ScienceExamCER: A High-Density Fine-Grained Science-Domain Corpus for Common Entity Recognition

Hannah Smith, Zeyu Zhang, John Culnan, Peter Jansen
School of Information, University of Arizona
Tucson, Arizona, USA
pajansen@email.arizona.edu

Abstract

Named entity recognition identifies common classes of entities in text, but these entity labels are generally sparse, limiting utility to downstream tasks. In this work we present ScienceExamCER, a densely-labeled semantic classification corpus of 133k mentions in the science exam domain where nearly all (96%) of content words have been annotated with one or more fine-grained semantic class labels including taxonomic groups, meronym groups, verb/action groups, properties and values, and synonyms. Semantic class labels are drawn from a manually-constructed fine-grained typology of 601 classes generated through a data-driven analysis of 4,239 science exam questions. We show an off-the-shelf BERT-based named entity recognition model modified for multi-label classification achieves an accuracy of 0.85 F1 on this task, suggesting strong utility for downstream tasks in science domain question answering requiring densely-labeled semantic classification.

Keywords: named entity recognition, corpus, science

1. Introduction

Named entity recognition (NER) (Grishman and Sundheim, 1996) is a common natural language processing task that aims to abstract or categorize common classes of noun phrases in text, such as identifying “Arthur” as a person or “Montreal” as a location. This high-level categorization of important entities in text is a staple of most modern NLP pipelines, and has a variety of applications for higher-level tasks including information extraction (Valenzuela-Escárciga et al., 2016), knowledge base population (Dredze et al., 2010), and question answering (Abujabal et al., 2017).

Named entity recognition identifies common classes of entities in text, but these entity labels are generally sparse (typically occurring for between 10% to 20% of words in a corpus, see Section 3.4.), limiting utility to downstream tasks. In this work, we introduce the idea of common entity recognition (CER), which aims to tag all content words in text with an appropriate fine-grained semantic class. CER allows text to be automatically annotated with a much more high-level categorization in the context of standardized science exams is shown in Figure 1. While CoreNLP (Manning et al., 2014) does not locate any entities in the sentence “Rolanda is growing tomato plants in her garden”, our CER annotation and system abstracts this sentence to “[Rolanda]Human [is] StateOfBeing [growing] Growth/ActionsForAgriculture [tomato] Food [plants] Plant [in] RelativeLocation her [garden] ManmadeLocation”.

We detail corpus and ontology/typology construction in Section 3, including a comparison of mention density with other common corpora. Automated evaluations of CER performance are shown in Section 4, including an analyses of the training data requirements of this fine-grained classification, as well as an error analysis.

2. Related Work

Common sets of entity labels (or typologies) have expanded from early experiments with a single label, organization (Rau, 1991), to the 7 common MUC-6 types (Grishman and Sundheim, 1996) typically used by NER systems, including named entities (person, organization, location), temporal mentions (date, time), and numeric categories (money, percent). Subsets of the MUC-6 types have been included in the typologies of benchmark NER corpora, including CoNLL-2003 (Sang and De Meulder, 2003), OntoNotes (Weischedel et al., 2013), and BBN (Weischedel and Brunstein, 2005).

Sekine et al. (2002) proposed an extended hierarchy of MUC-6 types expanded to include 150 open-domain category labels. While most of these category labels are named entities, Sekine et al. include 10 measurement categories (e.g. weight, speed, temperature) and 3 high-level natural object categories (animal, vegetable, mineral) that most closely relate to the 601 fine-grained science categories in this work. A subsequent version, the Extended Named Entity (ENE) Ontology (Sekine, 2008), expands the

arXiv:1911.10436v1 [cs.CL] 24 Nov 2019
Rolanda is growing tomato plants in her garden. She has created a compost pile and has been adding compost around her tomato plants to help fertilize them. Compost is solid waste in which organic material is broken down by microorganisms in the presence of oxygen to where it can be safely stored, handled, and applied to the environment.

On what does Rolanda primarily rely in order for composting to work?

(A) producers
(B) consumers
(C) scavengers
(D) decomposers

Figure 1: An example standardized science exam question densely annotated with one or more fine-grained semantic categories for nearly each word. This 4-choice multiple choice question (here, under the curriculum topic “The Interdependence of Life > The Food Chain > Decomposers”) is one of 4,239 drawn from the ARC corpus and densely annotated in this work.

typology to 200 classes, including 19 fine-grained expansions of the natural_object type, such as bird or reptile, as well as adding 5 meronym categories, such as plant_part, that further relax the working definition of named entities from proper names to include other categories such as the.

While open-domain typologies are common, domain-specific typologies and corpora are also popular, occasionally making use of existing domain ontologies to reduce the burden in manually generating fine-grained typologies, such as the manual creation of the fine-grained science-domain typology in this work. An extreme example of fine-grained NER is the MedMentions corpus (Murty et al., 2018), which contains 246k mentions labelled with Universal Medical Language System (UMLS) (Bodenreider, 2004) categories, a fine-grained ontology of over 3.5 million medical concepts. Similarly, large knowledge bases can be filtered to automatically produce fine-grained typologies (as in FIGER (Ling and Weld, 2012) and HYENA (Yosef et al., 2012)), or used to bootstrap the entity classification process in manually-generated typologies. Magnini et al. (2002) demonstrate combining WordNet predicates (Fellbaum, 1998) with approximately 200 handcoded rules can achieve an F1 score of 0.85 on recognizing 10 common entity types, while Ritter et al. (2011) use distantly supervised topic modeling over Freebase entities (Bollacker et al., 2008) to perform named entity recognition on social media posts, achieving an F1 score of 0.59 on 10 common entity types. With respect to larger typologies, Del Corro et al. (2015) perform super-fine-grained entity typing using the 16k fine-grained WordNet types under the high-level taxonomic categories of person, organization, and location, achieving a manually-evaluated precision of 59.9% on the CoNLL corpus and 28.3% on New York Times news articles. For smaller manually-generated typologies, Mai et al. (2018) demonstrate a model combining LSTMs, CNNs, CRFs, and dictionary-based methods can achieve an F1 of 83.1 on an in-house corpus labeled with Sekine’s 200-class ENE ontology.

NER has historically been approached using a wide variety of methods, including rules (Hanisch et al., 2005), feature-based machine learning systems (Mayfield et al., 2003), conditional random fields (Greenberg et al., 2018), contextualized embeddings, (Peters et al., 2018), and combinations thereof. Qu et al. (2016) demonstrate that it is possible to use a conditional random field model to transfer NER performance between datasets, at least in part. Ma et al. (2016) show embedding models can transfer performance in zero-shot settings on fine-grained named entity classification. Expanding on this, recent transformer models (Peters et al., 2018) have shown strong transfer performance on a variety of text classification tasks including named entity recognition using large pretrained contextualized embeddings that are fine-tuned on comparatively small in-domain corpora. In this work we make use of an off-the-shelf bidirectional transformer (BERT) NER system modified to support multi-label classification, and demonstrate strong performance on the fine-grained common entity recognition task.

3. Data and Annotation

3.1. Corpus

We annotate fine-grained semantic classes on standardized science exam questions drawn from the Aristo Reasoning Challenge (ARC) corpus (Clark et al., 2018), which contains 7,787 elementary and middle school (3rd through
We conducted a large data-driven analysis of the 4,239 science exam questions with the aim of identifying a set of high-level semantic categories that would provide near total coverage for classifying or grouping nearly all of the 156k words found across the question and answer text in this corpus. While named entity recognition typically focuses on proper names with specific referents (Nadeau and Sekine, 2007), in the end we arrived at creating 601 fine-grained categories spanning 6 classes of groups:

**Taxonomic Groups:** High-level categories expressing taxonomic membership, such as that a hummingbird is a kind of bird. This (or stricter interpretations) is the common form of entity classification in most named entity recognition corpora.

**Meronym Groups:** Categories expressing part-of relations, such as that a fin is a part of an aquatic animal, an x-axis is a part of a graphical representation, or that an individual is a part of a group.

**Action Groups:** Collections of action words that tend to describe similar ideas. For example, decrease, increase, contract, expand, inflate, deflate, accelerate, decelerate, lower, raise all describe a group of actions that involve increasing or decreasing quantities.

**Thematic word groups:** Groups of words that surround a particular topic. For example, observe, conduct an experiment, compare, study, consider, test, collect, record, gather, examine, and research are some of the words included in the performing research using the scientific method semantic class.

**Properties and Values:** Common science-domain properties of objects, such as mass, size, or conductivity, typically grouped with common values they might take, such as soft, brittle, or hard in the case of hardness.

**Synonyms:** Groups of words that tend to express similar ideas in the context of science exams. For example, disease, infection, and sick all convey the notion of illness.

To identify specific instances of these categories in the science exam domain, we first sorted questions into fine-grained curriculum topics using the 406 detailed science-domain question classification labels of Xu et al. (2019), noting that common categories of words tended to emerge
upon detailed manual inspection when questions on similar topics were examined together. We proceeded through several iterations of this process, recording candidate high-level semantic classes, as well as seed words that belonged to those categories. After assembling a large list of candidate categories, we further enumerated the seed words with encyclopedic knowledge manually through web searches. For example, while the annotators may have only observed the words Sun and Proxima Centauri in the corpus for the Star category, we would manually expand this to also include other nearby stars such as Vega, Polaris, and Wolf 359.

As a final step, we automatically expanded the seed word list to include lexical variations of each manually added word by first using pretrained GLoVe embeddings (Pennington et al., 2014) to compute the top-N most similar words to a given seed word using cosine similarity, then using several low-precision high-recall heuristics to identify words that had the potential to be lexical variations of an existing word on the seed list. We then generated a frequency histogram of any word present in the corpus that did not yet belong to at least one semantic category, and either placed it in an existing category, or formed a new category for that word and repeated the expansion process for seed words. This detailed manual category development process required approximately three weeks of annotator time, ultimately arriving at a list of 601 high-level semantic categories, with an extensive list of both manually and automatically populated seed words for each category. The full list of semantic categories and seed words is included in the supplementary material.

### 3.3. Annotation Procedure

Annotating a large set of semantic classes onto more than one hundred thousand words presents challenges with annotation consistency and tractability. It would be challenging for crowdsworkers to learn a detailed set of 601 fine-grained semantic categories, and extremely time consuming for research assistants to traditionally annotate a collection at this scale. To overcome these challenges, we modified the annotation task to automatically preannotate the entire corpus using the large set of bootstrapped seed words associated with each semantic class, effectively pre-annotating each word with a set of possible semantic category labels. These preannotated labels are effectively low-precision and high-recall, most often containing the correct label(s) for a given mention, but also containing other incorrect labels that must be manually removed by an annotator. A total of 226k preannotated mentions were generated (an average of 1.5 per word), which was reduced to 133k mentions (0.9 per word) after incorrect labels were removed by the annotator. We used the BRAT annotation tool (Stenetorp et al., 2012) for the label removal step. To ease the annotator’s need for switching semantic contexts, questions were presented to the annotator sorted by curriculum topic using the question classification annotation of Xu et al. (2019). The annotation procedure took approximately 2.5 minutes per question, for a total of 200 hours.

A clear question with this “preannotate-then-filter” annotation protocol is how well this procedure is able to provide both coverage and accurate labels for the words in the corpus. Our analysis in Section 3.4 shows that after annotation, 96% of content words and 75% of all words have at least one gold semantic category label, suggesting this protocol allows for near-complete coverage of content words at a fraction of the time required to make accurate 601-class annotation judgements at scale. Both our interannotator agreement (included below) and automatic classification performance are high, suggesting adequate annotated label accuracy.

### Label distribution: Named entity corpora often have many labels in their taxonomies, but the majority of mentions tend to cluster around a small set of possible labels (Choi et al., 2018). The distribution of most frequent labels after annotation is shown in Table 1. The usage of the 601 total semantic class labels in this corpus is well distributed, with the 356 most-frequent types covering 95% of the total mentions, while 479 types cover 99% of mentions. At the 99% level, categories (for example, Geometric Qualities, such as angle, slope, or circumference) still contain 16 mentions, highlighting the scale of the corpus.

### Interannotator agreement: Each question was annotated by a single annotator. A second annotator was trained in the annotation procedure and re-annotated 50 questions totalling 1,756 tokens. Between both annotators, a total of 1,369 mentions were annotated with semantic class labels. Total percent agreement across both annotators was 76%. Upon inspection, labeling multi-word sequences as either a single mention or multiple smaller mentions was a frequent source of disagreement. When these cases were removed, percent agreement rose to 83%.

### 3.4. Mention density comparison

To increase the utility of our common entity corpus for downstream tasks, one of the design goals was to provide at least one high-level semantic category to nearly every word in the corpus. To measure this we define the notion of the mention density of a corpus as the proportion of words that contain at least one entity label. We compare the mention density of this corpus with the English subsets of the four benchmark named entity recognition corpora listed below:

---

1. Because the bootstrapped preannotation procedure reduces the set of possible labels for a given mention from 601 to an average of approximately 2 (the average number of preannotated labels per annotated word), Cohen’s Kappa (Cohen, 1960) would either be artificially inflated (if treating the annotation as a 601 class labeling problem) or reduced (if treating annotation as a 2 class problem). As such we report raw percent agreement, which (as critiqued by Cohen) has known problems when dealing with highly skewed frequency distributions of labels, particularly when few labels are present. Here, the number of label categories is high, and (as shown in Table 1) the frequency of labels is well distributed across the label set. As such, the inflation of the percent agreement statistic is likely to be minimal.

2. Specifically, the proportion of non-punctuation tokens in a BIO-formatted corpus that are labelled with either a B (beginning) or I (inside) tag.
The analysis of mention density is shown in Table 2. Overall, the mention density of this science corpus is 75%, meaning that 75% of all words in the corpus are annotated with at least one high-level semantic category. When considering only content words (here, determined to be nouns, verbs, adjectives, adverbs, and numbers), this proportion increases to 96%. The mention density for the named entity corpora examined in Table 2 ranges between 12% and 59% for all words, and 18% to 62% when considering only content words. At 62%, the GUM corpus contains the next-nearest mention density to the ScienceExamCER corpus, however a large portion of those mentions (40.5%) of words use the high-level object or abstract labels, and as such are of limited informativeness to downstream tasks. BBN, the corpus with the next-nearest mention density to GUM, has labels for only 36% of it content words, and 25% of all words.

Table 2: Summary statistics including mention density for the ScienceExamCER corpus, as well as for four other common benchmark corpora. At 96%, the ScienceExamCER is significantly more densely labeled than the next-nearest corpus. (*) denotes that approximately 16k spans have multiple labels, and as such the total mentions exceeds the total labeled words.

| Measure                | ScienceExamCER | OntoNotes 5 | BBN | GUM | CoNLL 2003 |
|------------------------|----------------|-------------|-----|-----|------------|
| Entity Categories      | 601            | 18          | 64  | 11  | 4          |
| Total Mentions         | 133k*          | 162k        | 172k| 11k | 35k        |
| Words                  | 156k           | 2.44M       | 1.05M| 55k | 264k       |
| Labeled Words          | 117k           | 284k        | 257k| 33k | 51k        |
| Mention Density (overall) | 75%          | 12%         | 25% | 59% | 19%        |
| Content Words          | 104k           | 1.39M       | 677k| 34k | 190k       |
| Labelled Content Words | 100k           | 255k        | 243k| 21k | 50k        |
| Mention Density (Content Words) | 96%           | 18%         | 36% | 62% | 27%        |

Table 3: Summary statistics for the training, evaluation, and test sets used for evaluating semantic category classification.

CoNLL (Sang and De Meulder, 2003): The CoNLL 2003 Named Entity Recognition Shared Task corpus, which includes 4 entity labels that are a subset of the MUC-6 typology: person, location, organization and miscellaneous.

OntoNotes 5.0 (Weischedel et al., 2013): A large multi-genre corpus of news media, blog, newsgroup, and conversational text, annotated with 18 entity labels, including the MUC-6 types.

BBN (Weischedel and Brunstein, 2005): A corpus of news text annotated with 21 course entity types, including 12 named entity types (e.g., person, organization, product) and 7 numeric types (e.g., date, percent, cardinal number). The full set of entity labels includes 64 fine-grained types.

GUM (Zeldes, 2017): An open-domain corpus annotated with a collapsed set of OntoNote entities reduced to 11 entity types, such as person, organization, or place. Two additional catch-all tags are added, object and abstract, which provide high-level but minimally informative categorical information for large noun phrases. Approximately 40.5% of the labelled words in this corpus are labelled as either object or abstract.

4. Experimental Results

4.1. Model

Our semantic class labeling task is conceptually similar to named entity recognition or entity typing, only requiring a label for nearly every word in an input sentence. In light of this, here we use an off-the-shelf named entity recognition model, and show it also performs well on the densely-labeled common entity recognition task.

Recently, pretrained bidirectional encoder representations from transformer (BERT) models (Devlin et al., 2018) have shown state-of-the-art performance at both named entity recognition as well as a variety of other token-level classification tasks. In this work, we use an off-the-shelf implementation of a BERT-based named entity recognition system, BERT-NER[^1]. Most approaches to named entity recognition model the task as a single-label prediction task, where each word has at most one label. We modify the BERT-NER implementation to allow for multi-label predictions using the following method.

Given a sentence \( S = (x_1, x_2, ..., x_L) \), the original BERT-based token classification model generates \( L \) respective \( M \)-dimensional encodings \((x_1, x_2, ..., x_L)\), one for each token. These encodings then pass through a softmax layer and make use of a multi-classes cross entropy loss function that generates a single class prediction per token. We adapt this system to multi-label classification by using a sigmoid function and binary cross entropy in place of the original loss function to allow the classifications for each token to return non-zero values for more than one class. More formally, our implementation performs the following steps for each token \( x_i \):

\[
\mathcal{L}(x_i) = -\sum_{c 

[^1]: [https://github.com/kamalkraj/BERT-NER](https://github.com/kamalkraj/BERT-NER)
loss function becomes:

\[
L_{\text{multilabel}} = -\frac{1}{M} \sum_{m=1}^{M} \left[ \tilde{y}_m \cdot \log\sigma(x_l^m) + (1 - \tilde{y}_m) \cdot \log(1 - \sigma(x_l^m)) \right]
\]

where \( M \) is the number of total classes, \( x_l \) is \( M \)-dimensional encoding for the \( l \)-th token in sentence, \( \tilde{y}_m \) is the \( l \)-th token’s gold label vector, and \( \sigma \) is the sigmoid activation function.

Folds: Because of the expense associated with annotating a large corpus, only the training and development subsets of the ARC corpus were manually annotated with semantic class labels. As such we repurpose the original development set for testing, and hold out 20% of the training corpus for development. Summary statistics on these folds are provided in Table 5.

Hyperparameters: We make use of the pre-trained English BERT-Base-cased model \( \text{[1]} \) with a maximum sequence length of 64. The threshold for the sigmoid activation layer was tuned on the development set, with a value of 0.4 found to provide good performance. The large number of possible class labels in our task compared with typical named entity recognition datasets, combined with the modified multi-label loss function, necessitated significantly longer training times for the model to converge. We empirically found that the model tended to converge by 140 epochs, which took approximately 5 hours to train using dual RTX2080Ti GPUs. Classification of the entire test dataset is comparatively fast, providing semantic class labels at a rate of approximately 900 questions (35,000 words) per minute, enabling the pre-trained model to be run on other science-domain corpora (for example, textbooks, study guides, Simple Wikipedia, or other grade-appropriate knowledge resources) at scale.

4.2. Evaluation

The results for our semantic classification task on the ScienceExamCER corpus using the 601-class fine-grained typology are shown in Table 4. We evaluate entity classification performance using the standard definitions of Precision, Recall, and F1. Overall classification performance is high, reaching 0.85 F1 on the held-out test set. This suggests the common entity recognition performance is sufficiently high to be useful for a variety of downstream tasks. To further characterize performance, we investigate how the availability of training data affects this fine-grained classification task, as well as common classes of prediction errors the BERT-NER model makes.

4.3. Performance vs Training Data

Manually annotating fine-grained mentions in large corpora is expensive and time consuming. To investigate how classification performance varies with availability of training data, we randomly subsampled smaller training sets from our full training corpus that were 25%, 50%, or 75% as large, corresponding to spending approximately 50, 100, or 150 hours at the manual annotation task, respectively. The results are shown in Figure 2. With only 25% of training data available, F1 performance decreases dramatically from 0.85 to 0.50. 50% of training data decreases classification performance by 7 points, while 75% of available training data decreases classification performance by 2 points. This suggests that the scale of training data generated in this work is provides near saturated performance using the BERT-NER model, and that annotating the remainder of available standardized science exam questions in the ARC corpus would likely result in only a minimal increase on classification performance.

4.4. Error Analysis

To better understand the sources of error in our model, we conducted an analysis of the first 100 errorful predictions on the test set, with the results shown in Table 6. Nearly one third of errors are due to issues with the annotation, such as a mention missing an additional label that is also good (24% of errors), or the manually annotated gold label being incorrect (7% of errors). For a substantial portion of errors (24%), no single semantic class rose to meet the activation threshold of the sigmoid layer and the model did not produce a prediction for that word, while, similarly, in 21% of cases only one label of a multi-label word was produced. The remaining errors broadly cluster around technical challenges in determining the semantics of each cate-

Table 4: Performance on the 601-category fine-grained semantic classification task on the development and test folds using the BERT-NER model.

| Model   | Fold  | Prec. | Recall | F1   |
|---------|-------|-------|--------|------|
| BERT-NER | dev   | 0.84  | 0.85   | 0.84 |
| BERT-NER | test  | 0.85  | 0.86   | 0.85 |

Figure 2: Classification performance (F1) versus the number of training epochs when training the model with less data. Series represent training the model with the entire training set, or randomly subsampled proportions of training data summing to 75%, 50%, and 25% of the original training set size. Each point represents the average of 5 randomly subsampled training sets.
Table 5: An analysis of common categories of model prediction errors, as a proportion of the first 100 errors on the test set. Note that a given errorful prediction may belong to more than one category, and as such the proportions do not sum to 100%.

| Error Class | Prop. |
|-------------|------|
| Predicted label also good | 24% |
| Model did not generate prediction | 24% |
| Multiple gold labels, one found | 21% |
| Predicted label semantically near gold label | 17% |
| Gold label incorrect | 7% |
| Multi-word Expression | 6% |
| Predicted label using incorrect word sense | 5% |

5. Conclusion

We present ScienceExamCER, a densely annotated corpus of science exam questions for common entity recognition where nearly every word is annotated with fine-grained semantic classification labels drawn from a manually-constructed typology of 601 semantic classes. We demonstrate that BERT-NER, an off-the-shelf named entity recognition model, achieves 0.85 F1 on classifying these fine-grained semantic classes on unseen text in a multi-label setting. The data and code are released with the goal of supporting downstream tasks in question answering that are able to make use of this dense semantic category annotation.

6. Supplementary Material

The annotated corpora, fine-grained typology, and pre-trained models for this work are available at http://cognitiveai.org/explanationbank/ A truncated version of the typology is included in the Appendix below.

7. Acknowledgements

This work was supported by the Allen Institute for Artificial Intelligence and the National Science Foundation (NSF Award #1815948, “Explainable Natural Language Inference”, to PJ). We thank Peter Clark for thoughtful discussions on this work and comments on an earlier draft.

8. Bibliographical References

Abujabal, A., Yahya, M., Riedewald, M., and Weikum, G. (2017). Automated template generation for question answering over knowledge graphs. In Proceedings of the 26th international conference on world wide web, pages 1191–1200. International World Wide Web Conferences Steering Committee.

Bodenreider, O. (2004). The unified medical language system (umls): integrating biomedical terminology. Nucleic acids research, 32(suppl_1):D267–D270.

Bollacker, K., Evans, C., Paritosh, P., Sturge, T., and Taylor, J. (2008). Freebase: a collaboratively created graph database for structuring human knowledge. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data, pages 1247–1250. AcM.

Choi, E., Levy, O., Choi, Y., and Zettlemoyer, L. (2018). Ultra-fine entity typing. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 87–96, Melbourne, Australia, July. Association for Computational Linguistics.

Clark, P., Cowhey, I., Etzioni, O., Khot, T., Sabharwal, A., Schoenick, C., and Tafjord, O. (2018). Think you have solved question answering? try arc, the ai2 reasoning challenge. arXiv preprint arXiv:1803.05457.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1):37–46.

Del Corro, L., Abujabal, A., Gemulla, R., and Weikum, G. (2015). Finet: Context-aware fine-grained named entity typing. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 868–878.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Dredze, M., McNamee, P., Rao, D., Gerber, A., and Finin, T. (2010). Entity disambiguation for knowledge base population. In Proceedings of the 23rd International Conference on Computational Linguistics, pages 277–285. Association for Computational Linguistics.

Fellbaum, C. (1998). WordNet. Wiley Online Library.

Greenberg, N., Bansal, T., Verga, P., and McCallum, A. (2018). Marginal likelihood training of bilstm-crf for biomedical named entity recognition from disjoint label sets. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pages 2824–2829.

Grishman, R. and Sundheim, B. (1996). Message understanding conference-6: A brief history. In COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics.

Hainisch, D., Fundel, K., Mevissen, H.-T., Zimmer, R., and Fluck, J. (2005). Prominer: rule-based protein and gene entity recognition. BMC bioinformatics, 6(1):S14.

Ling, X. and Weld, D. S. (2012). Fine-grained entity recognition. In Twenty-Sixth AAAI Conference on Artificial Intelligence.

Ma, Y., Cambria, E., and Gao, S. (2016). Label embedding for zero-shot fine-grained named entity typing. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 171–180.

Magnini, B., Negri, M., Preverte, R., and Tanev, H. (2002). A WordNet-based approach to named entities recognition. In COLING-02: SEMANET: Building and Using Semantic Networks.

Mai, K., Pham, T.-H., Nguyen, M. T., Nguyen, T. D., Bol-
legal, D., Sasano, R., and Sekine, S. (2018). An empirical study on fine-grained named entity recognition. In Proceedings of the 27th International Conference on Computational Linguistics, pages 711–722.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D. (2014). The stanford corenlp natural language processing toolkit. In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pages 55–60.

Mayfield, J., McNamee, P., and Piatko, C. (2003). Named entity recognition using hundreds of thousands of features. In Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003 - Volume 4, CONLL ’03, pages 184–187, Stroudsburg, PA, USA. Association for Computational Linguistics.

Murty, S., Verga, P., Vilnis, L., Radovanovic, I., and McCallum, A. (2018). Hierarchical losses and new resources for fine-grained entity typing and linking. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 97–109.

Nadeau, D. and Sekine, S. (2007). A survey of named entity recognition and classification. Lingvisticae Investigationes, 30(1):3–26.

Pennington, J., Socher, R., and Manning, C. (2014). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pages 1532–1543.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., and Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365.

Qu, L., Ferraro, G., Zhou, L., Hou, W., and Baldwin, T. (2016). Named entity recognition for novel types by transfer learning. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 899–905, Austin, Texas, November. Association for Computational Linguistics.

Rau, L. F. (1991). Extracting company names from text. In [1991] Proceedings. The Seventh IEEE Conference on Artificial Intelligence Application, volume 1, pages 29–32. IEEE.

Ritter, A., Clark, S., Etzioni, O., et al. (2011). Named entity recognition in tweets: an experimental study. In Proceedings of the conference on empirical methods in natural language processing, pages 1524–1534. Association for Computational Linguistics.

Sang, E. F. and De Meulder, F. (2003). Introduction to the conll-2003 shared task: Language-independent named entity recognition. arXiv preprint cs/0306050.

Sekine, S., Sudo, K., and Nobata, C. (2002). Extended named entity hierarchy. In LREC.

Sekine, S. (2008). Extended named entity ontology with attribute information. In Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC’08), Marrakech, Morocco, May. European Language Resources Association (ELRA).
| **Label**                  | **Examples**                                                                 |
|---------------------------|-------------------------------------------------------------------------------|
| Celestial Object          | celestial object, astronomical body, celestial body, extrasolar-body         |
| Asteroid                  | asteroid belt, Kuiper belt, asteroid, planetoid, Iris, Flora, Metis, Hygiea  |
| Black Hole                | super-massive black hole, black hole, Centaurus A, Sagittarius A*             |
| Comet                     | Halley’s comet, comet, Shoemaker-Levy, Great Comet of 1807                    |
| Constellation             | constellation, Leo, Little Dipper, pattern of stars, star pattern, ursa Major |
| Galaxy                    | Andromeda galaxy, Milky Way, M87, galaxy, Large Magellanic Cloud             |
| Galaxy Parts              | galactic region, halo, spiral arm, spiral arms, nuclear bulge                 |
| Light                     | light, ray, beam of light, ray of light, glow, radiance, corona, flash        |
| Celestial Light on Earth  | daylight, sunlight, starlight, moonlight, moonshine, sunshine, twilight       |
| Meteor                    | meteor, meteoroid, meteorite, Perseids, Lyrids, Quadrantids, Geminids         |
| Moon                      | moon, lunar, Deimos, Phobos, Europa, Ganymede, Rhea, Charon                   |
| Lunar Phases              | new moon, waxing crescent, first quarter, waxing gibbous, full moon, last quarter |
| Nebula                    | nebula, Cat’s Eye Nebula, Horseshoe Nebula, Orion Nebula                      |
| Particles                 | ice particle, dust, particle, particulate, cosmic dust, space dust, stardust   |
| Planet                    | planet, planets, rouge planet, planetary                                     |
| Dwarf Planets             | Pluto, Ceres, Haumea, Makemake, Eris                                         |
| Inner Planets             | Earth, Venus, Mars, Mercury, terrestrial planets, inner planet, inner         |
| Outer Planets             | Saturn, Jupiter, Neptune, Uranus, gas giants, outer planet, outer             |
| Planet Parts              | core, crust, mantle, ring, surface, axis, magnetic pole, atmosphere, magnetosphere |
| Satellite                 | satellite, Sputnik, communications satellite, GOES 15, Oceansat-1, Astrosat  |
| Solar System              | solar system, multiplanetary system, planetary system, planetary systems      |
| Space Probes              | Mars Rover, exploratory robot, Viking I Lander, space probe, Luna 9, Voyager 1 |
| Spacecraft (Human Rated)  | lunar module, spacecraf, Soyz, International Space Station, Apollo, Space Shuttle |
| Spacecraft Subsystem      | guidance, propulsion, support, suspension, structure, attitude control        |
| Star                      | star, Sun, Proxima Centauri, Polaris, Vega, VY Canis Majoris, Wolf 359       |
| Star Types                | giant, dwarf, main-sequence, supergiant, protostar, supernova, neutron, binary|
| Star Layers               | core, radiative zone, convection zone, chromosphere, photosphere, corona     |
| The Universe and Its Parts| universe, space, outer space, cosmos, observable universe, supercluster      |
| Vacuum                    | vacuum, in vacuo, vacuity                                                    |
| Celestial Events          | celestial event, solar flare, shooting star, meteor shower, transits, planetary |
| Eclipse Events            | solar eclipse, lunar eclipse, annular eclipse, partial eclipse, partial lunar eclipse |
| Force                     | weak force, strong force, magnetic force, centripetal force, friction, centrifuge |
| Gravity                   | gravitational pull, gravitational acceleration, gravitational force, gravitation |
| Inertia                   | inertia                                                                       |
| Magnetic Force            | magnetism, magnetic force, magnetic pull, magnetic field, electromagnetic force |
| Pressure                  | atmospheric pressure, vapor pressure, air pressure, water pressure, barometric |
| “Pulling” Forces          | air resistance, friction, traction, frictional force, sound barrier, drag, torsion |
| “Pulling” Actions         | pull, slow down, stop, attract, pulling, draw in, wrench, twist, twisted, pull back |
| “Pushing” Forces          | thrust, lift, compression, compressive force, compressive forces, normal force |
| “Pushing” Actions         | push, throw, toss, fall, sink, accelerate, motion, repel, compress, swing, exert |
| Energy                    | energy, light energy, radiation, kinetic energy, thermal energy, mechanical energy |
| Absorb Energy             | absorb, energy consumption, endothermic reaction, reabsorb, consume, uptake   |
| Electrical Energy         | electromagnetic energy, electrical charge, shock, electricity, electric current |
| Energy Waves              | radio wave, wave, ripple, light wave, seismic wave, electromagnetic wave, sound |
| Parts of Waves            | crest, trough, peak, amplitude, wavelength, frequency                          |
| Wave Perception           | Doppler effect, interference, wave-particle duality, sound perception         |
| Magnetic Energy           | electromagnetic energy, magnetic field, magnetic moment, ferromagnetism, dipole |
| Produce Energy            | fission, fusion, nuclear reaction, energy production, energy generation, create |
| Release Energy            | burn, glow, transmit, heat, surface cooling, distribute, exothermic reaction  |
| Sound Energy              | sound, sound energy, noise, vibration, vibrations, pascal, decible, echo, echoes |
| Examples of Sounds        | cluck, clucking, meow, meowing, humming, buzz, buzzing, shout, shouted, note, tune |
| Thermal Energy            | conduction, convection, radiation, heat, solar radiation, thermal, latent heat |
| Transfer Energy           | heat transfer, conduct, energy transfer, convection, convert, change into      |
| Spectra                   | spectrum, electromagnetic spectrum, continuum, frequency spectrum             |
| Electromagnetic Spectrum  | visible light, radio wave, radio waves, microwave, x-ray, infrared             |
| Living-thing              | organism, creature, extra-terrestrial life, bacteria, living, biological, plankton |
| Animal                    | animal, worm, predator, sponge, Animalia, heterotroph, dinosaur, snail, creature |
| Aquatic                   | fish, sea star, anemone, shellfish, anglerfish, otter, walrus, stout beardfish |
Foods: pizza, pepperoni, cheese, banana, corn, soybeans, wheat, tomatoes, apples
Plant Nutrients: humus, fertilizer, soil, carbon, hydrogen, oxygen, nitrogen, phosphorus
Actions for Nutrition: rehydration, dehydration, storage, overcook, pasteurization, cooking, absorb
Agriculture: crops, agriculture, agricultural, farming, food crops, farm land, livestock
Actions for Agriculture: growing, rotate, farming, harvest, irrigating, fertilization, grazing, raise

Measurements:
- Celestial Measurements: parallax, redshift, absolute magnitude, apparent magnitude, red shift
- Geometric Measurements: angle, curvature, circumference, compactness, dimension, position, reach
- Measures of Amount of Light: specific heat, heat capacity, thermal capacity, specific heat capacity
- Measuring Speed: slow, rate, speed, speed of light, constant, changing, fast, steady, increasing
- Unit: metric system
- Astronomy Units: light year, A.U., AU, ly, Parsec, pc, light second
- pH
- Hectare, square meters, m², square inches, in², square kilometers
- Atomic mass unit, gram, g, kg, lb, ton, pounds, kilogram, tonne
- Milliliter, milliliters, L, dl, deciliter, gallons, liter, fluid ounces
- Degrees Celsius, °C, °F
- Joules, calorie, watt, erg, erg/s
- Hz, megahertz, gigahertz, terahertz, hertz, kHz, MHz, GHz
- Percentage, percent
- Mohs hardness, Vickers hardness number, Rockwell hardness, Shore hardness

Manmade Objects:
- Golf ball, hammer, boots, solid, space suit, power lines, plate, balloon, box
- Lawn mower, household appliance, solar panel, hand dryer, fan, drill
- Gas grill, electric fry pan, microwave oven, solar cooker, electric stove
- Freezer, refrigerator, cold pack, air conditioner, fan, ice box, air cooling
- Electric toothbrush, plug, sewing machine, telephone, electric stove
- Pot, pan, graduated cylinder, jar, glass, container, reservoir, bucket
- Electrical circuit, circuit, parallel circuit, electric circuit, series circuit
- Cutting tool, heat engine, machine, device, radar, sonar, engine, outlet
- Prism, Triangular prism, Abbe prism, Pellin-Broca prism
- Filter, paper filter, coffee filter, sifter, surface filter, sieve
Computing Device calculator, computer, laptop, personal computer, driver
Light-producing Object light bulb, flashlight, incandescent light bulb, laser, penlight, lamp
Sound-producing Object tuba, bassoon, violin, guitar, drum, piano, flute, harp, recorder
Simple Machines simple machine, wheel and axle, lever, inclined plane, screw, pulley, wedge
System of Communication newspaper, Internet, telephone, radio, television, TV, walkie talkie
Technological Instrument calculator, computer, tripod, test tube, camera, recorder, radio, robot
Technological Component wire, power lines, attachment, filament, button, encoder, decoder, receiver
Chemical Product cleaners, laundry detergent, dish soap, adhesives, sealants, polymers
Vehicle mechanical system, vehicle, craft
Air Vehicle jet plane, plane, helicopter, airplane, glider, airship, blimp, hot air balloon
Land Vehicle car, automobile, bumper car, bus, bicycle, train, skateboard, motorcycle
Snow Vehicle snowmobile, motor sled, motor sledge, snow scooter, snow machine
Space Vehicle rocket, capsule, lunar lander, space shuttle, spacecraft
Water Vehicle boat, submarines, oceanliners, canoe, cable ferry, cable, cog, cutter, dugout
Traffic traffic, congestion, air traffic, pedestrian traffic, foot traffic
Clothes/Textiles biodegradable carpeting, clothes, shirt, skirt, pants, shoes, sock
Man-made Geographic Formations oil wells, wells, dams, aeration pond, canal, port, harbor, wharf

Property property, properties, characteristic, characteristics, nuclear property, trait
Age old, young, new, ancient, mature, prehistoric, maturity, old-growth, aged
Chemical Property chemical property, salinity, corrosive, nitrate levels, concentration
Ph (Acidity) ph, acid, base, acidic, basic
Flammability flammability, flammable, inflammable, combustible, combustible
Language Latin, English, Spanish, Greek, Hawaiian, Italian, Chinese, Mandarin, Japanese
Nationality/Origin American, North American, Hawaiian, foreign, Scottish, Chinese, European
Ability ability, skill, aptitude, capability, capability, potentiality
Other Organism Properties alive, multicellular, DNA, unicellular, autotrophic, dormant, fossilized
Behaviors behavior, conscious behavior, environmental behavior, conduct, comportment
Inherited Behavior instinct, inherited behavior, heredity, inherit, inherited
Learned Behavior routine, habit, learned behavior, acquired, modus operandi, habitual method
Other Animal Properties fertile, adaptable, endothermic, ectothermic, hairy, slimy, warm-blooded
Gender female, male, maleness, masculinity, androgyny, hermaphroditism, femaleness
Other Human Properties blood type, humor, honesty, leadership, handedness, kindness, wisdom, duty
Physical Property density, density, height, surface area, weight, physical property, conductivity
Conductivity conductivity, conducts heat, conducts electricity, conducts sound
Temperature temperature, cool, hot, warm, cold, room temperature, unevenly heated
Composition composition, chemical composition, metallic, rocky, icy, porous, concentration
Mass mass, heavy, light, biomass, lightweight, hefty, massive, ponderous, weighty
Distance distance, 100-meter
Shape shape, long, elliptical, spiral, irregular, oval, circular, convex, concave
Size big, small, large, size, diameter, radius, thin, thick, volume
Height low, tall, short, high, elevation, altitude, high-altitude, highest, height
Depth deep, shallow, deepest, depth, depthness, profundity, profundous, shallowness, shallowness
Width wide, narrow, thin, thick, thickness, width, breadth, wideness, breadth
Length long, short, length, longest, longer, shortest, longest
Wetness wet, dry, damp, driest, wetter, moist, bedewed, dewy, besprinkled, boggy, marshy
Texture smooth, rough, waxy, rocky, Slippery, porous, coarse, grainy, gritty
Material material
Synthetic Material plastic, glass, rubber, fiberglass, foam, Styrofoam, rayon, polyester, kevlar
Natural Material clay, soil, wood, paper, natural material, cardboard, ceramic, cotton, wool
Rigidity rigid, flexible, loose, brittle, rigidity, rigidity, inflexible, flexibility
Resistance/Strength water-resistant, resistant, heat-resistant, insulator, strong, weak, insulated
Hardness soft, brittle, hard, hardness, firmness, incompressibility, compressible
Permeability permeable, impermeable, semi-permeable, porous, pervious, impervious, leaky
Magnetic magnetic, nonmagnetic, ferromagnetic, magnetic field, magnetic flux, magnetize
Electrical Property electrical property, charge, electrical conductivity, electrical resistivity
Properties of Food fresh, shelf life, spoiled, rot, rotten, gone bad, unfermented, soured
Mineral Properties cleavage, fracture, hardness, luster, streak, structure, composition, color
Quality good, bad, useful, great, catastrophic, profound, adequate, best, crucial
Rarity typical, rare, common, commonly, abnormal, unusual, conventional, common enough
Speed fast, slow, quick, slowly, rapidly, rapid, immediate, gradual, faster, slower
Complexity simple, complex, directly, raw, complicated, composite, decomposable
| Visual Property          | reflective, shiny, appearance, dull, opaque, polished, symmetrical, milky |
|-------------------------|--------------------------------------------------------------------------|
| Color                   | orange, color, red, blue, white, yellow, grey, green, violet, black, sepia |
| Brightness              | brightness, luminosity, bright, glowing, lighted, sunny, dimmest          |
| Temporal Property       | long, short, length, variance, spontaneous, rapid, relatively short, long-term |
| Property of Motion      | speed, momentum, acceleration, velocity, rate, fast-flowing, movement, abrupt |
| Stability               | fixed, moveable, constant, stable, stability, steady, static, unchanging   |
| Position                | position, horizontal, parallel, perpendicular, sitting, standing, lying down |
| Properties of Waves     | wavelength, frequency, speed, amplitude                                   |
| Safety                  | safe, safer, safest, harmful, dangerous, reliable, danger, vulnerable     |
| Cost                    | expensive, inexpensively, affordable, efficiently, energy-saving           |
| Property of Production  | conventionally produced, organically produced, organic, coal-fired, manmade |
| Difficulty              | easily, easy, difficult, average, normal, hard, simple, trivial, arduous   |
| Other Properties        | layered, covered, distinctive, diversity, amniotic, crowded, divergent    |
| Numbers                 | number, amount, quantity, quantification, numerical, counting              |
| Cardinal Number         | one, two, three, four, 1, 12, 28, 7, 13, 130000, 2400, fifteen thousand    |
| Arithmetic Measure      | seven times, once, twice, 24 times, 365 times, ten millions, hundreds     |
| Relative Number         | several, abundance, fewer, lots, many, tankful, relative, too much, some, a few |
| Calculations            | x, times, divide, average, multiply, add, subtract, *, /, +, −              |
| Geography               | geographical, geography, human geography, physical geography, spatial analysis |
| Earth Parts (Gross)     | atmosphere, horizon, Northern Hemisphere, southern hemisphere, ocean, air  |
| Layers of the Earth     | crust, mantle, core, outer core, inner core, lithosphere, atmosphere      |
| Parts of Earth Layers   | tectonic plate, divergent boundaries, convergent boundaries, continental shelves |
| Tectonic Plates         | oceanic, continental, crustal, Pacific Plate, North American Plate        |
| Atmospheric Layers      | stratosphere, exosphere, thermosphere, mesosphere, troposphere, ozone layer |
| Fossils                 | fossil, fossils, remains                                                  |
| Archeological Process/Technique | dated, technique, radioactive dating, road cut, Law of Superposition     |
| Fossil Forming          | fossil-forming conditions, permineralization, authigenic mineralization   |
| Fossil Types            | index fossils, marine fossils, transitional fossils, microfossils, resin, amber |
| Cast Fossil/mold Fossil | coral fossil, coral fossils, endocast, concretions, mold fossil, cast fossil |
| Trace Fossil            | print, trace fossil, footprints, domicinia, fadinichnia, paschichnia       |
| True Form Fossil        | bone, bones, shell, shells, tooth, teeth, seashell, petrified wood, trilobite |
| Fossil Record/Timeline  | fossil record, geologic history, timeline                                 |
| Fossil Location         | Appalachian Mountains, Grand Canyon, Ohio, wooded area, desert, mountains  |
| Speciation              | speciation                                                               |
| Extinction              | extinct, mass extinction, mass extinctions                               |
| Geological Eons, Eras, Periods, Epochs | Mesozoic era, Cretaceous period, Precambrian, Paleozoic, Cenozoic, ice age |
| Natural Resources       | resource, supply, natural resource, natural resources, biotic resource    |
| Fossil Fuel             | fossil fuel, oil, coal, petroleum, gas, natural gas, gasoline, crude oil   |
| Other Energy Resources  | solar, wind, water, solar energy, flowing water, sunlight, wind power     |
| Changes to Resources    | restriction, conservation, loss, preservation, depletion, overconsumption  |
| Geographic Formations   | valley, mountain, volcano, highland, crater, sea, glacier, cliff, lake, fault |
| Geographic Formation Parts | peaks, slope, foot, caldera, crater, sill, conduit, cone, vent, ledge, hump |
| Bodies of Water         | pond, lake, puddle, ocean, spring, springs, groundwater, river, tributary |
| Specific Named Bodies of Water | Pacific Ocean, Atlantic Ocean, Mississippi River, Arctic ocean          |
| Types of Water in Bodies of Water | freshwater, saltwater, groundwater, brackish                      |
| Parts of Bodies of Water | riverbeds, basin, mouth, floor, wave, waves, delta, deltas, shoreline    |
| Currents                | current, ocean current, wind current, wind currents, Great Ocean Conveyor |
| Tides                   | high tide, low tide, tides, tidal, intertidal, highest astronomical tide  |
| Actions for Tides       | rise, fall, tidal action, come in, go out                               |
| Geographic Formation Process | geologic process, geomorphology, petrification, permineralization        |
| Change in Location      | collide, collision, distance, distancing, impact, shift, shifting, strike |
| Change in Composition   | chemical reactions, chemical reaction, burn, erupt, eruption, explode    |
| Constructive/Destructive Forces | deposit, deposition, erode, erosion, weather, weathering, compress       |
| Minerals                | gold, silver, mineral, crystals, copper, phosphorous, abelsonite, amberthine |
| Mineral Formations      | vein deposits, vein deposits, mineral deposits, mineral deposite, stalactite |
| Rock                    | rock, pebbles, gravel, lava, boulder, boulders, slab, gravel deposite      |
| Igneous                 | igneous, granite, igneous intrusion, basalt, volcanic, intrusive igneous  |
| Metamorphic             | marble, foliated, metamorphic, gnesis, anthracite, granulite, greenschist  |
| Sedimentary             | sedimentary, sediment, limestone, sandstone, shale, marine sediments      |
| Soil                    | soil, sand, topsoil, mud, clay, ground, soil covering, earth, dirt        |
| Properties of Soil      | porous, fertility, nutrients, texture, structure, porosity, chemical makeup |
| Natural Phenomena                | erode, flood, erupt, weathering, natural event, earthquake, glacial activity |
|---------------------------------|-----------------------------------------------------------------------------|
| Weather Phenomena               | storm, wind, high tide, tide, trade winds, cloud, greenhouse effect, weather |
| Weather Descriptions            | clear, cloudy, humid, stormy, sunny, snowy, rainy, freezing, balmy, nice      |
| Precipitation                   | snow, rain, precipitation, rainfall, snowfall, acid rain, sleet, fog, hail    |
| Seasons                         | season, winter, summer, fall, spring                                        |
| Environmental Phenomena         | environmental pressure, environmental changes, habitat change               |
| Ecosystems/Environment          | ecosystem, environment, climate, world, biosphere, biome, environmental      |
| Nonliving Parts of the Environment | abiotic element, abiotic factor, nonliving thing, inanimate objects        |
| Habitat                         | habitat, shelter, territory, surroundings, landscape, home ground, habitation |
| Examples of Habitats            | hive, hollow tree, dam, stream, nest, burrow, river bottom, forest soil, soil|
| Types of Terrestrial Ecosystems | desert, temperate, tropical, savanna, arctic, plain, tundra, grassland, prairie |
| Forests                         | rainforest, coniferous forest, deciduous forest, Alpine forest, wooded area  |
| Sky                             | sky, night sky, ozone layer, greenhouse gas, air mass, blue sky, aerospace   |
| Environmental Damage/Destruction| pollution, air pollution, chemical spills, logging, deforestation           |
| Underwater Ecosystem            | whale fall, black smoker, estuary, intertidal, reef, marine ecosystems       |
| Other Geographic Words          | volcanic, global, geological, oceanic, geologically, layers, buildup         |

| Matter                          | matter, nonliving matter, agent, material, dark matter, antimatter, ylem, thing |
|---------------------------------|--------------------------------------------------------------------------------|
| Compound                        | carbon dioxide, chemical composition, ammonia, methane, greenhouse gas          |
| Organic Compounds               | organic, organic compound, hexane, ozone, formaldehyde, acetic acid, alcohol   |
| Elemental Components            | atom, molecules, polar molecule, ion, formula unit, biomolecules              |
| Atom Components                 | proton, electron, nucleus, neutron, subatomic particles, particles             |
| Atomic Properties               | atomic mass, atomic radius, electrical charge, electric potential              |
| Molecular Properties            | covalent bond, cohesion, net charge, chemical bond, molecular speed, polarity  |
| Chemical Processes              | nitrification, denitrification, saturation, fixation, hydration, dehydration    |
| Element                         | element, radioactive isotope, isotope, fluorine, ammonium, hydrogen, helium    |
| Classes of Elements             | alkali metal, nonmetal, metalloid, noble gas, halogen, alkaline earth metals   |
| Mixtures                        | solution, mixture, suspension, colloid, alloy, blend, mix, azo trope, air       |
| Parts of a Solution             | solute                                                                         |
| Separating Mixtures             | chromatography, distillation, evaporation                                     |
| Phases of Water                 | water, frost, ice, steam, vapor, liquid water, ice crystals                     |
| State of Matter                 | solid, gas, liquid, plasma, state of matter, physical state                    |
| Solid Matter                    | ice, sulfur, flower, cloth, glass, wood paper, peanuts, match, top             |
| Granular Solids                 | sugar, sand, table salt, salt, pepper, baking soda, powder, dust, pepper        |
| Metal Solids                    | metal, nail, hammer, gold bar, magnesium, copper, car fender, wire, gold, iron |
| Liquid Matter                   | water, acid, carbonated water, milk, oil, vinegar, lemon juice                 |
| Capillary Action                | capillary action, capillarity, capillary motion, capillary effect              |
| Gaseous Matter                  | oxygen, air, nonreactive gas, gases, bubble, vapor, greenhouse gas emissions   |
| Phase Transition Point          | boiling point, freezing point, transition point, evaporation point             |
| Substances                      | substance, silver, magnesium, sulfur, aluminum, compounds, pure substance      |

| Changes                         | electrical, thermal, change, conversion, transform, chemical reaction         |
| Chemical Change                 | new/different substance be formed, chemical change, rust, light a candle, burn |
| Physical Change                 | physical change, change volume, change the shape, temperature change, diffusion |
| Phase Changes                   | phase change, change the state of matter, change to, change in the state of   |
| Phase-changing Actions          | melt, freeze, boil, evaporate, become steam, vaporize, condense, sublimate    |
| Reactions                       | chemical reaction, nuclear reaction, thermonuclear reaction                   |
| Parts of Chemical Reactions     | reactant, product, catalyst, inhibitor, positive feedback, negative feedback  |
| Types of Chemical Reactions     | endothermic, exothermic, combination, decomposition, single displacement       |

| Actions                         | act on, apply, apply to, interact, interaction, transfer, operate, attract    |
| Act Upon Something              | act on, apply, apply to, interact, interaction, transfer, operate, attract    |
| Alter                           | fix, affect, adapt, impact, shape, alter, modify, modified, regulate, recycle |
| Form-changing Actions           | tie, cut, crush, break, shred, dissolve, saw, filter, mix, slice, spread, roll |
| Color-changing Actions          | color, paint, polish, change color, stained, dyed, tinge, discoloring, colorize |
| Location-changing Actions       | drop, blow, spin, float, sink, bury, burying, dump, dumping, pump, pumped    |
| Amount-changing Actions         | increase, decrease, reduce, add, take, put, lloss, extend, release, lost      |
| Avoid/Reject                    | avoided, avoid, disregard, unattended, prevent, ignore, ignoring, evade, evaded |
| Believe                         | thought, believe, believed, conceived of, suspect, suspects, consider, hold    |
| Buy                             | buy, purchase, buy back, bought back, buys up, repurchase, repurchases, owns   |
| Change Into                     | change, converted, convert, become, replace, replacing, self-assembly, into    |
| Classify                        | classify, label, call, called, categorize, classification, reclassify, identify |
| Choose                  | decide, decision, opt, option, choose, choice, prefer, vote, determine |
|------------------------|------------------------------------------------------------------------|
| Clean Up               | clean, wipe up, cleaning, washing, wash, flush, wipe, dispose of, throw away |
| Collect                | gather, gathering, congregate, collect, accumulate, amass, massed, compile |
| Combine/Add            | stir, stir into, mix, place in, shake, add, cover with, pour into, assemble |
| Communicate            | communicate, imitate, mimic, mimicry, signal, discuss, discussing, message |
| Compete                | compete, competition, vie, content, try for, race, rival, go for, challenge |
| Contain/Be Composed of | contain, composed, together, consist, make up, accumulate, cover |
| Create                 | create, form, generate, produce, emit, replicate, formation, cause, make, grow |
| Differentiate          | distinguish, specialize, differentiate, differ, differred, differentiation |
| Examine                | compare, comparing, monitor, track, tracking, analyze, analyzing, study |
| Harm                   | destroy, damage, deplete, depletion, malfunction, contaminate, collapse, harm |
| Help                   | help, contribute, support, defend, benefit, helped, aid, beneficial, heal |
| Identify               | detect, find, notice, found, discover, discovered, identify, identification |
| Increase/Decrease      | decrease, increase, contract, expand, inflate, decline, accelerate, lower, thin |
| Indicate               | mark, marked, indicate, indicates, list, listed, designate, show, give evidence |
| Move                   | pass through, carry through, deposit, travel, redistribute, move, migrate |
| Gaseous Movement       | swirl, swirling, blowing, blow, rise, rose, airflow, sink, sinks, flight, float |
| Liquid Movement        | flow, flows, flowing, drain, drained, flood, flooding, overflow, seep, seeps |
| Mechanical Movement    | pull, push, pedal, roll, drop, locomotion, drive, drag, shove, cycle, cycling |
| Particle Movement      | rise, fall, condense, expand, move faster, move slower, move upward, collide |
| Transportation         | transport, deliver, ship, delivered, transporting, drive, driving, glide |
| Celestial Movement     | revolve, rotate, orbit, tilt, move, turn, revolution, rotation, movement, spin |
| Apparent Celestial Movement | rise, set, apparent motion, disappear, become, appear, ascend, ascension |
| Light Movement         | shine, refract, reflect, travel, block, transmit, strike, glow, produce, emit |
| Observe                | appear, watch, observe, observation, seem, be seen, monitor, monitoring, view |
| Occur                  | happen, occur, experience, coincide, exist, present, undergo, take place |
| Permit                 | allow, letting, permit, accept, accepts, consent, consented, give permission |
| Perform an Activity    | operate, dissect, express, expresses, repeat, coordinate, irradiate, perform |
| Preserve               | sustain, storage, store, protect, recycle, preserve, continue, keep, conserve |
| Represent              | represent, describe, represented, stand for, correspond, typify, symbolize |
| Require                | require, requires, required, need, needs, needing, rely on, depend upon |
| Separate               | separate, break down, decompose, settle, sort, release, separation, escape |
| Release                | discharge, discharging, release, emit, re-emit, spill, loosen, free, expel |
| Break                  | break, shatter, broken, crack, crumble, fracture, fall apart, come apart, burst |
| Divide                 | divide, differentiate, lose, loss, divide into, disperse, dissolve, split |
| Start                  | originate, begin, start, set out, commence, commenced, lead off, led off |
| Stop/Remove            | block, prevent, eliminate, withstand, kill, stop, extinguish, dispose |
| Succeed                | achieve, accomplish, flourish, complete, finish, succeed, succeeded, win, won |
| Surpass                | outstrip, surpass, pass by, bypass, outmaneuver, overrule, exceed, outmatch |
| Touch                  | make contact with, connect, rub, tap, touch, stick together, reach, reached |
| Uptake                 | trap, take up, hold, pick up, obtain, colonization, process, capture, take |
| Use                    | use, deplete, using, burn, consume, consumption, used, utilized, overuse |
| Associate              | associate, associated, match, link, linked, relate, related, lead to |
| Verify                 | make sure, ensure, verify, verified, verifying, validate, check, checking |
| Wait/Stay              | wait, remain, stay, hold off, attend, continue, continued, expect, expecting |

**Scientific Method**

- Hypothesizing: hypothesize, predict, thought, estimate, suggested, topic, question, expected
- Performing Research: observe, conduct an investigation, conduct an experiment, compare, study
- Analyzing Research: inferring, classifying, generalizing, determine, calculate, analyze, discover
- Concluding Research: conclusion, report, presentation, evidence, support, finding, share, shared
- Replicating Research: reproduce, repeat, redo, replicate, duplicate, reproducible, reduplicate, copy
- Question/Activity Type: question, activity, mission, report, research, work, fieldwork, project
- Response Type: statement, explanation, fact, suggestion, term, theory, law, sentence
- Experimentation: experimentation, experimental design, experiment, trials

**Groups**

- Control group

**Variables and Controls**

- Variable, independent variable, dependent variable, factor

**Validity**

- Valid, reliable, authoritative, validity, relevant, logical, legitimate

**Performing Experiments Well**

- Critically, critical, skeptical, cautiously, precaution, appraising, evaluative

**Words for Data**

- Data, information, metadata, raw data, data set, evidence, input, datum

**Scientific Meetings**

- Symposium, meeting, science fair, conventions, conference, seminar, colloquium

**Audiences**

- Audience, spectators, gallery, grandstand, house, gathering, assemblage

**Guidelines and Rules**

- Rule, laws, regulation, conventions, requirement, prescriptions, principle
Markers of Time
- period, time zone, time, timing, era, epoch, biological time, cosmic time
- night, day, evening, nighttime, daytime, sunrise, noon, sunset, mid-afternoon
Relative Time
- first, beginning, middle, end, never, during, throughout, between, past, span
Months
- January, February, March, April, May, June, July, August, September, October
Day
- Monday, Tuesday, Wednesday, Thursday, Friday, Saturday
Year
- year
Year Numerals
- 1971, 1953, 1990, 2020, 0, 45, 1266, 1496, 1692, 1777, 1787, 999, 1900, 1900s
Frequency
- daily, constantly, monthly, yearly, every night, perpetually, continuously

Locations
- location, land area, spot, place, region, position, setting, zone, point, site
Manmade Locations
- house, street, garden, building, town, factory, airport, radio station tower
Parts of a Building
- foundation, roof, floor, frame, walls, windows, window panes, beams, boards
Terrestrial Locations
- beach, Equator, shoreline, field, underground, sea-floor, polar snowcaps
Northern Hemisphere Locations
- Northern Hemisphere, Alaska, North Pole, New York State, Baltimore, Florida
Southern Hemisphere Locations
- Southern Hemisphere, Australia, South Pole, South America, Chile, South Africa
Relative Locations
- bottom, top, middle, between, surroundings, under, nearby, submerged, exposed
Directions
- direction, path, route, trail, itinerary, way, via, course, trackway
Cardinal Directions
- west, east, south
Relative Direction
- upward, downhill, right, left, direct, clockwise, counterclockwise, western
Prepositional Directions
- across, toward, around, through, away from, up, down, across from, along, among
Geopolitical Locations
- county, Yellowstone National Park, Mojave Desert, Chesapeake Bay, Knight Island
Continents
- Antarctica, North America, South America, Africa, Asia, Australia, Europe
Countries
- nations, industrialized nations, country, Afghanistan, Albania, Algeria
Cities
- Baltimore, Port Orange, city, Tucson, Yuma, Flagstaff, Winslow, Boston
States
- Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware
Verbs for Locate
- located, placed, extend, transport, locate, navigate, circumnavigate, find

Comparisons
- difference, in common, comparison, distinct, different, identical, similar
Visual Comparison
- look like, resemble, looking similar, bear resemblance to, take after
Quality Comparison
- better, best, good, poor, advantage, negative, benefit, improvements
Amount Comparison
- fewer, less, more, quantity, a small amount, ratio, most, level, all, maximum
Importance Comparison
- primary, primarily, main, dominant, of importance, of import, crucially
Distance Comparison
- further, closest, closer, close, equal distances, nearer, farther

Scientific Theory, Experimentation, and History
- science, scientific, scientific theories, scientific terms, topic area, history
Theory of Matter
- law of conservation of mass, law of conservation of energy
Representing Elements and Molecules
- orbitals, models, chemical formula, chemical equation
Astronomy/Aeronautics
- astronomy, astronomical theory, aerodynamics, Milankovitch cycles, Tusi couple
Space Agencies
- NASA, National Aeronautics and Space Administration, Air Force Space Command
Space Missions
- Kepler Mission, Apollo 14, manned space exploration, mission
Observation Places
- Maryland Space Grant Observatory, observatory, Royal Observatory
Observation Instruments
- Morris W. Offit telescope, telescope, Hubble space telescope
Astronomical Distance Units
- Astronomical Units, light year, A.U., AU, ly, Parsec, pc, light second
Cosmological Theories
- Big Bang Theory, theory, Heliocentric Theory, Earth-centered universe theory
Cosmological Theory Thematic Words
- contract, contracting, contracts, form, expand, expanding, expands, change
Theory of Physics
- first law of motion, second law of motion, third law of motion, physics
Occupation
- Professor, police, firefighter, teacher, doctor, nurse, baker, lumberjack
Scientists
- Thomas Edison, Galileo, observer, student, Darwin, geologist, Jonas Salk
Groups of Scientists
- students, scientists, observers, NASA, paleontologists, surveyors, researchers
Biology
- selective breeding, cell theory, endosymbiotic theory, molecular biology
Natural Selection
- natural selection, survival of the fittest, Darwin’s Theory of evolution
Observation Techniques
- color staining
Meteorology
- weather forecasts, air-quality control, Saffir-Simpson scale, Coriolis effect
Meteorological Models
- station model, atmospheric model, Mesoscale Model, NAM, Global Forecast System
Geologic Theories
- Law of Superposition, law of crosscutting relationships, continental drift
Conservation Laws
- The Water Quality Act of 1987, Clean Air Act, Clean Water Act
Discovery
- discovery, invention, scientific advancement, advances, scientific discovery
Undiscovered
- undiscovered, unknown, unidentified, undetected, unexplored, lost, hidden
| Generic Terms                          | terms, terminology, generic, items, words, definitions, language, referents  |
|----------------------------------------|-----------------------------------------------------------------------------|
| Ability/Availability                   | unable, usable, useable, potential, room, able, available, unavailable      |
| Relations                              | independent, together, homologous, relationship, imbalance, interaction      |
| System and Functions                   | system, machine, subsystem, activity, function, network, practice, programs |
| Feedback Mechanism                     | feedback mechanism, positive feedback, negative feedback, regulatory feedback |
| System Parts                           | source, power source, structure, structural, boundary, functional, unit      |
| System/Process Stages                 | step, sequence, stage, aspects, procedure, phase, degree, level, point      |
| Representation                         | image, diagram, chart, sign, model, prototype, drawing, list, instructions   |
| Parts of a Representation              | x-axis, y-axis, axes, labels, title, bar, line, point, coordinates, key, grid |
| Belief/Knowledge                       | dogma, understanding, knowledge, learning, logical, attitude, belief, religion |
| Classification                         | classification, taxonomy, categorization, compartmentalization, assortment   |
| Pattern                                | pattern, sequence, cycling, distribution, arrangement, trend, order, intervals |
| Gaps and Cracks                        | cracks, gap, crack, fractured, openings, grooves, pockets, diastema, hiatus  |
| Exemplar                               | kind, example, type, breed, medium, nature, version, variant, variation      |
| Emergency Services                    | 911, emergency services, police, fire department, EMS                      |
| Method                                 | method, way, fashion, strategy, technique, practice, plan, methodology      |
| State of Being                         | condition, state, scenario, presence, role, lifestyle, format, formats      |
| Event                                  | phenomenon, episode, occurrence, exhibit, event, practices, process, phenomena |
| Types of Event                         | race, party, test, class, explosion, club meeting, meeting, conference, match |
| Geometric/Spatial Objects              | plane, sphere, incline, object, body, ramp, equilateral triangle           |
| Object Part                            | end, center, core, surface, edge, rim, corner, middle, top, bottom, outside |
| Object Quantification                  | piece, sample, grain, component, whole, sheet, percent, group, chunk, layer |
| Negations                              | not, no, non, lacks, cannot, except, neither, nor, lack, never, nobody, none |
| Result                                 | by-product, buildup, following, effect, outcome, product, impact, reward     |
| Goal                                   | goal, objective, solution, end, finish, destination, aim, target, object    |
| Cause                                  | stimulus, internal stimulus, external stimulus, reason, factor, demand       |
| Source                                 | source, reserve, supply, origin, root, beginning, rootage, head              |
| Response                               | response, stress, reflex, symptom, reaction, answer, reply, aftereffect     |
| Relevant                               | appropriate, applicable, germane, pertinent, relevant                       |
| Group                                  | group, system, collection, cluster, list, combination, series, nature, council |
| Groups of Organisms                    | population, populations, community, residents, colony, the public, society  |
| Parts of a Group                       | member, individual, leader, teams                                           |
| Opportunities and Their Extent         | opportunities, advancement, limitation, opportunity, limits                 |
| Probability and Certainty              | approximately, exactly, about, correctly, likely, true, accurate, average    |
| Level of Inclusion                     | complete, some, few, all, every, each, both, certain, part, partial, incomplete |
| Problem                                | flaw, disorder, danger, negative effect, defect, accident, issue, impurities |
| Value                                  | value, worth, cost, price, profit, rate, expense, appraisal, assessment, charge |
| Separation                             | barrier, separation, wall, membrane, divider, blockade, roadblock, block     |
| Viewpoint                              | perspective, angle, attitude, mindset, viewpoint, headset, point of view     |
| Business/Industry                      | supplier, companies, company, businesses, enterprise, movie studio, industry |
| Business Names                         | further, closest, closer, close, equal distances, nearer, farther           |
| Scientific Associations/Administrations| American Dental Association, Food and Drug Administration, government agencies |
| Advertising                            | advertisement, commercial, advertise, ad, marketing, announcements, broadcast |
| Parts of a Business                    | distribution, mass marketing, public relations, research, quality control    |
| Products                               | merchandise, goods, brands, products, services, commodities, stock, effects  |
| Money Terms                            | funds, money, credited, funded, financial gain, shopping, lottery, income, fees |
| Patents                                | patented                                                                     |
| Employment                             | unemployment, employment, full-time, part-time, temporary, seasonal          |
| Media                                  | encyclopedia, world almanac, science textbooks, scientific journal, book     |
| Popular Media                          | news report, music, public speaking, CD, radio, radio show, website, blog    |
| Written Media                          | wall poster, brochure, article, magazine, book, newspapers, printed media     |

Table 6: The full list of semantic categories used in the ScienceExamCER common entity recognition corpus, as well as example words for each category.