# Linking Knowledge Graphs across Languages with Semantic Similarity and Machine Translation

**John P. McCrae, Mihael Arčan, Paul Buitelaar**
Insight Centre for Data Analytics
National University of Ireland Galway
Galway, Ireland
john@mccr.ae, mihael.arcan@insight-centre.org,
paul.buitelaar@insight-centre.org

## Abstract

Knowledge graphs and ontologies underpin many natural language processing applications, and to apply these to new languages, these knowledge graphs must be translated. Up until now, this has been achieved either by direct label translation or by cross-lingual alignment, which matches the concepts in the graph to another graph in the target languages. We show that these two approaches can, in fact, be combined and that the combination of machine translation and cross-lingual alignment can obtain improved results for translating a biomedical ontology from English to German.

## 1 Introduction

Knowledge graphs, including large databases such as DBpedia [18], have found a wide range of use cases in many domains and in many languages. The applications that depend on these knowledge graphs frequently use natural language and thus to adapt them to new languages it is necessary to add labels for concepts in this language. Moreover, sometimes it is the case that there already exists a suitable knowledge graph created in that language and as such it would be preferable to align the two knowledge graphs. The task of cross-lingual alignment has generally required statistical machine translation (SMT) and for the most part has simply translated the vocabulary and applied a monolingual alignment, although some authors have suggested translating into a third 'pivot' language [34]. In most cases there is not a full and exactly similar counterpart to a knowledge graph on the Web, but there is a large amount of relevant multilingual data already on the Web and the ability to reuse this in machine translation has still not be well exploited [25].

In this paper, we propose a hybrid approach that combines dataset alignment techniques and ontology translation techniques in order to translate a dataset to a new language. In particular, we consider three cases: firstly, we use dataset alignment techniques to link the European Health Records ontology[1] with a highly multilingual resource, namely DBpedia to translate this ontology from English into German. Secondly, we apply the OTTO machine translation system for ontologies [1] to translate the labels of the ontology directly. Finally, we combine these two approaches showing a 30 point increase in BLEU score over the alignment model and 2 point increase over the machine translation model.

## 2 Related Work

The growth of semantic information published in recent years has been stimulated to a large extent by the emergence of linked data, which showed to be the best way to expose, share and connect data in a language-independent fashion on the Semantic Web [15]. Although the idea of the Semantic Web depends on language independent information access, most of the ontologies published are expressed in English, leading to a biased access to the Semantic Web. On the other hand, the information on the traditional Web can be found in different languages, hence is language-specific. For these reasons, approaches have to be put in place to interlink and lexicalise these resources on a multilingual basis. Therefore, language communities started publishing data in different languages, so that no matter in which language ontologies or documents are published, access to knowledge will be supported [5].

Nevertheless, multilingualism is one of the

[1] http://trajano.us.es/~isabel/EHR/

biggest challenges for the Semantic Web since it is important to access information in the language of the user's choice. Most of the previous related work tackled this problem by using multilingual lexical resources, like EuroWordNet or IATE [7, 8]. Their work focuses on the identification of the lexical overlap between the ontology labels and the multilingual resources, which guarantees a high precision but a low recall. Therefore, external translation services like *BabelFish*, *SDL FreeTranslation* tool or *Google Translate* were used to overcome this issue [11, 13]. Additionally, [11] and [23] performed ontology label disambiguation, where the ontology structure is used to annotate the labels with their semantic senses. Similarly, [24] show positive effects of different domain adaptation techniques, i.e., using Web resources as additional bilingual knowledge, re-scoring translations with Explicit Semantic Analysis and language model adaptation for automatic ontology translation.

The task on ontology alignment frequently builds on a label translation system and thus converts the task of finding a cross-lingual alignment to that of finding a monolingual alignment among translated labels [34]. [13] used *Google Translate* in their work to translate the labels nested within the context of a larger label. The authors stress the importance of the translation quality to generate high-quality matching results for monolingual matching tools but do not provide a deeper evaluation on the translation part. In their follow-up work [14], the authors illustrate a semantic oriented cross-lingual ontology mapping, with disambiguating labels on the target side. Their strategy is based on a manual processing of the ontology labels and the usage of the knowledge extracted from bilingual corpora. Additionally their approach also linguistically enriches the labels in the ontology. Their semantic-oriented cross-lingual ontology mapping approach (SOCOM) uses off-the-shelf machine translation systems, such as *Google Translate*. They use the ontology relations, if the ontology is rich in structure, otherwise, nesting approaches of near labels were used as contextual information. Nevertheless, they observe that resource constraints (granularity, size of ontologies) have an impact on the translation selection approach. [21] describe an instance-based interlinking method that mostly relies upon machine translation technology. Their task focuses on finding whether two occurrences in different languages refer to the same object. The authors evaluate the suitability of *Microsoft Translator* to interlink RDF resources, represented as text documents in English and Chinese. The results demonstrated that the method can identify most of the correct matches using minimum information in a resource description with precision over 98%. [27] demonstrates the synergy of Wikipedia and Princeton WordNet [12] structure as a wide-coverage multilingual ontology, called BabelNet. This resource is created by linking Wikipedia entries and Wordnet synsets together. Since the lexical knowledge is the key for understanding and decoding a continuously changing world, they use commercial translation systems to fill the lexical gaps.

There has also been research in the process of finding translingual semantic representations, starting with methods that merge parallel corpora to obtain a translingual representation by means of latent topic modelling methods such as Latent Semantic Analysis [10] and Latent Dirichlet Allocation [26]. These methods can be used to estimate the latent similarity between ontology labels in different languages, but until recently such approaches did not yield results that outperformed direct translation [32] in comparable tasks. In particular, recent methods such as Orientated Principle Component Analysis [32], Kernel Canonical Correlation Analysis [36, CCA] and Orthonormal Explicit Topic Analysis [22, ONETA] calculate a translingual representation by means of estimating the correlation between term frequencies in a document-aligned corpus and this has been shown to outperform dictionary-based machine translation [17]. Additionally, ontology label disambiguation was performed by [23], where the structure of the ontology along with existing multilingual ontologies was used to annotate the labels with their semantic senses. In a similar direction, [6] derived cross-lingual alignments between English and Chinese WordNet by means of a distributional similarity approach.

## 3 Methodology

### 3.1 Dataset linking

Dataset linking is the process of discovering equivalent entries between entities in two datasets and this is based on a process of identifying relative 'facets' of entities features that can then be

processed into a feature that represents the similarity between the facets of a candidate pair of entities. Many such features of these candidates are then combined into a single score representing the similarity of two entities, and finally a 'matching' algorithm is used to find the linking between the datasets. As an example, we may extract the label of each entity in a dataset (the 'facets'), and then compute similarity features of the labels such as edit distance, a set of such features are then combined with a function learnt using a typical supervised machine learning procedure to a similarity score. Finally, the optimal mapping is calculated according to some constraint such as the *bipartite* assumption, which states that no entity in a dataset is connected to more than one entity in the dataset, and this can be efficiently computed by algorithms such as the Hungarian Algorithm [20]. This workflow is realized by a tool called NAISC (Nearly Automatic Integration of Schema)[2]. NAISC can extract a number of 'facets' from a dataset including, labels of terms, descriptions of terms, and most similar label of direct or indirect superterms. Then for each pair of strings extracted this way we develop a number of features as follows:

- **Jaccard, Dice, Containment:** We consider the two strings both as a set of words and a set of characters and compute the following functions:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|},$$

$$D(A,B) = \frac{2|A \cap B|}{|A| + |B|},$$

$$C(A,B) = \frac{|A \cap B|}{\min(|A|,|B|)}.$$

- **Smoothed Jaccard:** Smoothed Jaccard is calculated only on the word level for the concatenated labels facet. It is calculated as follows:

$$J_\sigma(A,B) = \frac{\sigma(|A \cap B|)}{(\sigma(|A|) + \sigma(|B|) - \sigma(|A \cap B|))},$$

$$\sigma(x) = 1.0 - exp(-\alpha x).$$

  This is a variant of Jaccard that can be adjusted to distinguish matches on shorter texts; it tends to Jaccard at $\alpha \to 0$.

[2] 'Naisc' means 'links' in Irish and is pronounced 'nashk'

- **Length Ratio:** The ratio of the number of tokens in each sentence. For symmetry this ratio is defined as:

$$\rho(x,y) = \frac{\min(x,y)}{\max(x,y)}.$$

- **Average Word Length Ratio:** The average length of each word in the text are also compared as above.

- **Negation:** One if both texts or neither text contain a negation word ('not', 'never', etc.), zero otherwise.

- **Number:** One if all numbers (e.g, '6') in each text are found in the other, zero otherwise.

- **GloVe Similarity:** For each word in each text we extract the GloVe vectors [30] and calculate the cosine similarity between these words' vectors $v_i$ and the $v_j$ using cosine similarity as follows:

$$g(x,y) = \frac{1}{n} \sum_{i=1,\ldots n} \max_{j=1,\ldots n} \cos(v_i, v_j),$$

  where $n$ is the length of the first string and $m$ is the length of the second.

- **LSTM:** We calculate a similarity using the LSTM approach described by [35].

Having extracted the features for each pair, we learn the similarity using a regression SVM [31] and learn the optimal alignment using the Hungarian algorithm [20].

#### 3.1.1 Alignment as Machine Translation

Alignment cannot itself translate an ontology or knowledge graph, but we can use the alignment procedure to align the ontology we wish to translate to another resource, which has the relevant translations. In particular, we align to the DBpedia resource, which is derived from Wikipedia, and as such contains a large number of translations for concepts given by the interlingual links between pages. As such, if we can correctly select the DBpedia entity for an element in our input ontology, we can generate a translation for this term in the target language.

### 3.2 Translating labels with Machine Translation

A similar approach to localizing ontologies is translating domain-specific labels within the relevant contextual information, which is selected from a pool of generic sentences based on the lexical[3] and semantic[4] overlap with the labels to be translated. The goal is to identify sentences that are domain-specific in respect of the target domain and contain as much as possible relevant words that can allow the SMT system to learn the translations of the labels. For instance, with the sentence selection we aim to retain only English sentences where the word *injection* belongs to the medical domain, but not to the technical domain. This selection process demonstrated translation improvement of labels, since we try to translate them within disambiguated sentences that belong to the targeted domain. For this work we used our ontology translation system, called OTTO[5], which has been evaluated by translating the labels of several ontologies, i.e. *Organic.Lingua*, DOAP, GeoSkills, STW, TheSoz, into different languages [2].

Nonetheless, some of the domain-specific labels within knowledge graphs may not be automatically translatable with SMT, due to the fact that the bilingual information is missing and cannot be learned from the parallel sentences.

### 3.3 Linking and Translating through Multilingual Linked Data

Since the task of translating labels in knowledge graphs often needs to deal with domain-specific expressions, we require lexical knowledge of the domain. Although SMT systems nowadays are suitable to translate very frequent expressions with an acceptable quality, they fail in translating infrequent domain-specific expressions. This mostly depends on the lack of domain-specific parallel data from which the SMT systems can learn. A valuable solution to handle domain-specific terms is represented by multilingual linked data resources, e.g. DBpedia or IATE, which are continuously updated and can be easily queried. For this reason, the identification, disambiguation of domain-specific expressions is a crucial step towards increasing the translation quality in highly specific domains. With the usage of disambiguated DBpedia entries showed significant translation improvements of domains-specific expressions and ontology labels [3, 4], which confirms the applicability of such techniques in a real scenario. Therefore, we combine the linking approach described in Section 3.1 and automatic label translation with SMT in Section 3.2, whereby we give preference to the translations provided by the linking approach and automatically translate the ontology labels with OTTO, which were not aligned with the DBpedia entries.

[3] string, substring and character overlap

[4] measuring cosine similarity of word embeddings

[5] http://server1.nlp.insight-centre.org/otto/

## 4 Experimental Setting

In this section, we give an overview on the dataset and the translation toolkit used in our experiment. Furthermore, we describe the SMT evaluation techniques, considering the translation direction from English to German.

### 4.1 Statistical Machine Translation and Training Dataset

To align labels within knowledge graphs across different languages, we use machine translation techniques, where we wish to find the best translation, of a source string, given by a log-linear model combining a set of features. The translation that maximizes the score of the log-linear model is obtained by searching all possible translations candidates. The decoder or search procedure, respectively, provides the most probable translation based on statistical translation model learned from the training data.

For a broader domain coverage of datasets necessary to train an SMT system, we merged several parallel corpora, e.g. JRC-Acquis, Europarl, DGT (translation memories generated by the Directorate-General for Translation), KDE, GNOME and TED talks among others, into one parallel dataset (Table 1). For the translation approach, the Moses toolkit [19] and GIZA++ [28] for word alignment generation were used, whereby a language model was build with KenLM [16].

**DBpedia as Multilingual Alignment Resource** The DBpedia project [18] aims to extract structured content from the knowledge added to the Wikipedia repository. DBpedia allows users to semantically query relationships and properties of Wikipedia resources, including links to other related datasets.

| | English | German |
|---|---|---|
| # of tokens | 145,757,562 | 135,627,679 |
| # of types | 561,667 | 1,156,379 |
| # of parallel sent. | 10,065,235 | |

Table 1: Statistics on the merged English-German parallel corpus.

Since some of the domain-specific ontology labels may not be automatically translatable with SMT, due to the fact that the bilingual information is missing and cannot be learned from the parallel sentences. We therefore use the information contained in the DBpedia knowledge base[6] to improve the translation of expressions which may not be known to the SMT system.

**Evaluation dataset** For the evaluation campaign of the proposed experiment we used the Electronic Health Records (EHR) ontology[7], which is based on the work of openEHR[8]. The ontology, which holds 75 labels, tries to overcome the issue of patient information access, which is distributed among several independent and heterogeneous systems that may be syntactically or semantically incompatible. Therefore the EHR ontology focuses on the developed within the framework of a federated information system.

**Machine Translation Evaluation** For the automatic evaluation we used the BLEU [29], METEOR [9] and chrF [33] metrics.

**BLEU** (Bilingual Evaluation Understudy) is calculated for individual translated segments (n-grams) by comparing them with a data set of reference translations. Considering the shortness of the entries in the targeted ontology, we report scores based on the unigram overlap (BLEU-1). Those scores, between 0 and 100 (perfect translation), are then averaged over the whole *evaluation data set* to reach an estimate of the translation's overall quality. **METEOR** (Metric for Evaluation of Translation with Explicit ORdering) is based on the harmonic mean of precision and recall, whereby recall is weighted higher than precision. Along with exact word (or phrase) matching it has additional features, i.e. stemming, paraphrasing and synonymy matching. In contrast to BLEU, the metric produces good correlation with human judgement at the sentence or segment level. **chrF3** is a character n-gram metric, which has shown very good correlations with human judgements.

[6] http://wiki.dbpedia.org/downloads-2016-10

[7] http://trajano.us.es/~isabel/EHR/

[8] http://www.openehr.org/

| | BLEU-1 | METEOR | Chrf |
|---|---|---|---|
| NAISC linking | 25.0 | 17.03 | 27.81 |
| OTTO | 53.6 | 29.53 | 52.18 |
| NAISC-then-OTTO | 55.5 | 30.63 | 53.92 |

Table 2: Translation evaluation based ion the EHR ontology labels.

## 5 Evaluation

The automatic evaluation on the translated EHR ontology labels provided by Dbpedia and the OTTO translation system is based on the correspondence between the automatically generated output and reference translations (gold standard). For the automatic translation evaluation of the EHR labels we used the BLEU, METEOR and chrF metrics.

Table 2 presents the evaluation scores of the translations of the EHR ontology. We observed, that the NAISC linking approach using DBpedia on the whole EHR dataset performs worse compared to the OTTO translation system. This is due the low number of established alignments between the target ontology, i.e. EHR, and the Dbpedia resource, where only 32 (out of 75) alignments with translations into German were found. The linking approach found label alignments and their translations like, *organiser–Veranstalter*, *observation–Beobachtung* or *entry–Eingabe* , but did not provide any alignments for *clinical context* or *list folder*. In the next step, we engage the OTTO translation system to translate all labels of the EHR ontology. With this approach all evaluation scores significantly increase compared to the DBpedia linking method. This can be deducted to the possibility of translating all labels stored in the EHR ontology. Finally, we combined the provided multilingual knowledge from both approaches, i.e. data linking and automatic translation, respectively. In this experiment, we first give preference to translation of a labels provided by NAISC and if no translation is provided, we use the OTTO system to fill the lexical gap in German. Due to the precise translation provided by NAISC

and the possibility to translate the remaining EHR labels, we further improve the translation quality in terms of the used evaluation metrics.

## 6 Conclusion

In this work we propose an approach to link different knowledge graphs across languages. To enable the usage of knowledge graphs for multilingual NLP applications, these resources, which are mostly represented in English only, have to be linked to other multilingual resource. This can be performed by using existing multilingual knowledge graphs, e.g. DBPedia, or by the usage of machine translation. Our ongoing work focuses on the synergy of the linking approaches and machine translation, specifically, how dataset linking can benefit from machine translation and how disambiguation and translation quality of labels can be improved from existing multilingual knowledge graphs. The proposed approach provides the possibility to translate highly specific vocabulary without particular in-domain parallel data. Furthermore, we observed that more than 25% of the identified lexical knowledge consists of multiword-expressions, e.g. "fatal familial in-somnia". For this reason, we also focus on the alignment of nested knowledge inside those expressions (composing partial translation).

## Acknowledgments

This work was supported by the Science Foundation Ireland under Grant Number SFI/12/RC/2289 (Insight).

## References

[1] M. Arcan, K. Asooja, H. Ziad, and P. Buitelaar. OTTO Ontology Translation System. In *ISWC 2015 Posters & Demonstrations Track*, volume 1486, Bethlehem (PA), USA, 2015.

[2] M. Arcan, M. Dragoni, and P. Buitelaar. Translating ontologies in real-world settings. In *Proceedings of the 15th International Semantic Web Conference (ISWC-2016)*, Kobe, Japan, 2016.

[3] M. Arcan, C. Giuliano, M. Turchi, and P. Buitelaar. Identification of Bilingual Terms from Monolingual Documents for Statistical Machine Translation. In *Proceedings of the 4th International Workshop on Computational Terminology (Computerm)*, Dublin, Ireland, 2014.

[4] M. Arcan, M. Turchi, and P. Buitelaar. Knowledge portability with semantic expansion of ontology labels. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics*, Beijing, China, July 2015.

[5] P. Buitelaar, K.-S. Choi, P. Cimiano, and E. H. Hovy. The Multilingual Semantic Web (Dagstuhl Seminar 12362). *Dagstuhl Reports*, 2(9):15–94, 2013.

[6] M. Carpuat, G. Ngai, P. Fung, and K. Church. Creating a bilingual ontology: a corpus-based approach for aligning WordNet and HowNet. In *Proceedings of the 1st Global WordNet Conference*, 2002.

[7] P. Cimiano, E. Montiel-Ponsoda, P. Buitelaar, M. Espinoza, and A. Gómez-Pérez. A note on ontology localization. *Appl. Ontol.*, 5(2):127–137, 2010.

[8] T. Declerck, A. G. Pérez, O. Vela, Z. Gantner, and D. Manzano. Multilingual lexical semantic resources for ontology translation. In *In Proceedings of the 5th International Conference on Language Resources and Evaluation*, Saarbrücken, Germany, 2006.

[9] M. Denkowski and A. Lavie. Meteor universal: Language specific translation evaluation for any target language. In *Proceedings of the EACL 2014 Workshop on Statistical Machine Translation*, 2014.

[10] S. T. Dumais, T. A. Letsche, M. L. Littman, and T. K. Landauer. Automatic cross-language retrieval using latent semantic indexing. In *AAAI spring symposium on cross-language text and speech retrieval*, volume 15, page 21, 1997.

[11] M. Espinoza, E. Montiel-Ponsoda, and A. Gómez-Pérez. Ontology localization. In *Proceedings of the Fifth International Conference on Knowledge Capture*, NY, USA, 2009.

[12] C. Fellbaum. Wordnet. In *Theory and applications of ontology: computer applications*, pages 231–243. Springer, 2010.

[13] B. Fu, R. Brennan, and D. O'Sullivan. Cross-lingual ontology mapping–an investigation of the impact of machine translation. In *The Semantic Web*, pages 1–15. Springer, 2009.

[14] B. Fu, R. Brennan, and D. O'Sullivan. Cross-lingual ontology mapping and its use on the multilingual semantic web. In *Proceedings of the 1st International Workshop on the Multilingual Semantic Web*, pages 13–20, 2010.

[15] J. Gracia, E. Montiel-Ponsoda, P. Cimiano, A. Gómez-Pérez, P. Buitelaar, and J. McCrae. Challenges for the multilingual web of data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 11, 2012.

[16] K. Heafield. KenLM: faster and smaller language model queries. In *Proceedings of the EMNLP 2011 Sixth Workshop on Statistical Machine Translation*, pages 187–197, Edinburgh, Scotland, United Kingdom, July 2011.

[17] J. Jagarlamudi, R. Udupa, H. Daumé III, and A. Bhole. Improving bilingual projections via sparse covariance matrices. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 930–940. Association for Computational Linguistics, 2011.

[18] Jens Lehmann and Robert Isele and Max Jakob and Anja Jentzsch and Dimitris Kontokostas and Pablo N. Mendes and Sebastian Hellmann and Mohamed Morsey and Patrick van Kleef and Sören Auer and Christian Bizer. DBpedia a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web Journal*, 6(2):167–195, 2015.

[19] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst. Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, Stroudsburg, PA, USA, 2007.

[20] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, (2):83–97, 1955.

[21] T. Lesnikova, J. David, and J. Euzenat. Interlinking English and Chinese RDF data using BabelNet. In *Proceedings of the 2015 ACM Symposium on Document Engineering*, DocEng '15, pages 39–42, New York, NY, USA, 2015. ACM.

[22] J. McCrae, P. Cimiano, and R. Klinger. Orthonormal explicit topic analysis for cross-lingual document matching. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1732–1740, 2013.

[23] J. McCrae, E. Montiel-Ponsoda, G. Aguado de Cea, M. J. Espinoza Mejía, and P. Cimiano. Combining statistical and semantic approaches to the translation of ontologies and taxonomies. In *Proceedings of 7th Workshop on Syntax, Structure and Semantics in Statistical Translation*, pages 116–125, 2011.

[24] J. P. McCrae, M. Arcan, K. Asooja, J. Gracia, P. Buitelaar, and P. Cimiano. Domain adaptation for ontology localization. *Web Semantics*, 36:23–31, 2016.

[25] J. P. McCrae and P. Cimiano. Mining translations from the web of open linked data. In *Proceedings of the Joint Workshop on NLP&LOD and SWAIE: Semantic Web, Linked Open Data and Infromation Extraction*, pages 9–13, 2013.

[26] D. Mimno, H. M. Wallach, J. Naradowsky, D. A. Smith, and A. McCallum. Polylingual topic models. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2*, pages 880–889. Association for Computational Linguistics, 2009.

[27] R. Navigli. BabelNet goes to the (Multilingual) Semantic Web. In *Proceedings of the 3rd Workshop on the Multilingual Semantic Web*, 2012.

[28] F. J. Och and H. Ney. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29, 2003.

[29] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 311–318, 2002.

[30] J. Pennington, R. Socher, and C. D. Manning. Glove: Global vectors for word representation. In *EMNLP*, volume 14, pages 1532–1543, 2014.

[31] J. Platt. Fast training of support vector machines using sequential minimal optimization. In B. Schoelkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*. MIT Press, 1998.

[32] J. C. Platt, K. Toutanova, and W.-T. Yih. Translingual document representations from discriminative projections. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 251–261. Association for Computational Linguistics, 2010.

[33] M. Popović. chrF: character n-gram F-score for automatic MT evaluation. In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 392–395, Lisbon, Portugal, September 2015. Association for Computational Linguistics.

[34] D. Spohr, L. Hollink, and P. Cimiano. A machine learning approach to multilingual and cross-lingual ontology matching. *The Semantic Web–ISWC 2011*, pages 665–680, 2011.

[35] K. S. Tai, R. Socher, and C. D. Manning. Improved semantic representations from tree-structured long short-term memory networks. *arXiv preprint arXiv:1503.00075*, 2015.

[36] A. Vinokourov, N. Cristianini, and J. S. Shawe-taylor. Inferring a semantic representation of text via cross-language correlation analysis. In *Advances in neural information processing systems*, pages 1473–1480, 2002.