Ten Years of BabelNet: A Survey

Roberto Navigli[1], **Michele Bevilacqua**[1], **Simone Conia**[1],
Dario Montagnini[2] and **Francesco Cecconi**[2]
[1]Sapienza NLP Group, Sapienza University of Rome, Italy
[2]Babelscape, Italy
{roberto.navigli, michele.bevilacqua, simone.conia}@uniroma1.it
{montagnini, cecconi}@babelscape.com

Abstract

The intelligent manipulation of symbolic knowledge has been a long-sought goal of AI. However, when it comes to Natural Language Processing (NLP), symbols have to be mapped to words and phrases, which are not only ambiguous but also language-specific: multilinguality is indeed a desirable property for NLP systems, and one which enables the generalization of tasks where multiple languages need to be dealt with, without translating text. In this paper we survey BabelNet, a popular wide-coverage lexical-semantic knowledge resource obtained by merging heterogeneous sources into a unified semantic network that helps to scale tasks and applications to hundreds of languages. Over its ten years of existence, thanks to its promise to interconnect languages and resources in structured form, BabelNet has been employed in countless ways and directions. We first introduce the BabelNet model, its components and statistics, and then overview its successful use in a wide range of tasks in NLP as well as in other fields of AI.

1 Introduction

Natural Language Processing (NLP), a key field of AI, deals with text and aims at enabling the intelligent understanding and generation of written language. A fundamental issue when processing language is scaling multilingually, that is, being able to carry out the same task in multiple languages while ideally avoiding the need to repeat the same annotation activity (e.g., providing training data) in each new language. It is certainly true that the recent revolution in NLP – often known as its "ImageNet moment" – sparked by the availability of large language models, such as ELMo [Peters *et al.*, 2018] and BERT [Devlin *et al.*, 2019], pretrained on massive amounts of textual data, has enabled high performance in a large variety of tasks. Thanks to transfer learning, these models have also made it possible to tackle several problems across different languages in which task-specific data is not available.

However, pretraining is not enough. Despite the tremendous increase in the number, size and variety of pretrained language models, there is a growing consensus on the need to integrate symbolic knowledge into neural architectures [d'Avila Garcez and Lamb, 2020]. The rationale is that the use of, and linkage to, symbolic knowledge can not only enable interpretable, explainable and accountable AI systems, but it can also increase the degree of generalization to rare patterns (e.g., infrequent meanings) and promote better use of information which is not explicit in the text.

Symbolic knowledge requires that the link between form and meaning be made explicit, connecting strings to representations of concepts, entities and thoughts. Historical resources such as WordNet [Miller, 1995] are important endeavors which systematize symbolic knowledge about the words of a language, i.e., lexicographic knowledge, not only in a machine-readable format, but also in structured form, thanks to the organization of concepts into a semantic network. More recent efforts take advantage of crowdsourcing and have led to the creation of incredibly valuable resources, such as Wikipedia and Wiktionary. Wikipedia, for instance, is at the heart of much ongoing research thanks to its continuously growing wealth of encyclopedic knowledge in hundreds of languages [Hovy *et al.*, 2013]. However, a key issue concerning many sources of knowledge, including lexico-semantic resources, is their lack of inter-resource links. For instance, the Spanish Wikipedia, the Chinese Wiktionary and the English WordNet provide complementary knowledge while being mostly disconnected from each other. To tackle this problem ten years ago, Navigli and Ponzetto [2010; 2012] put forward BabelNet, a large-scale multilingual resource which integrates knowledge coming from heterogeneous sources and languages, such as the aforementioned Wikipedia, Wiktionary and wordnets, into a unified multilingual semantic network.[1]

Prior to the initial wave of pretrained language models, BabelNet enabled the development of semantics-rich approaches in situations where multilingual data was available in small quantities only, or not at all, scenarios in which approaches based on machine translation were not viable options [Moro *et al.*, 2013; Chakraborty *et al.*, 2016; Klein *et al.*, 2017]. Although the advent of widely available pretrained models has disrupted the landscape of NLP and AI, BabelNet is today receiving constant, even increased, attention from multiple fields crossing the frontiers of NLP,

[1]Available via API and for download at https://babelnet.org.

including Computer Vision and electronic lexicography.

In this survey we provide an overview of BabelNet, its model, contents and statistics, as well as where it has found success and where it is currently being used, why it is still relevant and why it is still growing, with an eye to its future developments and possible applications.

2 BabelNet

BabelNet is a multilingual semantic network that brings together heterogeneous resources such as WordNet, Wikipedia, Wikidata, Wiktionary, and many others. Similarly to a jigsaw puzzle, the goal is to integrate different bits of information so as to provide as complete a picture as possible of the lexical and semantic knowledge available from the integrated resources. In this Section, we review the main components of BabelNet and highlight how it has changed over its ten years of life, providing an overview of its semantic model (§2.1), its sources of knowledge (§2.2), their integration into BabelNet and a summary of its evolution in numbers (§2.3).

2.1 The BabelNet Model

Multilingual synsets. BabelNet represents each meaning based on the WordNet notion of a synset, which is the set of synonymous words (i.e., senses) that can be used to express the same meaning in a given language. For example, in WordNet the concept of DOG is represented by the set of words { *dog*, *domestic dog*, *Canis familiaris* } whereas the named entity NEW YORK is defined as the set { *New York*, *New York City*, *Greater New York* }. BabelNet extends this notion to include synonymous lexicalizations in multiple languages. For instance, the multilingual synset of DOG would also contain the terms { $chien_{\mathrm{FR}}$, $cane_{\mathrm{IT}}$, $Hund_{\mathrm{DE}}$, 犬$_{\mathrm{ZH}}$, …, 개$_{\mathrm{KO}}$ }. The adoption of the synset model offers two main advantages. First, this abstraction provides a unified interface between the concept or named entity and its associated lexical-semantic knowledge, independently of the language and the resource in which such knowledge is expressed. Second, the multilingual extension of synsets enables language generalization by just linking language-specific content to multilingual synsets.

The graph model. Analogously to WordNet, BabelNet can be viewed as a graph where synsets are nodes and edges are semantic relations between them. The relations in BabelNet stem from the underlying resources which provide them. For instance, these range from WordNet-based semantic relations such as *hypernymy* (generalization or is-a), *meronymy* (part-whole) and *antonymy* (opposite-of), to instance-focused relations such as *author*, *location* and *occupation*.

2.2 The BabelNet Sources

The BabelNet graph provides a single access point to navigate its underlying heterogeneous resources. While the original release of BabelNet was centered around the integration of WordNet and Wikipedia, BabelNet 5.0 currently draws knowledge from 51 sources. In the following we describe the main resources that contribute to today's BabelNet.

The Princeton WordNet and Other Wordnets

The original Princeton WordNet [Miller, 1995], or simply WordNet, has been a part of BabelNet since its conception and is still one of its cornerstones. WordNet is the first large-scale lexicographic inventory of English primarily meant to be machine-readable. Even today it is one of the largest resources of its kind, featuring around 82K nominal synsets, 13.8K verbal synsets, 18K adjectival synsets and 3.6K adverbial synsets, totaling around 117K synsets, each of which comes with an associated definition. Each synset is connected to other synsets through semantic relations such as hypernymy and hyponymy. BabelNet includes WordNet 3.0 and inherits all its synsets, definitions and relations.

While the Princeton WordNet is no longer actively updated, McCrae *et al.* [2020] created the English WordNet, an open-source, up-to-date, refined version of the original project, which introduces new synsets, removes duplicates and updates definitions. BabelNet 5.0 includes both the Princeton WordNet 3.0 and the 2020 version of the English WordNet, thanks to synsets in the latter resource being linked to the former. The huge impact of the Princeton WordNet in NLP inspired other researchers to create analogous resources in many other languages, most of which having one-to-one links to the English synsets.[2] BabelNet integrates different wordnets in 33 different languages, including many from the Open Multilingual WordNet [Bond and Foster, 2013].

Wikipedia

Wikipedia is arguably the largest freely-available encyclopedia. For this reason it is an obvious choice as a source of knowledge for BabelNet, which has, therefore, included this encyclopedia since its first release. Although at first glance Wikipedia and WordNet-like resources may not seem very much alike, BabelNet's seminal intuition was to treat each Wikipage as a synset whose multilingual lexicalizations are provided by the page titles connected to each other via interlanguage links (e.g., $Fork_{\mathrm{EN}}$, $Forchetta_{\mathrm{IT}}$, $Fourchette_{\mathrm{FR}}$, etc.), but excluding domain/topic labels specified within parentheses, e.g., $Spring_{\mathrm{EN}}$ instead of *Spring (season)*$_{\mathrm{EN}}$. Redirections to such pages are also included as lexicalizations (e.g., *Voiture automobile*$_{\mathrm{FR}}$ redirects to $Automobile_{\mathrm{FR}}$). Moreover, BabelNet retains the Wikipedia relational structure formed by the hyperlinks that connect a Wikipedia page to other semantically-relevant pages. Pictures, first-sentence definitions and categories are also collected.

One concern of relying on an ever growing, continuously updated resource such as Wikipedia is that BabelNet's knowledge can rapidly become outdated. However, during its existence, BabelNet has been steadily updated to include newer information from Wikipedia – BabelNet 5.0 is built on top of the Wikipedia dump of November 2020 – with up-to-date knowledge and an increased number of supported languages.

Wikidata

Designed to provide structured data for several Wikimedia projects, Wikidata has quickly become the largest online open knowledge base. In the Wikidata graph, each node represents a concept or a named entity and is identified by a label, i.e., its most common lexicalization, and a set of aliases, i.e., other lexicalizations in multiple languages. Thanks to the Wikipedia interlanguage links, each Wikidata node is mapped

[2] http://globalwordnet.org/resources/wordnets-in-the-world/

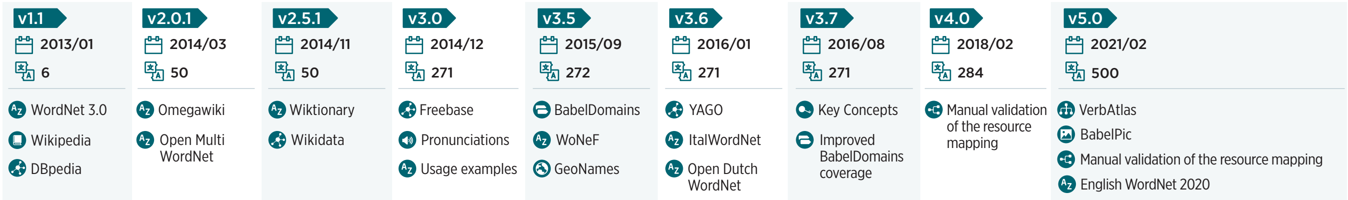

Figure 1: Timeline of the BabelNet development and evolution (release version, date, number of languages, changes), from version 1.1 to the current 5.0 version. Full statistics available at https://babelnet.org/statistics.

to the corresponding BabelNet multilingual synset. BabelNet also inherits the property-value statements that connect Wikidata nodes. For instance, the Wikidata nodes for the *Sherlock Holmes* novels feature an *author* property whose value is the Wikidata node corresponding to *Arthur Conan Doyle*.

The Wikipedia Bitaxonomy

Because being able to generalize concepts (e.g., *utility bond* → *municipal bond* → *bond* → *debt instrument* etc.) is useful in a wide variety of tasks, BabelNet integrates a large taxonomy of concepts and entities, namely the Wikipedia Bitaxonomy [Flati *et al.*, 2016], which complements the hypernymy information available in WordNet and Wikidata with is-a relations between Wikipedia pages and categories.

Other Lexicographic Resources

BabelNet also draws its multilingual knowledge from other lexicographic resources such as OmegaWiki and the English-language Wiktionary, added to BabelNet as part of its 2.0 and 2.5 releases, respectively. OmegaWiki is a lexical knowledge base centered around a WordNet-like multilingual synset model. While OmegaWiki can optionally encode semantic relations for meaning entries, Wiktionary is even less structured: each Wiktionary page enumerates the possible uses of an English word and its possible translations in other languages, without organizing different senses into a graph.

2.3 The Making of BabelNet

From its very conception, the core of BabelNet has always been its algorithm to automatically link Wikipedia articles and WordNet senses. During its life, however, BabelNet has undergone major changes in several other key areas. In the following, we review the mapping algorithms and describe the manual efforts aimed at improving BabelNet's quality.

Mapping Wikipedia to WordNet

While Wikipedia and WordNet are complementary, due to their respective encyclopedic and lexicographic nature, they also share a considerable number of entries. BabelNet leverages this overlap and integrates the two graphs by mapping and bringing together around 47K WordNet synsets and Wikipedia pages that describe the same concept or entity.

A mapping $\phi_{\text{WIKI}\to\text{WN}}$ is produced by maximizing, for a given Wikipedia page w, the conditional probability $P(s|w)$ of selecting the most suitable WordNet sense s (or nothing if not mappable, i.e., ϵ):

$$\phi_{\text{WIKI}\to\text{WN}}(w) = \operatorname*{argmax}_{s \in \text{SENSES}_{\text{WN}}(w) \cup \{\epsilon\}} P(s|w)$$

The above probability is estimated by defining i) the context of s as its neighbors in the WordNet graph and the word senses (i.e., WordNet nodes) that appear in its WordNet definition, and ii) the context of a Wikipedia page as the WordNet senses that appear in the page titles of links and redirections to w, the domain/topic labels in the title of w and its Wikipedia categories; after this, a Personalized PageRank-based algorithm is applied to the WordNet graph to determine the probability of reaching the senses in the context of s (re)starting from senses in the context of w.

Mapping Wiktionary to WordNet

The mapping of Wiktionary in BabelNet 5.0 offers an example of how the methodologies used to update BabelNet have evolved together with NLP: Wiktionary entries are now integrated automatically using a BERT-based neural model, fine-tuned to associate a word sense definition to its correct synset, attaining an F_1 score of 92% on a manually annotated test set.

Manually Improving BabelNet

One of the major, yet often understated, improvements in BabelNet over its first, completely automatic releases, is the huge manual effort that has been put into increasing the quality of the semantic network. For example, in BabelNet 5.0 more than 90% of the mapping between Wikipedia pages and WordNet synsets has been manually validated by experts, resulting in an overall mapping precision above 99.5%. Not only this, the set of domains in BabelNet 5.0 has also been overhauled. These were introduced as part of the BabelNet 3.5 release and were initially derived from the Wikipedia Categories [Camacho-Collados and Navigli, 2017]. Following recent studies [Lacerra *et al.*, 2020], the new set of domains is now more precise, provides wider coverage and, most importantly, is used to manually label all the nominal synsets of WordNet.

BabelNet in Numbers

Figure 1 and Table 1 show the steady evolution of BabelNet from its first public release to its latest iteration. Over its ten years of development, there has been a massive growth in the number of supported languages (from 6 to 500), sources (from 4 to 51), senses (from 22M to 1429M) and synsets (from 6M to 20M). Importantly, BabelNet 5.0 synset IDs are backward compatible with its previous releases, meaning that concepts preserve their identity over time as much as possible, and a system or resource that has been developed with a previous version of BabelNet can easily be updated.

Release	Date	Lang.	Sources	Senses	Synsets	Definitions	Images
1.1	13/01	6	4	22M	5.6M	8.4M	6.5M
2.0	14/03	50	5	50M	9.3M	18.0M	7.8M
2.5	14/11	50	7	68M	9.3M	21.8M	7.8M
3.0	14/12	271	7	117M	13.8M	40.3M	11.0M
3.5	15/09	272	13	119M	13.8M	40.6M	10.8M
3.6	16/01	271	13	746M	13.8M	40.7M	10.8M
3.7	16/08	271	14	746M	13.8M	40.7M	10.8M
4.0	18/02	284	47	809M	15.8M	91.2M	54.2M
5.0	21/02	500	51	1,429M	20.3M	135.3M	51.3M

Table 1: Evolution of BabelNet in numbers, from its first (1.1) to its latest public release (5.0).

3 Applications

Although BabelNet in and of itself is already a ready-to-use, innovative multilingual dictionary (e.g., for language learners), the research community has used it not only in computational lexical semantics, but also in an array of very different tasks. Here we review a selection of these: Word Sense Disambiguation (WSD) (§3.1), sense, concept and named entity representation (§3.2), diagnostic tasks (§3.3), resources (§3.4), and miscellanea (§3.5).

3.1 Empowering Word Sense Disambiguation

WSD is a task with a very long history in NLP [Navigli, 2009], which consists in associating a word or expression in context with the most fitting meaning among those listed in the sense inventory. BabelNet has so far been key to enabling WSD to scale beyond English.

Multilingual WSD Data

Sense-annotated data are the holy grail of supervised WSD. In English, where WordNet is the de facto standard, large datasets such as SemCor and the WordNet Tagged Gloss Corpus are commonly used. In other languages there exist many annotated datasets [Petrolito and Bond, 2014] whose usage, however, is often hindered by the lack of standardization in both format and underlying sense inventory. This is particularly harmful for cross-lingual models, the best in terms of scalability, which benefit from a shared multilingual output space. In this landscape, BabelNet, by linking together resources in many different languages within a stable sense inventory, offers a practical solution that has, indeed, enabled continued multilingual evaluations [Raganato *et al.*, 2017; Scozzafava *et al.*, 2020; Bevilacqua and Navigli, 2020; Scarlini *et al.*, 2020a; Scarlini *et al.*, 2020b; Luan *et al.*, 2020] on datasets such as SemEval-2013 Task 12 [Navigli *et al.*, 2013] and SemEval-2015 Task 13 [Moro and Navigli, 2015] which cover multiple languages.

Due to the paucity of large gold training datasets for WSD in languages other than English, BabelNet has also been usefully exploited in the construction of many silver-quality datasets. The Train-o-Matic [Pasini and Navigli, 2017] corpus was built for six different languages using heuristics based on Personalized PageRank to extract sentences from Wikipedia and the UN Parallel Corpus. Similarly, OneSeC [Scarlini *et al.*, 2019] exploited the "One Sense per Wikipedia Category" heuristic to collect noun annotations from Wikipedia, making training data available for French, German, Italian, and Spanish. In MuLaN [Barba *et al.*, 2020], multilingual contextualized embeddings and BabelNet are exploited together to project sense annotations from English to other languages. Finally, XL-WSD [Pasini *et al.*, 2021] uses BabelNet as the underlying inventory to propose a unified framework for multilingual WSD in 18 different languages, including many non Indoeuropean ones like Basque, Chinese, Hungarian, Japanese and Korean.

Multilingual WSD Systems

Beyond being used as a multilingual sense inventory, BabelNet has often been employed to build automatic multilingual WSD systems [Bevilacqua *et al.*, 2021]. First, BabelNet is the backbone of many so-called *knowledge-based* WSD systems, i.e., those that do not rely on corpus supervision but, instead, exploit other forms of lexical-semantic information. For example, BabelNet, thanks to its seamless integration of both concepts and named entities, enables Babelfy [Moro *et al.*, 2014] to tackle two similar yet different tasks, i.e., WSD and Entity Linking, jointly in hundreds of languages. SyntagRank [Scozzafava *et al.*, 2020] exploits the BabelNet graph along with additional edges from SyntagNet [Maru *et al.*, 2019] to perform multilingual WSD with a method based on Personalized PageRank.

While knowledge-based systems are certainly good in terms of scalability – indeed, there are still few competitors to Babelfy in its capability to jointly perform language-agnostic WSD and Entity Linking – they are not on a par with so-called *supervised* systems, which rely on training corpora to learn a model that maps target words in context to senses from the inventory. In addition to the use of silver data built from BabelNet, knowledge from the resource has been incorporated into supervised systems. For example, BabelPic images (§3.4) have been used to create multimodal synset embeddings that are more performant for WSD than unimodal counterparts [Calabrese *et al.*, 2020a]. Recent work [Luan *et al.*, 2020] has also shown that the joint use of machine translation and BabelNet can indeed improve the performance of both knowledge-based and supervised systems. This is achieved by interpolating between the probability distribution produced by the disambiguation system, a learned unconditional probability distribution [Pasini *et al.*, 2020] (also produced exploiting BabelNet), and a soft probability distribution computed by matching the target expression with its automatic translations.

The recent trend of incorporating knowledge-base information in supervised WSD (which has enabled performances exceeding the 80% ceiling set by the inter-annotator agreement [Bevilacqua and Navigli, 2020; Conia and Navigli, 2021]) makes it likely that BabelNet will see even more usage in the field in the future, especially in the multilingual setting.

3.2 Semantic Representation Learning

Representing meaning at the word level poses a problem, since traditional word embedding techniques [Mikolov *et al.*, 2013; Pennington *et al.*, 2014] suffer from the so-called *meaning conflation deficiency problem* [Camacho-Collados and Pilehvar, 2018]: a word may have multiple meanings that are, however, pooled into a single representation in which the most common meaning may overshadow the others. This is why, in parallel with the development of contextualized word embeddings, representation learning techniques for individual word senses, concepts and named entities are still being actively investigated, with BabelNet playing a fundamental role in enabling multilinguality in semantic representations.

Multilingual Sense Embeddings

One of the first notable applications of BabelNet for representation learning was SensEmbed [Iacobacci *et al.*, 2015], a word2vec-based approach that takes advantage of BabelNet to learn dense sense representations, or sense embeddings, from a Wikipedia corpus where each content word is replaced by a BabelNet sense. More recently, SensEmBERT [Scarlini *et al.*, 2020a] and its extension, ARES [Scarlini *et al.*, 2020b], showed that contextual embeddings from pretrained language models can be enriched by using BabelNet to provide targeted context for each sense, attaining state-of-the-art results in multilingual WSD. Furthermore, LessLex [Colla *et al.*, 2020] successfully exploited BabelNet to produce sense embeddings that can be used not only to capture the meaning of single words in context, but also to represent and compare the semantics of larger portions of text across languages.

Interpretable Meaning Representations

While the aforementioned studies were concerned with dense vector representations, BabelNet has also enabled the development of high-quality sparse vectors, which allow a higher level of interpretability and flexibility, *desiderata* that are becoming increasingly necessary in production-ready systems. In particular, NASARI [Camacho-Collados *et al.*, 2016] takes advantage of the BabelNet graph to represent a concept through a vector whose higher-scoring components correspond to the most important terms that appear in the Wikipedia pages describing the concept. Conception [Conia and Navigli, 2020] builds upon NASARI and BabelNet to create state-of-the-art language-agnostic sparse representations that describe a concept through its most related concepts, therefore abstracting away from individual languages.

3.3 Probing Semantic Capabilities of NLP Systems

Ambiguity is a pervasive problem in the processing of natural language. This makes it all the more relevant, given the current dependence of NLP on pretrained contextualized embeddings, to test the capability of such models to deal with this phenomenon. To this end, Pilehvar and Camacho-Collados [2019] put forward the Word-in-Context (WiC) task and datasets, whose construction relies heavily on BabelNet's mapping to other resources. WiC probes whether a system is able to discern when a word conveys the same meaning in two different contexts. This capability is so important that WiC is part of the popular SuperGLUE test suite [Wang *et al.*, 2019a]. XL-WiC [Raganato *et al.*, 2020a] and MCL-WiC [Martelli *et al.*, 2021] – multilingual and cross-lingual expansions of WiC – have also exploited BabelNet as a multilingual dictionary to obtain candidate lemmas to be validated in the annotation phase. BabelNet has also been used as a source of semantic relations in the Linguistic Diagnostics Toolkit (LDT) [Rogers *et al.*, 2018], an open-source tool that quantifies interpretable characteristics of a word representation model.

Interpretability in Machine Translation

Solving ambiguity has also long been a dream of Machine Translation since its conception in the late 1940s – for example, the English word *ball* can correspond to either *ballo* (dance) or *palla* (round object) in Italian, depending on the context. Thanks to its multilingual semantic network, BabelNet has successfully been employed in the creation of two evaluation benchmarks that assess the ability of Neural Machine Translation (NMT) models to implicitly disambiguate and, therefore, correctly translate ambiguous words. In the large-scale MuCoW benchmark [Raganato *et al.*, 2019; Raganato *et al.*, 2020b] for NMT, BabelNet is used to retrieve translations for homographs in 10 language pairs. In a similar fashion, Emelin *et al.* [2020] use BabelNet to retrieve English-German sentences and build an NMT benchmark that focuses on *adversarial attractors*, i.e., cue words that are good predictors for some specific translation, but are actually misleading in some other cases, e.g., *hot* for *spring* in *John met his wife in the hot spring of 1988*.

Towards Semantics-Aware Measures

In addition to the construction of challenge sets, explicit semantic knowledge can be used to enable better measures of performance for Natural Language Generation (NLG) systems. Indeed, metrics such as BLEU [Post, 2018] based on string matching are still commonplace in experimental comparisons, despite their overly simplistic assumptions that result, for example, in *spring* and *leap* being considered as completely different words in all possible contexts. One example of a measure that exploits a knowledge graph is METEOR [Banerjee and Lavie, 2005], which takes synonyms from WordNet into account. While METEOR tends to "overgeneralize", as there is no guarantee that synonyms are contextually appropriate, METEOR-WSD [Apidianaki and Marie, 2015] tries to address this issue by explicitly integrating WSD, producing context-aware synonym substitutions using BabelNet. It would be interesting to see whether modern neural WSD models could be integrated into this evaluation framework.

3.4 Resources

Linguistic Resources

The multilingual mapping between different resources makes BabelNet an appealing candidate to work as hub on which to ground new, or already existing, linguistic resources: BabelNet is at the core of a dictionary matrix under development within the EU-funded ELEXIS project (https://elex.is/), aimed at interlinking lexicographic resources such as professional or proprietary dictionaries in tens of languages and showing their usefulness in NLP tasks.

The multilingual synset model of BabelNet allows researchers to expand their resources and support an increasing number of languages. For example, HurtLex [Bassignana *et al.*, 2018], a very popular catalogue of expressions of hate speech, was built by exploiting BabelNet to expand a preexisting Italian resource multilingually. BabelNet has also been exploited to overcome the language specificity of existing predicate-argument structure inventories used extensively in Semantic Role Labeling (SRL) and/or Semantic Parsing such as FrameNet [Baker *et al.*, 1998], VerbNet [Schuler, 2005] and PropBank [Kingsbury and Palmer, 2002]. While these proved to be successful for the language they were devised for, their application to multilingual scenarios was not straightforward as they were often language-specific. To address this issue, Di Fabio *et al.* [2019] proposed VerbAtlas, which associates semantically explicit roles (e.g., *Agent*, *Location*) with coarse-grained predicate frames, obtained by manually clustering verbal synsets which are therefore independent of any specific lexical realization. Thanks to BabelNet, the VerbAtlas synset-based predicate-argument structures can potentially be employed across hundreds of languages automatically using cross-lingual transfer techniques [Conia *et al.*, 2021], overcoming the years-long process of creating inventories separately for each individual language and advancing multilingual Natural Language Understanding [Navigli, 2018]. Finally, another resource that takes advantage of the BabelNet synset model is SyntagNet [Maru *et al.*, 2019], which manually connects concepts that often appear together through over 78,000 syntagmatic relation edges, boosting performance of multilingual WSD systems [Scozzafava *et al.*, 2020; Scarlini *et al.*, 2020b].

Multimodal Resources

It is worth noting that BabelNet has been functioning as a knowledge hub for areas that go beyond language and NLP. For example, BabelNet is linked to ImageNet, a database containing over 15 million images. While ImageNet mostly covers concrete objects, e.g., *ladder* and *apple*, BabelNet 5.0 also includes a reliable mapping to images for abstract concepts, e.g., *empathy* and *birthday*, thanks to BabelPic [Calabrese *et al.*, 2020b]. Another step in the representation of non-concrete meanings is the linking of BabelNet to IMAGACT [Gregori *et al.*, 2016], which illustrates the so-called *verbs of action* in multiple languages by way not only of static images but also of short videos. Finally, BabelNet has also been linked to EmojiNet [Wijeratne *et al.*, 2016], the largest machine-readable sense inventory of emojis which are a fundamental piece of (modern) human communication that lies at the intersection of language, vision and pragmatics. The use of emojis is, just like regular words, highly polysemous: for example, the 😂 emoji can be used to signal that one is happy or as a laughing response to a joke. Not only that, emojis often, but not always, convey the same meaning across languages, making their linkage to BabelNet's multilingual synsets a natural choice.

And this is not all, as BabelNet is currently enabling the creation and expansion of several other multilingual and multimodal resources, from VisualSem [Alberts *et al.*, 2020], a high-quality knowledge graph for vision and language, to MultiSubs [Wang *et al.*, 2021], a large multilingual corpus of subtitles in which words are grounded to images.

3.5 Other Applications

In addition to all we have covered so far, BabelNet has also proved vitally useful for a range of other purposes. In what follows we present just a small collection of the many projects that have taken advantage of BabelNet's potential.

Enriching Knowledge Bases. BabelNet is invaluable for approaches to the enrichment of general-purpose knowledge bases. For example, it has been shown that the performance of a baseline link prediction model can be boosted by using BabelNet to perform data augmentation by providing translations of knowledge-graph triples [Klein *et al.*, 2017]. Also, WSD using BabelNet as a sense inventory can be exploited to filter out noisy relation extraction rules [Moro *et al.*, 2013], e.g., *PERSON met PERSON* for the relation *married*. Word Sense Induction (WSI) methodologies have also benefited from BabelNet to boost their interpretability: Panchenko [2016] was able to map approximately 40% of their automatically inducted senses to a BabelNet synset, with around 87% precision.

Thanks to its large coverage, BabelNet has also been used to model not only general-purpose, but also domain-specific knowledge. One recent approach in this direction is SciKGraph [Tosi and dos Reis, 2021], a framework to structure and analyze a scientific field as a knowledge graph by linking concepts and named entities that co-occur in a text using the semantic relations defined in BabelNet.

Malicious language. Another domain in which BabelNet has been utilized successfully is the prevention of malicious uses of language. For instance, it has been exploited to build input features in a model that discovers vulnerabilities in Internet-of-Things (IoT) deployments [Wang *et al.*, 2019b], aiming to prevent unauthorized access resulting from conflict in user-generated rules. The clickbait detection system of Chakraborty *et al.* [2016] has used BabelNet to build clusters of synsets from keywords, which are then used to decide whether a page is clickbait. Finally, the cross-lingual plagiarism detection system of Franco-Salvador *et al.* [2016] represents documents as graphs of BabelNet synsets, abstracting away from any specific language realization.

Computational Social Science. There have also been applications in computational social science: for example, a large-scale study of homophily [Faralli *et al.*, 2015] – the tendency to befriend individuals sharing the same interests – in social networks exploited BabelNet and Babelfy to perform entity linking of famous Twitter users, e.g., `@britneyspears`, producing Twixonomy, a DAG containing disambiguated users. As a result, a similarity score between arbitrary users can be computed in terms of their shared interests.

4 The Road Ahead

The work surveyed in this paper, though selective, will hopefully give a sense of how much ground has been covered over the ten years of BabelNet. However, the road ahead stretches far longer than the ten years already passed, with many promising directions that are leading us towards a tool which is ever more useful for both *humans* and *machines*. Indeed, the human consumer is going to benefit from a richer, more accessible resource that brings together information that would otherwise be scattered across heterogeneous resources, while the algorithms of the future are going to learn to use a vast repository of machine-readable information in addition to that which they store (implicitly) in their hundreds of billions of parameters. Let us now outline some steps to further these objectives.

Explicit language-agnostic representations. BabelNet enables multilinguality at the word level by providing a lexical-semantic interface for concepts, which can be seen as the building blocks for encoding facts and events. However, what is currently missing in NLP is a unified, language-agnostic, semantics-rich formalism to encode entire sentences or documents: for example, AMR [Banarescu *et al.*, 2013] still relies on language-specific resources such as the English PropBank and OntoNotes, while UCCA [Abend and Rappoport, 2013] only provides a semantic overlay to accompany language-specific text. A way forward is to extend these formalisms multilingually by replacing words in semantic parses with BabelNet concepts, therefore representing text with the same graph-like structure across languages.

From semantics to commonsense knowledge. As we saw in §2, BabelNet contains large amounts of structured encyclopedic knowledge through edges coming from its underlying resources. These relations can easily be used to verify, for example, that *platypuses* are *mammals*, and that *Wales* is part of the *UK*. However, what is currently missing is the commonsense information that is rarely registered explicitly in knowledge bases, such as that *a soccer ball has to be inflated*. The resources that do provide commonsense knowledge, such as ConceptNet [Speer *et al.*, 2017], are rich, but also neither explicitly semantic nor language-agnostic, since they contain triples of language-specific strings, not concepts. An important next step will be to automatically collect and integrate pieces of commonsense knowledge, providing a unified repository of semantics and world knowledge for the retrieval-augmented and/or neurosymbolic methods of tomorrow [Guu *et al.*, 2020; Lewis *et al.*, 2020].

A multimodally-grounded resource. Images and videos are fundamental in this vision of a unified repository of knowledge. BabelNet already provides links to visual information, and this mapping is being manually and automatically refined and enriched [Vannella *et al.*, 2014; Gregori *et al.*, 2016; Calabrese *et al.*, 2020b]. In the future, having a resource that unifies access to semantic and visual knowledge could open up new questions, scenarios and tasks. We think that going forward on this path is key for progress, since, as is becoming increasingly clear, NLP needs to break out of its sandbox, connecting to vision and other forms of perception/embodiment, if we want to have any chance of representing "meaning" as humans do, instead of just producing cunning statistical models [Bender and Koller, 2020].

5 Final Remarks

BabelNet has been available for over a decade: we have taken this opportunity to review how this popular resource has changed over the years and how it has contributed to advancing not only NLP but also other fields of AI, in its role as central, multilingual hub for knowledge about concepts and its integration into effective neurosymbolic approaches.

BabelNet aims to modernize content, scope and scale of past and present electronic dictionaries and encyclopedias. Today, it provides a unified view of heterogeneous, diverse resources in hundreds of languages, enabling the development of multilingual semantics-first resources, models and systems. Tomorrow, BabelNet, thanks to its ongoing integration of world knowledge and its increasing grounding to visual data, may constitute a fundamental building block towards explicit language-agnostic semantic representations to complement and enhance pretrained models and other data-driven methodologies.

Acknowledgments

The authors gratefully acknowledge the support of the MOUSSE ERC Consolidator Grant No. 726487 and the ELEXIS project No. 731015 under the European Union's Horizon 2020 research and innovation programme.

This work was supported in part by the MIUR under grant "Dipartimenti di eccellenza 2018-2022" of the Department of Computer Science of Sapienza University.

References

[Abend and Rappoport, 2013] Omri Abend and Ari Rappoport. Universal Conceptual Cognitive Annotation (UCCA). In *Proc. of ACL*, 2013.

[Alberts *et al.*, 2020] Houda Alberts, Teresa Huang, Yash Deshpande, Yibo Liu, Kyunghyun Cho, Clara Vania, and Iacer Calixto. VisualSem: a high-quality knowledge graph for vision and language. *arXiv preprint arXiv:2008.09150*, 2020.

[Apidianaki and Marie, 2015] Marianna Apidianaki and Benjamin Marie. METEOR-WSD: Improved sense matching in MT evaluation. In *Proc. of SSST*, pages 49–51, 2015.

[Baker *et al.*, 1998] Collin F. Baker, Charles J. Fillmore, and John B. Lowe. The Berkeley FrameNet project. In *Proc. of ACL-COLING*, pages 86–90, 1998.

[Banarescu *et al.*, 2013] Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. Abstract Meaning Representation for sembanking. In *Proc. of LAW-ID*, 2013.

[Banerjee and Lavie, 2005] Satanjeev Banerjee and Alon Lavie. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proc. of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for MT*, pages 65–72, 2005.

[Barba *et al.*, 2020] Edoardo Barba, Luigi Procopio, Niccolò Campolungo, Tommaso Pasini, and Roberto Navigli. MuLaN: Multilingual label propagation for word sense disambiguation. In *Proc. of IJCAI*, pages 3837–3844, 2020.

[Bassignana *et al.*, 2018] Elisa Bassignana, Valerio Basile, and Viviana Patti. Hurtlex: A multilingual lexicon of words to hurt. In *Proc. of CLiC-it*, 2018.

[Bender and Koller, 2020] Emily M. Bender and Alexander Koller. Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proc. of ACL*, pages 5185–5198, 2020.

[Bevilacqua and Navigli, 2020] Michele Bevilacqua and Roberto Navigli. Breaking through the 80% glass ceiling: Raising the state of the art in Word Sense Disambiguation by incorporating knowledge graph information. In *Proc. of ACL*, 2020.

[Bevilacqua *et al.*, 2021] Michele Bevilacqua, Tommaso Pasini, Alessandro Raganato, and Roberto Navigli. Recent trends in Word Sense Disambiguation: A survey. In *Proc. of IJCAI*, 2021.

[Bond and Foster, 2013] Francis Bond and Ryan Foster. Linking and extending an open multilingual Wordnet. In *Proc. of ACL*, pages 1352–1362, 2013.

[Calabrese *et al.*, 2020a] Agostina Calabrese, Michele Bevilacqua, and Roberto Navigli. Evilbert: Learning task-agnostic multimodal sense embeddings. In *Proc. of IJCAI*, pages 481–487, 2020.

[Calabrese *et al.*, 2020b] Agostina Calabrese, Michele Bevilacqua, and Roberto Navigli. Fatality killed the cat or: BabelPic, a multimodal dataset for non-concrete concepts. In *Proc. of ACL*, pages 4680–4686, 2020.

[Camacho-Collados and Navigli, 2017] Jose Camacho-Collados and Roberto Navigli. BabelDomains: Large-scale domain labeling of lexical resources. In *Proc. of EACL*, 2017.

[Camacho-Collados and Pilehvar, 2018] Jose Camacho-Collados and Mohammad Taher Pilehvar. From word to sense embeddings: A survey on vector representations of meaning. *J. Artif. Int. Res.*, 63(1):743–788, 2018.

[Camacho-Collados *et al.*, 2016] José Camacho-Collados, Mohammad Taher Pilehvar, and Roberto Navigli. NASARI: Integrating explicit knowledge and corpus statistics for a multilingual representation of concepts and entities. *Artif. Intell.*, 240:36–64, 2016.

[Chakraborty *et al.*, 2016] Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. Stop clickbait: Detecting and preventing clickbaits in online news media. In *Proc. of ASONAM*, 2016.

[Colla *et al.*, 2020] Davide Colla, Enrico Mensa, and Daniele P. Radicioni. LessLex: Linking multilingual embeddings to SenSe representations of LEXical items. *Computational Linguistics*, 46(2):289–333, 2020.

[Conia and Navigli, 2020] Simone Conia and Roberto Navigli. Conception: Multilingually-enhanced, human-readable concept vector representations. In *Proc. of COLING*, 2020.

[Conia and Navigli, 2021] Simone Conia and Roberto Navigli. Framing Word Sense Disambiguation as a multi-label problem for model-agnostic knowledge integration. In *Proc. of EACL*, 2021.

[Conia *et al.*, 2021] Simone Conia, Andrea Bacciu, and Roberto Navigli. Unifying cross-lingual Semantic Role Labeling with heterogeneous linguistic resources. In *Proc. of NAACL*, 2021.

[d'Avila Garcez and Lamb, 2020] Artur d'Avila Garcez and Luís C. Lamb. Neurosymbolic AI: The 3rd wave. *CoRR*, abs/2012.05876, 2020.

[Devlin *et al.*, 2019] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proc. of NAACL*, pages 4171–4186, 2019.

[Di Fabio *et al.*, 2019] Andrea Di Fabio, Simone Conia, and Roberto Navigli. VerbAtlas: a novel large-scale verbal semantic resource and its application to Semantic Role Labeling. In *Proc. of EMNLP-IJCNLP*, pages 627–637, 2019.

[Emelin *et al.*, 2020] Denis Emelin, Ivan Titov, and Rico Sennrich. Detecting Word Sense Disambiguation biases in machine translation for model-agnostic adversarial attacks. In *Proc. of EMNLP*, pages 7635–7653, 2020.

[Faralli *et al.*, 2015] Stefano Faralli, Giovanni Stilo, and Paola Velardi. Large scale homophily analysis in Twitter using a twixonomy. In *Proc. of COLING*, page 2334–2340, 2015.

[Flati *et al.*, 2016] Tiziano Flati, Daniele Vannella, Tommaso Pasini, and Roberto Navigli. MultiWiBi: The Multilingual Wikipedia Bitaxonomy project. *Artif. Intell.*, 241:66–102, 2016.

[Franco-Salvador *et al.*, 2016] Marc Franco-Salvador, Paolo Rosso, and Manuel Montes y Gómez. A systematic study of knowledge graph analysis for cross-language plagiarism detection. *Information Processing & Management*, 52(4), 2016.

[Gregori *et al.*, 2016] Lorenzo Gregori, Alessandro Panunzi, and Andrea Amelio Ravelli. Linking IMAGACT ontology to BabelNet through action videos. In *Proc. of CLiC-it*, 2016.

[Guu *et al.*, 2020] Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat, and Ming-Wei Chang. REALM: retrieval-augmented language model pre-training. *CoRR*, abs/2002.08909, 2020.

[Hovy *et al.*, 2013] Eduard Hovy, Roberto Navigli, and Simone Paolo Ponzetto. Collaboratively built semi-structured content and Artificial Intelligence: The story so far. *Artif. Intell.*, 194:2–27, 2013.

[Iacobacci *et al.*, 2015] Ignacio Iacobacci, Mohammad Taher Pilehvar, and Roberto Navigli. SensEmbed: Learning sense embeddings for word and relational similarity. In *Proc. of ACL-IJCNLP*, pages 95–105, 2015.

[Kingsbury and Palmer, 2002] Paul Kingsbury and Martha Palmer. From TreeBank to PropBank. In *Proc. of LREC*, 2002.

[Klein *et al.*, 2017] Patrick Klein, Simone Paolo Ponzetto, and Goran Glavaš. Improving neural knowledge base completion with cross-lingual projections. In *Proc. of EACL*, 2017.

[Lacerra *et al.*, 2020] Caterina Lacerra, Michele Bevilacqua, Tommaso Pasini, and Roberto Navigli. CSI: A coarse sense inventory for 85% Word Sense Disambiguation. In *Proc. of AAAI*, pages 8123–8130, 2020.

[Lewis *et al.*, 2020] Patrick S. H. Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive NLP tasks. *CoRR*, abs/2005.11401, 2020.

[Luan *et al.*, 2020] Yixing Luan, Bradley Hauer, Lili Mou, and Grzegorz Kondrak. Improving Word Sense Disambiguation with translations. In *Proc. of EMNLP*, pages 4055–4065, 2020.

[Martelli *et al.*, 2021] Federico Martelli, Najla Kalach, Gabriele Tola, and Roberto Navigli. SemEval-2021 Task 2: Multilingual and Cross-lingual Word-in-Context Disambiguation (MCL-WiC). In *Proc. of SemEval*, 2021.

[Maru *et al.*, 2019] Marco Maru, Federico Scozzafava, Federico Martelli, and Roberto Navigli. SyntagNet: Challenging supervised word sense disambiguation with lexical-semantic combinations. In *Proc. of EMNLP-IJCNLP*, pages 3534–3540, 2019.

[McCrae *et al.*, 2020] John Philip McCrae, Alexandre Rademaker, Ewa Rudnicka, and Francis Bond. English WordNet 2020: Improving and extending a WordNet for English using an open-source methodology. In *Proc. of MMW@LREC*, 2020.

[Mikolov *et al.*, 2013] Tomás Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. In *Proc. of ICLR*, 2013.

[Miller, 1995] George A. Miller. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41, 1995.

[Moro and Navigli, 2015] Andrea Moro and Roberto Navigli. SemEval-2015 task 13: Multilingual all-words sense disambiguation and entity linking. In *Proc. of SemEval*, 2015.

[Moro *et al.*, 2013] Andrea Moro, Hong Li, Sebastian Krause, Feiyu Xu, Roberto Navigli, and Hans Uszkoreit. Semantic rule filtering for web-scale relation extraction. In *The Semantic Web (ISWC)*, pages 347–362, 2013.

[Moro *et al.*, 2014] Andrea Moro, Alessandro Raganato, and Roberto Navigli. Entity Linking meets Word Sense Disambiguation: a unified approach. *TACL*, 2:231–244, 2014.

[Navigli and Ponzetto, 2010] Roberto Navigli and Simone Paolo Ponzetto. BabelNet: Building a very large multilingual semantic network. In *Proc. of ACL*, pages 216–225, 2010.

[Navigli and Ponzetto, 2012] Roberto Navigli and Simone Paolo Ponzetto. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artif. Intell.*, 193:217–250, 2012.

[Navigli *et al.*, 2013] Roberto Navigli, David Jurgens, and Daniele Vannella. SemEval-2013 task 12: Multilingual Word Sense Disambiguation. In *Proc. of SemEval*, pages 222–231, 2013.

[Navigli, 2009] Roberto Navigli. Word Sense Disambiguation: A survey. *ACM Comput. Surv.*, 41(2):10:1–10:69, 2009.

[Navigli, 2018] Roberto Navigli. Natural language understanding: Instructions for (present and future) use. In *Proc. of IJCAI*, pages 5697–5702, 2018.

[Panchenko, 2016] Alexander Panchenko. Best of both worlds: Making Word Sense Embeddings interpretable. In *Proc. of LREC*, pages 2649–2655, 2016.

[Pasini and Navigli, 2017] Tommaso Pasini and Roberto Navigli. Train-O-Matic: Large-scale supervised Word Sense Disambiguation in multiple languages without manual training data. In *Proc. of EMNLP*, pages 78–88, 2017.

[Pasini *et al.*, 2020] Tommaso Pasini, Federico Scozzafava, and Bianca Scarlini. CluBERT: A cluster-based approach for learning sense distributions in multiple languages. In *Proc. of ACL*, 2020.

[Pasini *et al.*, 2021] Tommaso Pasini, Alessandro Raganato, and Roberto Navigli. XL-WSD: An extra-large and cross-lingual evaluation framework for word sense disambiguation. In *Proc. of AAAI*, 2021.

[Pennington *et al.*, 2014] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In *Proc. of EMNLP*, pages 1532–1543, 2014.

[Peters *et al.*, 2018] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *Proc. of NAACL*, pages 2227–2237, 2018.

[Petrolito and Bond, 2014] Tommaso Petrolito and Francis Bond. A survey of wordnet annotated corpora. In *Proc. of GWC*, pages 236–245, 2014.

[Pilehvar and Camacho-Collados, 2019] Mohammad Taher Pilehvar and Jose Camacho-Collados. WiC: the word-in-context dataset for evaluating context-sensitive meaning representations. In *Proc. of NAACL*, pages 1267–1273, 2019.

[Post, 2018] Matt Post. A call for clarity in reporting BLEU scores. In *Proc. of WMT*, pages 186–191, 2018.

[Raganato *et al.*, 2017] Alessandro Raganato, Claudio Delli Bovi, and Roberto Navigli. Neural sequence learning models for word sense disambiguation. In *Proc. of EMNLP*, pages 1156–1167, 2017.

[Raganato *et al.*, 2019] Alessandro Raganato, Yves Scherrer, and Jörg Tiedemann. The MuCoW test suite at WMT 2019: Automatically harvested multilingual contrastive Word Sense Disambiguation test sets for Machine Translation. In *Proc. of WMT*, pages 470–480, 2019.

[Raganato *et al.*, 2020a] Alessandro Raganato, Tommaso Pasini, José Camacho-Collados, and Mohammad Taher Pilehvar. XL-WiC: A multilingual benchmark for evaluating semantic contextualization. In *Proc. of EMNLP*, pages 7193–7206, 2020.

[Raganato *et al.*, 2020b] Alessandro Raganato, Yves Scherrer, and Jörg Tiedemann. An evaluation benchmark for testing the Word Sense Disambiguation capabilities of machine translation systems. In *Proc. of LREC*, pages 3668–3675, 2020.

[Rogers *et al.*, 2018] Anna Rogers, Shashwath Hosur Ananthakrishna, and Anna Rumshisky. What's in your embedding, and how it predicts task performance. In *Proc. of COLING*, pages 2690–2703, 2018.

[Scarlini *et al.*, 2019] Bianca Scarlini, Tommaso Pasini, and Roberto Navigli. Just "OneSeC" for producing multilingual sense-annotated data. In *Proc. of ACL*, pages 699–709, 2019.

[Scarlini *et al.*, 2020a] Bianca Scarlini, Tommaso Pasini, and Roberto Navigli. SensEmBERT: Context-enhanced sense embeddings for multilingual Word Sense Disambiguation. In *Proc. of AAAI*, pages 8758–8765, 2020.

[Scarlini *et al.*, 2020b] Bianca Scarlini, Tommaso Pasini, and Roberto Navigli. With more contexts comes better performance: Contextualized sense embeddings for all-round Word Sense Disambiguation. In *Proc. of EMNLP*, pages 3528–3539, 2020.

[Schuler, 2005] Karin Kipper Schuler. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD thesis, 2005.

[Scozzafava *et al.*, 2020] Federico Scozzafava, Marco Maru, Fabrizio Brignone, Giovanni Torrisi, and Roberto Navigli. Personalized PageRank with syntagmatic information for multilingual Word Sense Disambiguation. In *Proc. of ACL*, 2020.

[Speer *et al.*, 2017] Robyn Speer, Joshua Chin, and Catherine Havasi. ConceptNet 5.5: An open multilingual graph of general knowledge. In *Proc. of AAAI*, pages 4444–4451, 2017.

[Tosi and dos Reis, 2021] Mauro Dalle Lucca Tosi and Julio Cesar dos Reis. SciKGraph: A knowledge graph approach to structure a scientific field. *Journal of Informetrics*, 15(1), 2021.

[Vannella *et al.*, 2014] Daniele Vannella, David Jurgens, Daniele Scarfini, Domenico Toscani, and Roberto Navigli. Validating and extending semantic knowledge bases using video games with a purpose. In *Proc. of ACL*, pages 1294–1304, 2014.

[Wang *et al.*, 2019a] Alex Wang, Yada Pruksachatkun, Nikita Nangia, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman. SuperGLUE: A stickier benchmark for general-purpose language understanding systems. In *Proc. of NeurIPS*, pages 3261–3275, 2019.

[Wang *et al.*, 2019b] Qi Wang, Pubali Datta, Wei Yang, Si Liu, Adam Bates, and Carl A. Gunter. Charting the attack surface of trigger-action IoT platforms. In *Proc. of CCS*, 2019.

[Wang *et al.*, 2021] Josiah Wang, Pranava Madhyastha, Josiel Figueiredo, Chiraag Lala, and Lucia Specia. MultiSubs: A large-scale multimodal and multilingual dataset. *CoRR*, 2021.

[Wijeratne *et al.*, 2016] Sanjaya Wijeratne, Lakshika Balasuriya, Amit P. Sheth, and Derek Doran. EmojiNet: Building a machine readable sense inventory for emoji. *CoRR*, 2016.