Toward the automatic extraction of knowledge of usable goods

Mei Uemura, Naho Orita, Naoaki Okazaki and Kentaro Inui
Tohoku University, Japan
{mei.uemura, naho, okazaki, inui}@ecei.tohoku.ac.jp

Abstract

Knowledge of usable goods (e.g., *toothbrush is used to clean the teeth* and *treadmill is used for exercise*) is ubiquitous and in constant demand. This study proposes semantic labels to capture aspects of knowledge of usable goods and builds a benchmark corpus, USABLE GOODS CORPUS, to explore this new semantic labeling task. Our human annotation experiment shows that human annotators can generally identify pieces of information of usable goods in text. Our first attempt toward the automatic identification of such knowledge shows that a model using conditional random fields approaches the human annotation (F score 73.2%). These results together suggest future directions to build a large-scale corpus and improve the automatic identification of knowledge of usable goods.

1 Introduction

A rich body of information extraction techniques focuses on acquiring knowledge from a huge amount of text data (Nickel et al. 2016). This allows large-scale knowledge bases to cover a broad range of knowledge. However, an important subfield of knowledge is not fully addressed: knowledge about use of objects such that *hand sanitizer is used to kill bacteria* and *dental floss is used to remove plaque*. Every object that humans create has its own purpose and function. We call these pieces of information **knowledge of usable goods**. Knowledge of usable goods is ubiquitous and in constant demand. People use search engines to find information on effect caused by using a new product, its proper way to use, and so on.

Knowledge sources that contain such information would also be beneficial for various kinds of natural language processing tasks, such as question answering systems and textual entailment. However, knowledge of usable goods is not thoroughly covered by current knowledge bases because these resources focus on entities (e.g. person or organization) and their relations (e.g. *IsPresidentOf*). Section 4.3 shows the gap between kinds of knowledge available in the current knowledge bases and the ones that we aim to acquire.

To fill in this gap, this study proposes a set of semantic labels to capture knowledge of usable goods and builds a benchmark corpus, USABLE GOODS CORPUS, to explore the automatic extraction of such knowledge. This work begins with focusing on information of health care and household goods such as *air freshener*, *rice cooker*, and *nasal strip*.

We assume that one of the most important aspects of knowledge of usable goods is about effects caused by using/consuming them as in (1). [1]

(1) a. Fish-oils ... are known to <u>reduce inflammation in the body</u>, ... (`fish oil`)

b. Alcohol-based hand sanitizers are more effective at <u>killing microorganisms</u> than

[1]Throughout this paper, each typewriter word in a round bracket (e.g. `toothbrush`) indicates a name of a usable good that corresponds to the title of Wikipedia article.

soaps... (`hand sanitizers`)

c. BB cream and CC cream are both tinted moisturizers ... (`CC cream`)

d. ... the American Dental Association reports that up to 80% of plaque can be eliminated with this method. (`dental floss`)

Humans can easily understand what the effects of these goods are: fish-oils reduce inflammation in the body (1a), hand sanitizers kill microorganisms (1b), BB cream tints and moisturizes skin (1c), and dental floss eliminate plaque (1d). However, the automatic extraction of such knowledge is challenging in that these effects can be expressed in various ways such as a verb phrase (1a), gerund (1b), noun phrase (1c), and clause (1d). This poses a problem that superficial linguistic patterns would not help identifying these kinds of expressions. To gauge difficulties of the automatic acquisition of these pieces of information, we conduct human annotation (Section 4) and automatic identification experiments (Section 5).

The major contributions of this work are: (i) We define a set of semantic labels to capture knowledge of usable goods, suggesting a new semantic labeling task. (ii) We experimentally build a benchmark corpus (Usable Goods Corpus) to explore the automatic extraction of knowledge of usable goods. The corpus and guidelines will be available when this paper is presented. (iii) We present our initial attempts toward the automatic extraction of such knowledge using a sequence labeling method. The results in this experiment provide measures to estimate the complexity of this task and suggest future directions to build a large-scale corpus.

2 Related work

To our knowledge, there is no resource that focuses on knowledge of usable goods. There are manually constructed and relatively accurate lexical resources such as WordNet (Miller 1995) and FrameNet (Baker et al. 1998), but their coverage is inevitably limited and these ontologies do not contain knowledge of our interest. Current large-scale knowledge bases focus on knowledge of entities and their relations, but the coverage of knowledge of usable goods is still sparse as shown in Section 4.3. OpenIE systems (Etzioni et al. 2011) such as TextRunner (Etzioni et al. 2008) and ReVerb (Fader et al. 2011) extract a large number of relations such as ⟨*treadmill, burns, more calories*⟩ using lexico-syntactic patterns from massive corpora drawn from the Web. Though these systems cover a wide variety of relational expressions, they do not intend to extract information of usable goods.

As for extracting information of objects, there is a body of research on the acquisition of telic and agentive roles in the context of generative lexicon theory (Pustejovsky 1991). Pustejovsky proposes qualia structures that define prototypical aspects of word's meaning (Pustejovsky et al. 1993). Of four semantic roles in the qualia structures, the telic role describes the purpose or function of an object (e.g. *read* is a typical telic role for *book*). Computational approaches are suggested to automatically extract expressions of this role from text (Yamada et al. 2007, Cimiano and Wenderoth 2007), but these models tend to focus on taking paraphrases of "*using X*", rather than the expressions of purpose or function of objects. While the telic roles cover a broader range of expressions (probably due to the unspecified definition of telicity in the original theory), our work focuses on effects caused by using/consuming objects, standing as complementary to these previous studies.

Information extraction research in biomedical domains concerns effects caused by using drugs such that drug X causes adverse effect Y (Gurulingappa et al. 2012). This kind of information may overlap with what we aim to acquire, but ontologies in these studies are domain-specific such as protein interactions and adverse effects, contrary to our interest, which is more generic.

In summary, neither existing resources nor methods focus on knowledge of usable goods. In the next section, we propose a set of semantic labels that captures aspects of knowledge of usable goods.

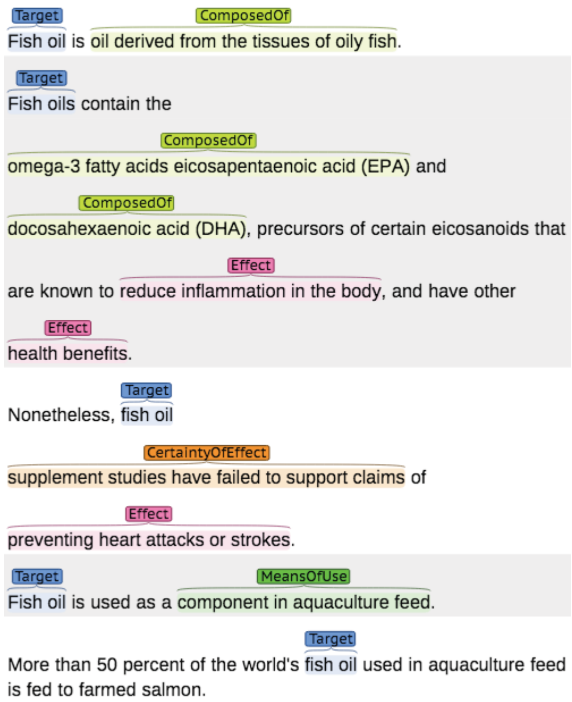

Figure 1: Assigned semantic labels for an excerpt from Wikipedia article on `fish oil` [3].

3 Semantic labels for capturing knowledge of usable goods

To capture aspects of knowledge of usable goods, we define semantic labels as in Table 1 based on observation of 25 Wikipedia lead sections on health care and household goods. The Wikipedia lead [4] is normally a summary of its most important contents, and therefore it may allow us to get rich information from relatively small amount of data.

As shown in (1), we assume that one of the most important aspects of knowledge of usable goods is about effects caused by the use of goods. We also observe that there are various kinds of information that express degree/certainty of effects and conditions for the occurrence of effects. Semantic labels in Table 1 are intended to capture these kinds of information. In addition to these semantic labels, we define a label TARGET for name and other expressions that refer to a usable good in the article. Names of usable goods essentially correspond to titles of Wikipedia articles, which refer to the topic of the text. Figure 1 shows how these labels are assigned to pieces of information about `fish oil`.

The annotation guidelines are designed to increase consistency. We define rules for segmentation in the guidelines, along with definition and examples of each label. To capture various linguistic expressions as illustrated in (1), we do not define a particular syntactic category for each label. All labels can take any type of linguistic constituent, but function words that do not contribute to the meaning are not included in each segment to avoid inconsistency. For example, we ask annotators mark *define the eyes* in *Eyeliner is a cosmetic used to define the eyes* as EFFECT (i.e., *to* is not included).

The set of semantic labels in Table 1 proposes a new semantic labeling task. To gauge the complexity of this task, we conduct human annotation experiment (Section 4) and automatic identification experiment (Section 5) as follows.

4 Annotation

We conduct a pilot annotation experiment to measure the complexity of this task. Measures of inter-annotator agreement and distributional analysis of the annotated data provide indications to improve the annotation schema for building a large-scale corpus in the future. This pilot corpus is also used for the automatic identification in Section 5. The following describes our annotation experiment in details.

4.1 Data: snippets from Wikipedia leads

We collect 200 English Wikipedia articles for annotation. Each article is about a health care or household goods such as *toothpaste*, *tea cosy*, and *dishwasher*. We choose these items using Ama-

[3]The phrase *precursors of certain eicosanoids* in Figure 1 is not COMPOSED OF for `fish oil` because this phrase just denotes an explanation of the constituents of `fish oil`, *omega-3 fatty acids eicosapentaenoic acid (EPA)* and *docosahexaenoic acid (DHA)* are constituents of `fish oil`.

[4]It is also known as the introduction of a Wikipedia article, the section before the table of contents and the first heading.

Label	Definition	Example
TARGET	expression referring to a target object, including aliases and pronouns	BB cream stands for blemish balm, blemish base (`BB cream`)
EFFECT	effect caused by using TARGET	to decorate and protect the nail plates (`nail polish`)
NULL EFFECT	description that states there is no EFFECT	The myth of its effectiveness (`bear's grease`)
DEGREE OF EFFECT	description that states a degree of EFFECT	poor substitute for protective clothing (`barrier cream`)
CERTAINTY OF EFFECT	description that states a certainty/reliability of EFFECT	a have not been proven to give lasting or major positive effects (`anti-aging cream`)
MEANS OF USE	description of how TARGET is used	is applied around the contours of the eye(s) (`eye liner`)
COMPOSED OF	material/ingredient that composes of TARGET	consisting mainly of triglycerides (`egg oil`)
PART OF	material/object that TARGET is a part of	Cinnamon is a spice obtained from the inner bark (`cinnamon`)
LOCATION	description of where TARGET is used	often used where sunlight can impair seeing (`eye black`)
TIME	description of when TARGET is used	soon after birth (`kohl`)
USER	description of who uses/receives EFFECT	mothers would apply kohl to their infants' eyes (`kohl`)
VERSION	different version of TARGET	It is distributed as a liquid or a soft solid (`lip gloss`)

Table 1: Semantic labels to capture knowledge of usable goods

zon categories and products lists. [5] All of chosen items are expressed as common nouns. We exclude any company-specific product.

We extract the lead section of each Wikipedia article for annotation. We use at most the first 5 sentences of the lead to even out the number of sentences, ending up 792 sentences in total from 200 lead snippets.

Each annotator annotates same 100 snippets using brat (Stenetorp et al. 2012). Figure 1 shows an example of annotation. In addition to these 100 snippets, one of the two annotators annotates another 100 snippets, resulting in 200 annotated snippets. We use this set of 200 annotated snippets as the gold standard dataset in the following automatic identification experiment.

4.2 Evaluation

Two annotators were given the guidelines and a short training on texts not included in the corpus. Their task is to annotate linguistic expressions that correspond to the semantic labels in Table 1.

Table 2 shows F scores for inter-annotator agreement. We compute these scores in two ways: (i) **strict match**: the starting and ending of the segment to be the same, (ii) **lenient match**: the starting and ending of the segment do not have to be the same but they overlap. We obtain Kappa coefficient of 0.57 in the lenient match, suggesting moderate agreement (Landis and Koch 1977). F score in the strict match (micro average 36.8%) seems to be reasonable because we give annotators unparsed raw text to explore the range of linguistic expressions. Most segmentation disagreements occur in deciding whether to include function words (e.g. *to protect skin* or *protect skin*).

Type of match	F-score (%)
lenient match (micro average)	**77.2**
lenient match (macro average)	52.5
strict match (micro average)	36.8
strict match (macro average)	27.1

Table 2: Inter-annotator agreement

In addition, there are label disagreements accounting for 20% of segment pairs that either partially or completely match. For example, one annotator marks *hair and skin care* in (2) as EFFECT and the other does so as MEANS OF USE, where both labels seem to be appropriate.

[5] `https://www.amazon.com/gp/help/customer/display.html`

Label	Annotator A	Annotator B
EFFECT	195 (31.7%)	189 (32.8%)
CERTAINTY OF EFFECT	32 (10.1%)	19 (3.3%)
DEGREE OF EFFECT	13 (2.1%)	13 (2.1%)
NULL EFFECT	0 (0.0%)	0 (0.0%)
MEANS OF USE	115 (18.7%)	59 (9.6%)
COMPOSED OF	98 (15.9%)	112 (19.4%)
PART OF	12 (1.9%)	14 (2.3%)
LOCATION	16 (2.6%)	26 (4.2%)
TIME	15 (2.4%)	16 (2.6%)
USER	19 (3.1%)	25 (4.1%)
VERSION	100 (16.2%)	103 (16.7%)
Total	616	576

Table 3: Numbers of the annotated labels

(2) It is used for topical applications such as hair and skin care. (`egg oil`)

This kind of disagreement may reflect differences in annotators' background knowledge. *Hair and skin care* does not explicitly denote the effect, but people usually have the relevant knowledge such that skin care improves skin elasticity.

The following (3) shows an example of disagreement between VERSION and COMPOSED OF.

(3) A wet wipe ... is a small moistened piece of paper or cloth ... (`Wet wipe`)

Paper and *cloth* in (3) could be VERSION of `wet wipe`, but they are also materials that compose of `wet wipe`. Both VERSION and COMPOSEDOF are valid in this example.

These examples of label disagreement suggest that single-label annotation would not be able to sufficiently capture the knowledge of usable goods. Allowing multi-labeling would be one direction for further improvement.

4.3 The distribution of the annotated data

We conduct distributional analysis to examine the extent to which the proposed semantic labels capture information of usable goods. Table 3 breaks up numbers of the annotated instances by two annotators. EFFECT results in the most frequent one, suggesting its significance at least in the domain of health care and household goods.

On the other hand, there are a few number of instances for CERTAINTY OF EFFECT, DEGREE OF EFFECT, NULL EFFECT, PART OF, LOCATION, TIME, and USER. This may due to the content of the Wikipedia leads. These kinds of more precise information would usually appear after the lead section. [6]

We further examine the syntactic distribution of EFFECT instances as in Table 4. The majority of EFFECT instances are represented as verb phrases and there is a variation in those instances such as *darken the eyelids* (`kohl`), *minimize shininess caused by oily skin* (`face powder`), *tones the face* (`face powder`), *reflect light at different angles* (`glitter`) and so on, in addition to typical causal expressions such as *causes anesthesia* (`anesthetic`), *prevent snoring* (`nasal strip`), and *promote oral hygiene* (`toothpaste`). An example of noun phrase in Table 4 suggests an interesting problem in that *lacquer* itself is a usable good but also means effect caused by using a nail polish. This kind of information structure has not been addressed in previous work on information extraction.

Overall, we find that 81.8% of instances occur with TARGET in the same sentence. The remaining cases involve long-distance dependencies across the sentence. This distribution suggests that we do not need to annotate the relation between TARGET and each label and we could exploit these inter-sentential relations in the automatic identification task. The following Section 5 shows our automatic identification experiment using this distributional property.

4.4 Comparison with current knowledge base

The above human annotation experiment shows that Wikipedia leads contain a reasonable amount of information on effects caused by us-

[6] Besides these semantic labels, there are other descriptions on the manufacturing process and history of usable goods as in (4).

(4) a. (herbal distillate) ... obtained by steam distillation or hydrodistillation (`herbal distillate`)

b. Modern perfumery began in late 19th century with the commercial synthesis (`perfume`)

Phrase type		# of instances	Example
Verb phrase	transitive	121 (60.8%)	... that can be applied to decorate and protect the nail plates (`nail polish`)
	intransitive	14 (7.0%)	It generally stays on longer than lipstick (`lip stain`)
Noun phrase		44 (22.1%)	Nail polish is a lacquer (`nail polish`)
Adjective phrase		19 (9.5%)	Choline is a water-soluble nutrient (`choline`)
Sentence		1 (0.5%)	... reports that up to 80% of plaque can be eliminated (`dental floss`)
Total		199	

Table 4: Syntactic distribution of EFFECT instances

ing goods. However, it is possible that existing knowledge bases might have already acquired such knowledge. To examine the coverage of the current knowledge base, we compare ConceptNet (Speer and Havasi 2012) with our corpus.

For comparison, we use 100 usable goods in our corpus such as *ice pack*, *hand sanitizer* and *perfume*. We then manually select 4 out of 39 pre-defined relations in ConceptNet that could be associated with effect expressions such as USED FOR, CAPABLE OF, CAUSES DESIRE, and CAUSES. Of 100 usable goods, 27 usable goods have pieces of knowledge that are expressed with the above relations such as ⟨*hand sanitizer*, CAUSES, *clean hand*⟩ and ⟨*Toothpaste*, CAPABLE OF, *help remove plaque*⟩.

In short, though ConceptNet contains information of our interest, the coverage is still not sufficient (27/100 usable goods). The automatic extraction of information of usable goods would help populate this kind of knowledge base. The next section shows our initial attempt toward the automatic extraction of knowledge of usable goods.

5 Sequence labeling model for identifying information of usable goods

This section presents our experiment for automatically identifying information of usable goods. The results provide baseline measures for this new semantic labeling task and suggest potential directions for improvement.

Section 4.3 shows that almost all instances in our corpus occur with TARGET in the same sentence. We exploit this distributional property by using TARGET words as a cue to find information of usable goods and pose this task as a sequence labeling problem. We use Conditional Random Fields (CRFs), a popular approach to solve sequence labeling problems (Lafferty et al. 2001). CRFsuite [7] is used as an implementation of CRF for our purpose.

5.1 Experimental Settings

The training and test data consists of 792 sentences from 200 Wikipedia snippets (see Section 4.1). We select the four most frequent labels in the corpus, EFFECT, MEANS OF USE, COMPOSED OF and VERSION, for evaluation.

For the data pre-processing, we first parse the raw text and assign a part of speech tag and a named entity tag to each word using Stanford CoreNLP (Manning et al. 2014). Then we add a semantic label to each word with BIO format (Beginning, Inside and Outside).

5.2 Features

Features shown in Table 5 are used for training. We use these features within a window of ±3 around the current word. Some of these features are used in combination with another feature as shown in Table 5.

In addition to standard features, we add three features to exploit the characteristics of this corpus: **Target**, **Disease** and **Repeat**. **Target** feature is true when the current word is same as the title of Wikipedia article. **Disease** feature is true when the current word is in a list of disease names that we create using Freebase (Bollacker et al. 2008). This feature is intended to capture effect expressions that include disease names such as *provoke allergy and asthma symptoms* (`air freshener`). **Repeat** feature is true when the current word has already been appeared in the sentence. This feature is intended to capture a parallel structure that is often used

[7] `http://www.chokkan.org/software/crfsuite/`

Feature	Definition	Example
Token	current word	Perfume
Lower	lowercased current word	perfume
POS	POS tag of the current word	NNS
NE	named entity type of the current word	O
Target	whether the current word is TARGET	True
Disease	whether the current word is a disease name	False
Repeat	whether the current word has been appeared in the sentence	False
Combination	Definition	Example
Token + Lower	current word and lowercased current word	(Perfume, perfume)
Token + POS	current word and its POS tag	(Perfume, NNS)
Lower + POS	lowercased current word and its POS tag	(perfume, NNS)
Disease + POS	POS tag and whether the current word is a disease name	(NNS, False)

Table 5: Features

to express VERSION and COMPOSED OF.

5.3 Evaluation

We compute precision, recall and F1 measure using ten fold cross validation. We compute these scores in two ways, lenient match and strict match as in the human annotation experiment (see Section 4.2). Table 6 shows results.

F score in the lenient match (73.2%) approaches the human annotation performance (81.9%). This suggests that the model is able to identify labels to some extent. For example, the model recognizes typical lexico-syntactic patterns such as *be used to* in *(wallpaper) is used to cover and decorate the interior walls* and *be designed to* in *(rice cooker) is designed to boil or steam rice*. Furthermore, the model captures various effect expressions such as an adjective phrase (5a), verb phrase (5b), and gerund (5c).

(5) a. Chandeliers are often ornate, and normally use... (`chandelier`)

b. A diuretic is any substance that promotes the production of urine. (`diuretic`)

c. An espresso machine brews coffee by forcing pressurized water near boiling point... (`espresso machine`)

On the other hand, the segmentation problem as discussed in the human annotation experiment influences the F score in the strict match (13.7%).

In sum, though there is the segmentation problem derived from the annotation, the results in the lenient match suggest that the model can identify information of usable goods to some extent. Improving the annotation schema and increasing the size of the corpus would be promising directions for future work.

6 Conclusion

This paper proposes semantic labels to capture aspects of knowledge of usable goods. We design annotation schema and build the benchmark corpus, USABLE GOODS CORPUS, based on the proposed semantic labels. Our human annotation experiment shows that (i) while there is the segmentation mismatch problem, human annotators can generally identify pieces of information of usable goods, and (ii) Wikipedia leads contain a reasonable amount of information on effects caused by using goods in contrast to the coverage of the current knowledge base. The automatic identification experiment shows that despite of the influence of the segmentation problem in the human annotation, the model can to some extent identify pieces of information of usable goods.

Our next steps are to alleviate the segmentation problem and increase the corpus size. With these goals in mind, we plan to revise the annotation schema as follows: (a) Some semantic labels do not seem to be important as seen in the statistics in Table 3. Reducing the variation of the semantic labels is a reasonable direction. (b) Defining a syntactic category for each label and giving annotators/models parsed text would increase consistency in the segmentation. (c) These simplifications (a,b) would allow us to

	Label	Precision (%)	Recall (%)	F score (%)
strict match	Effect	24.2	24.1	24.1
	Means of Use	10.3	4.9	6.6
	Composed of	13.0	13.0	11.4
	Version	15.9	8.5	11.1
	micro average	16.0	12.21	13.7
	macro average	20.2	15.4	17.4
lenient match	Effect	79.4	71.8	74.1
	Means of Use	75.0	58.1	60.2
	Composed of	71.8	60.6	63.1
	Version	75.9	64.6	66.5
	micro average	72.7	73.6	**73.2**
	macro average	51.9	39.2	41.7
Results in human annotation				
lenient match	micro average	81.9	81.9	81.9
	macro average	71.2	65.7	66.9

Table 6: 10-fold cross-validation

try crowdsourcing annotation to increase the size of the corpus.

Acknowledgements

This study is supported by CREST, JST and JSPS KAKENHI Grant Number JP15H05318.

References

Collin F Baker, Charles J Fillmore, and John B Lowe. The Berkeley FrameNet Project. In *Proceedings of the 17th International Conference on Computational Linguistics*, pages 86–90. Association for Computational Linguistics, 1998.

Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge. In *Proceedings of ACM Special Interest Group on Management of Data*, pages 1247–1250, 2008.

Philipp Cimiano and Johanna Wenderoth. Automatic Acquisition of Ranked Qualia Structures from the Web. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, pages 888–895, 2007.

Oren Etzioni, Michele Banko, Stephen Soderland, and Daniel S. Weld. Open information extraction from the web. *Communications of the Association for Computing Machinery*, 51 (12):68–74, 2008.

Oren Etzioni, Anthony Fader, Janara Christensen, Stephen Soderland, and Mausam. Open Information Extraction: the Second Generation. In *Internartional Joint Conference on Artificial Intelligence*, pages 3–10, 2011.

Anthony Fader, Stephen Soderland, and Oren Etzioni. Identifying relations for open information extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1535–1545, 2011.

Harsha Gurulingappa, Abdul Mateen Rajput, Angus Roberts, Juliane Fluck, Martin Hofmann-Apitius, and Luca Toldo. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. *Journal of biomedical informatics*, 45(5):885–892, 2012.

John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 282–289, 2001.

J Richard Landis and Gary G Koch. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159–174, 1977.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. The Stanford CoreNLP Natural Language Processing Toolkit. In *Association for Computational Linguistics System Demonstrations*, pages 55–60, 2014.

George A Miller. Wordnet: a lexical database for English. *Communications of the Association for Computing Machinery*, 38(11):39–41, 1995.

Maximilian Nickel, Kevin Murphy, Volker Tresp, and Evgeniy Gabrilovich. A Review of Relational Machine Learning for Knowledge Graphs. *Proceedings of the Instute of Electrical and Electronics Engineers*, 104(1):11–33, 2016.

James Pustejovsky. The Generative Lexicon. *Computational linguistics*, 17(4):409–441, 1991.

James Pustejovsky, Peter Anick, and Sabine Bergler. Lexical Semantic Techniques for Corpus Analysis. *Computational Linguistics*, 19 (2):331–358, 1993.

Robert Speer and Catherine Havasi. Representing General Relational Knowledge in ConceptNet 5. In *Language Resources and Evaluation Conference*, pages 3679–3686, 2012.

Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. BRAT: a Web-based Tool for NLP-Assisted Text Annotation. In *Proceedings of the Demonstrations at the 13th Eutopean Chapter of the Association for Computational Linguistics*, pages 102–107, 2012.

Ichiro Yamada, Timothy Baldwin, Hideki Sumiyoshi, Masahiro Shibata, and Yagi Nobuyuki. Automatic Acquisition of Qualia Structure from Corpus Data. *The Institute of Electronics, Information and Communication Engineers transactions on information and systems*, 90(10):1534–1541, 2007.