow; dbo:artist; S. C.), whereas the OIE triple expresses the information of the KB hit fact, but it contains additional information about the album. In 12% of the cases, the OIE triple is not able to express its best hit; i.e., either the KB triple is more specific—meaning, the KB triple cannot be inferred by the OIE triple—or the semantics of the OIE triple is different than the semantics of the KB fact. More precisely, in 7% of the cases the OIE triple is more generic than its KB hit; e.g., OIE triple *(Rhacophorus annamensis; "be species of"; Frog)* and KB hit (R. a.; dbo:order; Frog). Judging from the OIE triple only, it is not enough to infer the relation between the two entities (e.g., order, genus, kingdom, etc.). Finally, 5% of the triples have different semantics than their KB hit; e.g., *(Saab Automobile; "test V8 in"; Saab 99)* v.s. (Saab 99; dbo:manufacturer; Saab A.).

Is-a relation. We observed that OIE triples with *Is-a* relation are more specific than DBpedia types in roughly 2/3 of the cases (Fig. 2b). In only 1/3 of the OIE triples, the KB contains an equivalent type. There are almost no cases where either the OIE type is more generic than the KB type, nor when they are different. This suggests that OIE triples with *Is-a relation* can provide more fine-grained types for the KB; e.g. OIE triple *(Tony Blair; "be"; Prime minister)* is more fine-grained than DBpedia type (Tony Blair; type; OfficeHolder). From the type "Prime minister" one can infer the type "OfficeHolder", but not the other way around.

2.4 Qualitative Study

Argument type information within the OIE relation is frequent reason for an OPIEC triple to be more specific than its KB hit in DBpedia. In particular, the details in the relation refer to more fine grained types for the argument(s); e.g., OIE triple *(Strul; "is Swedish film directed by"; Jonas Frick)* and its KB hit (Strul; dbo:director; Jonas Frick). The type available for Strul is "film". If there was a type "Swedish film" in DBpedia, then this alignment would have been equivalent. For the cases where the OIE triple represents different information than the KB hit, we found that usually the information on both sides is somehow semantically related; e.g., OIE triple *(s; "be CEO of"; o)* and its KB hit (s; dbo:founder; o). In this example, "CEO" and "founder" are related concepts, but they are semantically different.

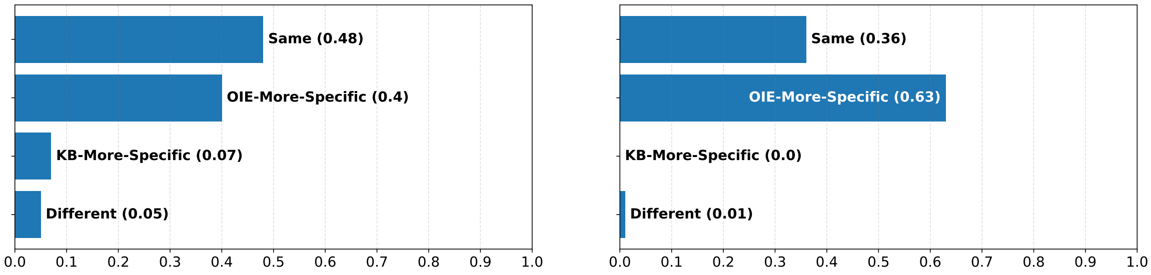
3 Expressibility of OPIEC triples with DBpedia

In the previous section, we evaluated the limits of aligning OIE triples from OPIEC for which KB facts exist in DBpedia. Such cases, however, comprise only $\frac{1}{4}$ of the data. In this section, we evaluate all cases: the limits of aligning *any* OPIEC triple with a DBpedia fact. Our goal is to answer the question of whether any OPIEC triple contains information which is relevant for DBpedia and, if so, *how* can it be expressed with the KB. We measured relevance by quantifying the information that can be expressed with KB language and we study how can such information be expressed.

3.1 One Triple Assumption

Many methods use large-scale outputs of OIE systems for downstream tasks by trying to express *one* OIE triple with *one* KB fact. This includes mapping open relations to a KB relation for improving slot filling (Soderland et al., 2013, 2015; Angeli et al., 2015; Yu et al., 2017), unifying open relations into a single KB schema (Bovi et al., 2015), canonicalizing open relations into relational synsets that are mapped to a KB relation (Galárraga et al., 2014), mapping open relations to lexical KBs (Grycner and Weikum, 2014), and mapping OIE triples to KB facts (Soderland et al., 2010; Zhang et al., 2019; Putri et al., 2019) which are used for KB population (Soderland et al., 2013; Dutta et al., 2013, 2015). Such methods implicitly make the *One Triple Assumption (OTA): "Any OIE triple can be expressed with one KB fact"*. For example, the OIE triple *(Emmanuel Macron; "be president of"; France)* can be expressed with the KB relation dbo:president: (Emmanuel Macron; dbo:president; France). Note that such mapping is possible even if this particular instance does not exist in the KB (e.g. in DBpedia, there is no KB fact stating that Emmanuel Macron is president of France).

Sometimes, an OIE triple cannot be expressed by a single KB fact, but it can be expressed by multiple



(a) *All relations*

(b) *Is-a relation*

Figure 2: Semantic relatedness between OIE triples from OPIEC and their DBpedia hits

KB facts or a first-order logic KB formula. Consider the OIE triple *(J. F. Kennedy; "be grandchild of"; P.J. Kennedy)*. This triple can be represented with the KB formula $\exists x$: (J. F. K.; dbo:parent; x) $\wedge$ (x; dbo:parent; P. J. K.), because there is no KB relation expressing "grandchild" relationship between two entities. Similarly, Das et al. (2016) use multi-hop reasoning between two entities in a KB to infer new relations, while Fu et al. (2019) do multi-hop reasoning over OIE data.

3.2 Expressibility Levels

To understand the semantic expressibility of an OIE triple w.r.t. KB facts, we differentiate three possible expressibility levels: *Fully-Expressible, Partly-Expressible* or *Not-Expressible*.

Fully-Expressible: Semantics of an OIE triple can be completely expressed with one KB fact; e.g., OIE triple *(E. Schmidt; "be chairman of"; Google)* and KB fact (Google; dbo:chairman; E. S.).

Partly-Expressible: The semantics of an OIE triple can be partly expressed with one KB fact, i.e. the OIE triple contains additional information which is not present in the KB fact. For example, the OIE triple *(Steffi Graf; "defeated"; Natasha Zvereva)* is *Partly-Expressible*, because there is no KB relation about one athlete defeating another; though it can be partly expressed with the KB fact (Steffi Graf; dbo:opponent; Natasha Zvereva).

Not-Expressible: The semantics of an OIE triple cannot be expressed with one KB fact, i.e. it is neither *Fully-Expressible* nor *Partly-Expressible*. For example, the OIE triple *(IBM; "has Color Paint for"; IBM PCjr)* cannot be expressed with a single KB fact, because the KB does not possess schemas for expressing such information in a single fact.

We make use of the above-defined expressibility levels to understand the semantic expressibility of an OIE triple w.r.t. KB formulas as well. For example, the OIE triple *(IBM; "has Color Paint for"; IBM PCjr)* is *Not-Expressible* w.r.t. a single KB fact, but it is *Fully-Expressible* w.r.t. KB formulas, because we can represent that OIE triple with the KB formula: (IBM; dbo:product; Color Paint) $\wedge$ (Color Paint; dbo:computingPlatform; IBM PCjr), i.e., two KB facts.

3.3 Study Design

The goal of the study is to evaluate whether the information found in any OPIEC triple is relevant for DBpedia. We do this by measuring the amount of OIE information which can be expressed with KB language and we study the different levels of expressibility. We constructed a random sample of 100 correctly extracted OIE triples from OPIEC with disambiguated arguments and an expert labeler evaluated the expressibility level for each OIE triple w.r.t. DBpedia fact and w.r.t. KB formula. First, we measured in how many cases an OPIEC triple can be expressed with *one* DBpedia fact fully or partially. Then, when the OPIEC triple is *Partly-Expressible* (or *Not-Expressible*), we investigate if it can become *Fully-Expressible* (or *Partly-Expressible/Fully-Expressible*) with a KB formula.

3.3.1 Expressibility of OPIEC Triple with a Single DBpedia Fact

Each OIE triple is presented to a human annotator along with: 1) argument types from DBpedia of the OIE triple, 2) a list of candidate DBpedia relations, 3) relevant information about the candidate DBpedia relations (descriptions, domain/range types, ...), 4) all other relevant information from DBpedia. The question asked was *"Can the OIE triple be expressed with one KB fact?"* Given all KB information available, the human annotator then assigned one of the three possible labels: *Fully-Expressible, Partly-Expressible* and *Not-Expressible*. Note that the assumption here is that we have a *perfect mapping* from OIE triple to KB fact. Thus, the labeler assigns the best possible mapping as a final label. The goal is to evaluate—given perfect mapping—the expressibility of an OIE triple via KB fact.

The KB relation candidates are generated by two methods: *KB hit counts* (aggregates *hit relations*) and *any relation* (uses any DBpedia relation satisfying the type constraints of the OIE arguments).

Hit relation. When possible, we aligned every OPIEC triple to DBpedia via KB hit statistics. In previous step, for every open relation, we counted the corresponding KB relations obtained from the KB hits. The counts are sorted in descending order.

Any relation. For aligning the OIE triples to KB facts, it is important that we go beyond the KB hits statistics, because such statistically-based methods are useful only for frequent open relations. To this end, we generate more candidates by additionally considering the DBpedia relations that fit the domain/range constraints imposed by the types of the OPIEC triple. In case the DBpedia types for the OIE arguments themselves are wrong or missing, the labeler corrects them with the appropriate DBpedia type. With this strategy, we ensure that we show every possible fitting candidate to the labeler.

3.3.2 Expressibility of OPIEC Triple with KB Formula

Since for many cases OPIEC triples cannot be fully expressed with a single DBpedia fact, an expert labeler manually generated KB formulas (when possible) that switch the expressibility level from *Partly-Expressible* to *Fully-Expressible* or from *Not-Expressible* to *Fully/Partly-Expressible*. For example, the OIE triple *(Garrett Davis; "is Representative from"; Kentucky)* is *Partly-Expressible* with DBpedia fact (G. D.; dbo:region; Kentucky), but it is fully expressible with KB formula: (G. D.; dbo:profession; State rep.) $\wedge$ (G. D.; dbo:state; K.).

3.4 Results and Discussion

If we consider only the hit relation candidates (light-blue bars, Fig. 3a), only 29% of the OPIEC triples can be fully expressed with a single DBpedia fact; partly expressed another 29%; and 42% of the OPIEC triples cannot be expressed at all. This suggests that KB hit relation counts contain signal for KB expressibility, but not enough to express all OPIEC triples. The main reason is that KB hit relation counts work well only for the triples having open relations with high frequency. Higher frequency of an open relation implies higher likelihood for a KB hit, thus higher likelihood for capturing the semantic content (fully or partially) of an OIE triple by one of the candidates. Many open relations are not frequent enough, which is the main reason why in 42% of the OIE triples the KB hit relations are not enough to express the OIE triple.

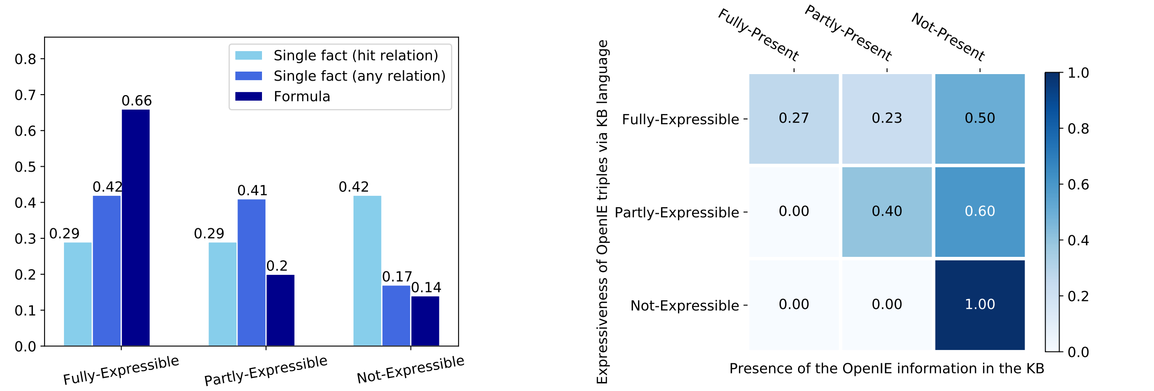
When we extend the limits of the candidates by including any DBpedia relation which respects the constraints of the argument types (represented as blue bars in the middle), then we significantly reduce the amount of OPIEC triples which cannot be expressed with one DBpedia fact (from 42% down to 17%). More precisely, 42% of the triples can be fully expressed and 41% can be partly expressed with one DBpedia fact. This study shows that most of the OPIEC triples are relevant for DBpedia, because more than 80% of them can be expressed with a single DBpedia fact. However, we observed that nearly half of these cases are only partly expressible, since the OPIEC triples contain additional details which cannot be expressed in DBpedia. The reason for this is because KB relations have very strict semantics, while open relations have the expressibility of natural language.

When we introduce KB formulas, the expressibility of the OPIEC triples is significantly improved. We observed that the number of OPIEC triples that can be fully expressed with DBpedia increased from 42% to 66%. This was mostly on the expense of the cases where an OPIEC triple is partly expressible w.r.t. DBpedia fact (it reduced these cases from 41% to 20%). Less significantly, the KB formulas allowed for some OPIEC triples that are not expressible via DBpedia to be expressible (percentage went down from 17% to 14%). We illustrate the expressibility of OIE triples with KB formulas in a few examples in Tab. 1.

Finally, 14% of the OPIEC triples cannot be expressed with DBpedia vocabulary neither with single DBpedia facts nor with KB formulas. Reasons include cases with tertiary relations or open relations which can be expressed via natural language, but not with DBpedia; e.g. OIE triple *(X; "sponsored"; Y)* cannot be expressed with DBpedia.

3.5 New Information for DBpedia

To evaluate how much of the OPIEC information is new for DBpedia, we used the OPIEC triples from the expressibility study. Based on the information content of the OIE triple and the content of DBpe-



(a) Can an OPIEC triple be expressed in DBpedia? (b) Does DBpedia contain information from the OPIEC triple?

Figure 3: Expressibility and presence of OIE information within the KB

#	OIE triple	KB formula
t_1	Temporal annotation	
	(Coral Fang; "was released by"; Sire Records) *Time: (in, 2003)*	(Coral Fang; dbo:recordLabel; Sire Records) ∧ (Coral Fang; dbo:releaseDate; 2003)
t_2	Complex formula	
	(Garrett Davis; "was Rep. from"; Kentucky)	(G. D.; dbo:profession; State representative) ∧ [(G. D.; dbo:region; K.) ∨ (G. D.; dbo:state; K.)]
t_3	Existential quantification	
	(Franz Liszt; "transcribed piece for"; Piano solo)	$\exists x$: (F. L.; dbo:write; x) ∧ (x; dbo:genre; P. solo)
t_4	Conjunctive formula	
	(Dick Ket; *"was Dutch magic realist painter noted for";* *Still life)*	(Dick Ket; dbo:nationality; Netherlands) ∧ (Dick Ket; dbo:genre; Magic realism) ∧ (Dick Ket; dbo:occupation; Painter) ∧ (Dick Ket; dbo:knownFor; Still life)

Table 1: Selected examples of OPIEC triples expressed with KB formulas

dia, an expert labeled each OIE triple with one of the three possible options: 1) completely present in the KB: there exists DBpedia fact or formula which fully expresses the OIE triple; 2) partly present in the KB: there exists DBpedia fact or formula which partly expresses the OIE triple; 3) not present in the KB: there is no existing DBpedia fact or formula which can fully or partially express the OIE triple.

We found that in 59% of the OIE triples the content is not present in DBpedia at all, in 23% is partly present and only in 18% it is fully present. This suggests that most OIE triples contain information which is either not present or not fully present in DBpedia. To investigate new *relevant*[3] information for DBpedia, we compared expressibility w.r.t. presence of the OIE information content in DBpedia (Fig. 3b). In general, we observed that most OIE information that is relevant for the KB is either not present or only partly present in DBpedia, showing the potential of such triples for downstream tasks such as KB population (Lin et al., 2020).

4 Transferability

In this section, we study whether and to what extent the results of the evaluations transfer to other OIE systems. We used three other popular OIE systems: Stanford OIE (Angeli et al., 2015), RnnOIE (Stanovsky et al., 2018) and OpenIE 5 (Saha et al., 2018). We ran these OIE systems on the provenance sentences of the sampled triples used in our evaluations—explained in Sec. 2 and Sec. 3—and compared their outputs to their OPIEC counterparts (extracted with MinIE (Gashteovski et al., 2017)).

Consider the OPIEC triple *(Turf Buccaneers; "be album by"; Mac Dre)*. We use the provenance sentence from which this triple was extracted and we run the other OIE systems on the same sentence. Then, we select the triples that match the argument pair of the OPIEC triple; i.e. *(Turf Buccaneers, Mac Dre)* in the example. As before, we keep only the triples that are correctly extracted. Then, an anno-

[3] OIE triple: *relevant* for KB if it is expressible in that KB

Label	Stanford	OpenIE 5	RnnOIE	All
	OIE triples and KB facts with same args.			
Hit category	0.98	0.84	0.77	0.86
	Expressibility of OIE triples with DBpedia			
Single fact (hit relation)	0.92	0.93	0.85	0.90
Single fact (any KB rel.)	0.88	0.86	0.77	0.84
KB formula	0.98	0.89	0.89	0.92

Table 2: Label equivalence ratio of the evaluations: labels from OPIEC triples (produced by MinIE) v.s. labels from triples produced by other OIE systems. *All* column considers all the labels for the triples produced by the other OIE systems combined.

tator evaluates these OIE triples w.r.t. DBpedia for either hit category (Sec. 2) or expressibility level (Sec. 3). Finally, we compared these labels with the original labels of our evaluations for OPIEC.

In Sec. 4.3 we examine to what extent other OIE systems extract different entities (and entity pairs) than MinIE, given the same provenance sentences used in our study. In particular, we measure how many entities each OIE system extracts in general, how similar they are w.r.t. the entities extracted by MinIE, and to what extent the entities extracted by other OIE systems are also extracted by MinIE. Such study is important for evaluating whether other systems extract different entities, which would influence the findings of our study.

4.1 Hit Categories

We compared the newly assigned labels for hit categories from the other OIE systems with the original labels of our evaluation presented in Sec. 2 (Tab. 2). Overall, we found that in 86% of the cases the labels were equivalent (*label equivalence ratio*). In most cases for which there was a mismatch of the labels, the OPIEC triple has same semantics as the KB fact, while the triple by the other OIE system is more specific than the KB fact. Hence, when moving to other OIE systems, one should expect that they may produce more specific OIE triples.

We also observed the label equivalence ratio of the OPIEC triples w.r.t. the other OIE systems individually (Tab. 2). We found that RnnOIE has lowest label equivalence ratio (77%). Again, the main reason for mismatch is because RnnOIE extracts more specific triples than MinIE. This is because the goal of MinIE is to produce shorter extractions, while RnnOIE does not aim at reducing the length of the extractions, thus producing more specific triples. Consequently, in many cases MinIE extracts triple having same semantics as the KB fact, while RnnOIE extracts a more specific triple.

On the other hand, Stanford OIE produced triples that have almost the same labels as OPIEC (98% of the labels are equivalent). The reason is that Stanford OIE was constructed with the slot filling task in mind, which results in producing shorter extractions (same goal as MinIE). Therefore, the specificity levels with MinIE are similar. OpenIE 5 is in between: it produces more specific triples than Stanford and less specific triples than RnnOIE.

4.2 Expressibility Levels

Following the same strategy as the labels for hit categories, we compared the newly assigned labels for expressibility levels from other OIE systems with the original labels of our evaluation for expressibility of OIE triples with DBpedia (Sec.3). We compared the labels for single fact (hit relation), single fact (any KB rel.) and KB formula (Tab. 2).

Our findings for the expressibility levels are similar to the findings discussed in Sec. 4.1. Overall, we found that the label equivalence ratio of the OPIEC triples and the triples produced by other OIE systems is relatively high. Again, most mismatches are because other OIE systems tend to produce more specific triples. Consequently, in such cases, when MinIE extracts triple that is *Fully-Expressible*, the other OIE systems extract triple that is *Partly-Expressible* with the KB.

4.3 Extracted Entities

To compare the entities extracted by MinIE with the entities extracted by the other OIE systems, we used the same provenance sentences from OPIEC's triples used in our studies (Sec. 2 and 3). From them, we extracted OIE triples with MinIE and the other OIE systems. Again, we kept only the triples generated by all systems that contain disambiguated arguments on both the subject and the object. We did not consider triples that contain more than one entity link per argument (e.g. some systems generate whole clauses as an object, which may contain more than one entity). For such extractions, it is not clear to which entity the argument is referring to. Finally, for each entity and entity pair, we computed counts, Jaccard distance w.r.t. MinIE and coverage by MinIE (Tab. 3).

For both the DSA and OTA sentences, we observed that MinIE extracts more arguments (and

DSA / OTA	*Entities*			*Entity pairs*		
	Count	Jaccard w.r.t. MinIE	Coverage by MinIE	Count	Jac. w.r.t. MinIE	Cov. by MinIE
MinIE	272 / 235	1.0 / 1.0	1.0 / 1.0	221 / 169	1.0 / 1.0	1.0 / 1.0
Stanford	156 / 120	0.44 / 0.45	0.83 / 0.92	99 / 80	0.23 / 0.29	0.61 / 0.70
OpenIE 5	70 / 81	0.21 / 0.32	0.83 / 0.95	38 / 47	0.11 / 0.21	0.68 / 0.81
RnnOIE	49 / 69	0.15 / 0.27	0.84 / 0.94	27 / 41	0.07 / 0.17	0.63 / 0.73

Table 3: Extracted entities by MinIE and other OIE systems for both studies: DSA (Sec. 2) / OTA (Sec. 3).

argument pairs) than other OIE systems. This is consistent with the findings of Gashteovski et al. (2017), where the authors report high recall for MinIE. Moreover, Lin et al. (2020) reported that MinIE extracts entities that are easier to disambiguate to KBs compared to other OIE systems, which is another reason why the number of extracted entities is lower in other systems.

Because of the lower amount of entities extracted by the other systems, the Jaccard distance between the entities extracted by MinIE and other systems is relatively low. If we turn to coverage by MinIE, however, we observed that most entities extracted by other OIE systems are also extracted by MinIE. This suggests that the extractions made by other OIE systems that are relevant for KBs were likely going to be extracted by MinIE as well. Based on these results, we conjecture that the findings of our study largely transfer over to other OIE systems.

4.4 Discussion

Overall, we found that OIE triples produced by other OIE systems tend to have very similar hit categories (as well as expressibility levels) with the OPIEC triples. Due to the fact that MinIE—OPIEC's underlying OIE system—is designed to produce less specific extractions, we observed that if one uses other OIE systems, it should be expected the extractions to be more specific. This, in turn, results in 1) producing larger fraction of triples that are more specific than the KB triple with the same argument pair; 2) producing larger fraction of triples that are *Partially-Expressible*.

As for the entities extracted by MinIE and other OIE systems, we found that MinIE extracts most of the entities that are extracted by the other OIE systems as well as additional entities. The reason for this observation is the high recall of MinIE as well as the compact extractions made by MinIE that contribute to extracting more KB-centric entities.

A limitation of this study is that we focus on the most common form of OIE extractions: OIE triples. Some OIE systems extract more complex structures—e.g. nested extractions (Bhutani et al., 2016)—which are not covered by this paper and require a separate study.

5 Main Findings and Conclusions

In this paper, we evaluated how OIE triples from the OPIEC corpus are related to the DBpedia KB w.r.t. information content. Both resources are automatically generated from same domain—OPIEC from the textual data and DBpedia from the semi-structured data of Wikipedia—which makes them compatible resources for evaluation.

First, we evaluated the semantic relatedness between OIE triples from OPIEC and DBpedia facts having the same arguments; i.e. the Distant Supervision Assumption (DSA). Such cases are important for downstream tasks and for bootstrapping OIE systems. In general, we found that such OPIEC triples are semantically related to the DBpedia facts, but quite often the OIE triples are more specific, thus capturing more information than the KB facts. Second, we evaluated the expressibility of any OPIEC triple w.r.t. DBpedia: whether (and *how*) an OPIEC triple can be expressed with a *single* DBpedia fact; i.e. the One Triple Assumption (OTA). OTA is an implicit assumption used in tasks such as slot filling. We found that expressing an OPIEC triple with a single DBpedia fact is often limited, and that the use of KB formulas improves the expressibility of OIE triples with KB language significantly. Third, we found that most OPIEC triples contain information which is not present in DBpedia, thus showing the potential of OIE triples for tasks such as KB population. Finally, we found that the findings of our case study—which was based on the OIE system MinIE—are likely to transfer over to other OIE systems.

Our study suggests that most information found in OIE triples is not present in the KB. One way to harvest such knowledge is to add OIE triples to a KB with universal schema (Riedel et al., 2013). We leave such evaluations for future work.

References

Gabor Angeli, Melvin Jose Johnson Premkumar, and Christopher D. Manning. 2015. Leveraging Linguistic Structure for Open Domain Information Extraction. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 344–354.

Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. DBpedia: A Nucleus for a Web of Open Data. In *The Semantic Web*, pages 722–735.

Michele Banko, Michael J Cafarella, Stephen Soderland, Matthew Broadhead, and Oren Etzioni. 2007. Open Information Extraction from the Web. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2670–2676.

Michele Banko and Oren Etzioni. 2008. The Tradeoffs between Open and Traditional Relation Extraction. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 28–36.

Nikita Bhutani, HV Jagadish, and Dragomir Radev. 2016. Nested Propositions in Open Information Extraction. In *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 55–64.

Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. 2009. DBpedia - a Crystallization Point for the Web of Data. *Journal of Web Semantics*, 7(3):154–165.

Claudio Delli Bovi, Luis Espinosa Anke, and Roberto Navigli. 2015. Knowledge Base Unification via Sense Embeddings and Disambiguation. In *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 726–736.

Samuel Broscheit, Kiril Gashteovski, and Martin Achenbach. 2017. OpenIE for Slot Filling at TAC KBP 2017-System Description. In *Proc. of the Text Analysis Conference (TAC)*.

Samuel Broscheit, Kiril Gashteovski, Yanjie Wang, and Rainer Gemulla. 2020. Can We Predict new Facts with Open Knowledge Graph Embeddings? A Benchmark for Open Link Prediction. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 2296–2308.

Lei Cui, Furu Wei, and Ming Zhou. 2018. Neural Open Information Extraction. *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 407–413.

Rajarshi Das, Arvind Neelakantan, David Belanger, and Andrew McCallum. 2016. Incorporating Selectional Preferences in Multi-hop Relation Extraction. In *Proc. of the Workshop on Automated Knowledge Base Construction (AKBC)*, pages 18–23.

Xin Dong, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. 2014. Knowledge Vault: A Web-scale Approach to Probabilistic Knowledge Fusion. In *Proc. of the International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 601–610.

Arnab Dutta, Christian Meilicke, Mathias Niepert, and Simone Paolo Ponzetto. 2013. Integrating Open and Closed Information Extraction: Challenges and First Steps. In *Proc. of Workshop NLP-DBPEDIA@ISWC*.

Arnab Dutta, Christian Meilicke, and Heiner Stuckenschmidt. 2015. Enriching Structured Knowledge with Open Information. In *Proc. of International Conf. on World Wide Web (WWW)*, pages 267–277.

Cong Fu, Tong Chen, Meng Qu, Woojeong Jin, and Xiang Ren. 2019. Collaborative Policy Learning for Open Knowledge Graph Reasoning. *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2672–2681.

Luis Galárraga, Geremy Heitz, Kevin Murphy, and Fabian M. Suchanek. 2014. Canonicalizing Open Knowledge Bases. In *Proc. of the International Conference on Information and Knowledge Management (CIKM)*, pages 1679–1688.

Kiril Gashteovski, Rainer Gemulla, and Luciano Del Corro. 2017. MinIE: Minimizing Facts in Open Information Extraction. In *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2630–2640.

Kiril Gashteovski, Sebastian Wanner, Sven Hertling, Samuel Broscheit, and Rainer Gemulla. 2019. OPIEC: An Open Information Extraction Corpus. *In Proc. of the Conference on Automated Knowledge Base Construction (AKBC)*.

Fabrizio Gotti and Philippe Langlais. 2019. Weakly Supervised, Data-Driven Acquisition of Rules for Open Information Extraction. In *Proc. of the Canadian Conference on Artificial Intelligence (CCAI)*, pages 16–28.

Adam Grycner and Gerhard Weikum. 2014. HARPY: Hypernyms and Alignment of Relational Paraphrases. In *Proc. of the International Conference on Computational Linguistics (COLING)*, pages 2195–2204.

Swapnil Gupta, Sreyash Kenkre, and Partha Talukdar. 2019. CaRe: Open Knowledge Graph Embeddings. In *Proc. of the Conference on Empirical Methods in Natural Language Processing and the International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 378–388.

Amina Kadry and Laura Dietz. 2017. Open Relation Extraction for Support Passage Retrieval: Merit and Open Issues. In *Proc. of International Conference on Research and Development in Information Retrieval (SIGIR)*, pages 1149–1152.

Keshav Kolluru, Samarth Aggarwal, Vipul Rathore, Soumen Chakrabarti, et al. 2020. IMoJIE: Iterative Memory-Based Joint Open Information Extraction. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 5871–5886.

Xueling Lin, Haoyang Li, Hao Xin, Zijian Li, and Lei Chen. 2020. KBPearl: A Knowledge Base Population System Supported by Joint Entity and Relation Linking. In *Proc. of the Very Large Data Base Endowment (PVLDB)*, pages 1035–1049.

Colin Lockard, Prashant Shiralkar, and Xin Luna Dong. 2019. OpenCeres: When Open Information Extraction Meets the Semi-Structured Web. In *Proc. of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 3047–3056.

Mausam, Michael Schmitz, Robert Bart, Stephen Soderland, and Oren Etzioni. 2012. Open Language Learning for Information Extraction. In *Proc. of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 523–534.

Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant Supervision for Relation Extraction without Labeled Data. In *Proc. of the Joint Conference of the Annual Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing (ACL-IJCNLP)*, pages 1003–1011.

Federico Nanni, Jingyi Zhang, Ferdinand Betz, and Kiril Gashteovski. 2019. EAL: A Toolkit and Dataset for Entity-Aspect Linking. In *Proc. of the Joint Conference on Digital Libraries (JCDL)*.

Harinder Pal and Mausam. 2016. Demonyms and Compound Relational Nouns in Nominal Open IE. In *Proc. of the Workshop on Automated Knowledge Base Construction (AKBC@NAACL-HLT)*, pages 35–39.

Rifki Afina Putri, Giwon Hong, and Sung-Hyon Myaeng. 2019. Aligning Open IE Relations and KB Relations using a Siamese Network Based on Word Embedding. In *Proc. of the International Conference on Computational Semantics (ICCS)*, pages 142–153.

Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M Marlin. 2013. Relation Extraction with Matrix Factorization and Universal Schemas. In *Proc. of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 74–84.

Swarnadeep Saha, Harinder Pal, and Mausam. 2017. Bootstrapping for Numerical Open IE. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 317–323.

Swarnadeep Saha et al. 2018. Open Information Extraction from Conjunctive Sentences. In *Proceedings of the International Conference on Computational Linguistics (COLING)*, pages 2288–2299.

Stephen Soderland, John Gilmer, Robert Bart, Oren Etzioni, and Daniel S. Weld. 2013. Open Information Extraction to KBP Relations in 3 Hours. In *Proc. of the Text Analysis Conference (TAC)*.

Stephen Soderland, Natalie Hawkins, John Gilmer, and Daniel S. Weld. 2015. Combining Open IE and Distant Supervision for KBP Slot Filling. In *Proc. of the Text Analysis Conference (TAC)*.

Stephen Soderland, Brendan Roof, Bo Qin, Shi Xu, Mausam, and Oren Etzioni. 2010. Adapting Open Information Extraction to Domain-Specific Relations. *AI magazine*, 31(3):93–102.

Gabriel Stanovsky, Ido Dagan, et al. 2015. Open IE as an Intermediate Structure for Semantic tasks. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 303–308.

Gabriel Stanovsky, Julian Michael, Luke Zettlemoyer, and Ido Dagan. 2018. Supervised open information extraction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 885–895.

Fei Wu and Daniel S. Weld. 2010. Open Information Extraction using Wikipedia. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 118–127.

Mohamed Yahya, Steven Whang, Rahul Gupta, and Alon Halevy. 2014. ReNoun: Fact Extraction for Nominal Attributes. In *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 325–335.

Dian Yu, Lifu Huang, and Heng Ji. 2017. Open Relation Extraction and Grounding. In *Proc. of the International Joint Conference on Natural Language Processing (IJCNLP)*, pages 854–864.

Dongxu Zhang, Subhabrata Mukherjee, Colin Lockard, Xin Luna Dong, and Andrew McCallum. 2019. OpenKI: Integrating Open Information Extraction and Knowledge Bases with Relation Inference. *Proc. of the Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT)*.

A Reference Corpora and Methodology

A.1 OIE Data and Methodology

OIE Corpus. One of the major problems of aligning OIE triples with KB facts is that the OIE triples are consisted of surface patterns, which makes the triples highly ambiguous. To make such alignments possible, it is necessary that the arguments of the OIE triples are disambiguated. For these reasons, we chose OPIEC-Linked (Gashteovski et al., 2019) as an OIE corpus for our study, because it is the biggest OIE corpus to date, containing 6M triples with disambiguated arguments. OPIEC-Linked was constructed by running the OIE system MinIE-SpaTe over the entire English Wikipedia. The links in the text added by Wikipedia authors were kept, which provides golden disambiguation links for the arguments.

OPIEC Filters. The goal of the study is to investigate the limits of aligning OIE triples with KBs. For this reason, we assume both a perfect extractor and perfect alignments between the OIE triples and the KB facts. To reduce the noise from OPIEC-Linked, we followed (Broscheit et al., 2020) and filtered out the triples having the following properties: 1) confidence score is less than 0.3; 2) extraction type is SVOO, SVOC or extractions are made from the *apposition* dependency parse relation. In a preliminary study, we found these triples to be very noisy. For the remainder of the paper, we will refer to this data as OPIEC for simplicity.

Sampling Correctly Extracted OIE Triples for the DSA Study. The DSA implies that for each (correctly extracted) OIE triple which has a KB-hit, the open relation expresses the same information as the KB relation. The goal of the study is to *investigate the limits* of such alignments, which is why we consider only extractions that are correctly extracted. For this reason, we constructed a random sample of 200 OIE triples from the OIE triples having KB-hits, which were labeled for correctness by an expert. To ensure that the information of the triple is complete, the triples which are not self-contained were labeled as "incorrectly extracted". For example, the triple *(Pope Clement VII; "named him inquisitor of"; Modena)* is not self-contained, because it is not clear to which entity "him" refers to. The labeler stopped at the 100th correctly extracted triple. Note that these 100 correctly extracted triples are also self-contained.

Study Design for *Is-a relation* OIE triples. The study for *Is-a relation* triples is similar with the one done on *All relations*. We sampled 100 correctly extracted triples from OPIEC-Typed (i.e. the subset of OPIEC containing triples of the form *(subj, "be"; obj)*). For each correctly extracted OPIEC-Typed triple, we matched the subject link with all the DBpedia entries for types. As a result, we have an OIE triple *(subj, "be"; obj)* and on the KB side we have (subj; type; T). The sampling and labeling logic is the same as the one explained in the previous paragraph.

A.2 DSA Study: KB Data and Methodology

Reference KB. For the alignments, it is very important that the KB contains the same information as the text corpus from which the OIE data was constructed (i.e. that both the OIE triples and the reference KB were automatically constructed from the same domain). This ensures that the information in the KB and the information content in the OIE triples is the same. In such settings, the OIE arguments have the same ID links as the KB entities, which makes the study comparable. For these reasons, we chose DBpedia (Auer et al., 2007) as a reference KB, because it is a well-established KB constructed from Wikipedia (the same resource from which OPIEC is constructed), and because it is the largest KB to date which is automatically constructed from Wikipedia. Prior work for aligning OIE triples with KB facts also exploited the combination of Wikipedia and DBpedia (Wu and Weld, 2010; Dutta et al., 2013, 2015; Yu et al., 2017; Gashteovski et al., 2019).

DBpedia-filtered. For our study, it is essential that both of the KB triple arguments are disambiguated. Therefore, from DBpedia, we filtered out any triples containing literals, abstracts, dates, etc. Many relations in DBpedia are extracted with generic infobox extraction. These KB relations tend to be noisier—sometimes even ambiguous—and they often lack important information describing the precise semantics of the KB relation (Bizer et al., 2009) (e.g. domain/range types or descriptions are often missing). For these reasons, we filtered out these KB triples as well. We retained only the triples that were extracted with mapping-based infobox extraction (i.e. with namespace `http://dbpedia.org/ontology`), because of their higher extraction quality and higher level of details they provide. This way, it is much clearer to an expert labeler to assess the alignments.