UC Davis
UC Davis Previously Published Works

Title

Using Word Embeddings to Learn a Better Food Ontology

Permalink

https://escholarship.org/uc/item/4t27j6gb

Authors

Youn, Jason

Naravane, Tarini

Tagkopoulos, Ilias

Publication Date

2020

DOI

10.3389/frai.2020.584784

Copyright Information

Peer reviewed

Using Word Embeddings to Learn a Better Food Ontology

Jason Youn[1,2], Tarini Naravane[2,3] and Ilias Tagkopoulos[1,2]*

[1]Department of Computer Science, University of California at Davis, Davis, CA, United States, [2]Genome Center, University of California at Davis, Davis, CA, United States, [3]Biological Systems Engineering, University of California at Davis, Davis, CA, United States

Food ontologies require significant effort to create and maintain as they involve manual and time-consuming tasks, often with limited alignment to the underlying food science knowledge. We propose a semi-supervised framework for the automated ontology population from an existing ontology scaffold by using word embeddings. Having applied this on the domain of food and subsequent evaluation against an expert-curated ontology, FoodOn, we observe that the food word embeddings capture the latent relationships and characteristics of foods. The resulting ontology, which utilizes word embeddings trained from the Wikipedia corpus, has an improvement of 89.7% in precision when compared to the expert-curated ontology FoodOn (0.34 vs. 0.18, respectively, p value = 2.6×10^{-138}), and it has a 43.6% shorter path distance (hops) between predicted and actual food instances (2.91 vs. 5.16, respectively, p value = 4.7×10^{-84}) when compared to other methods. This work demonstrates how high-dimensional representations of food can be used to populate ontologies and paves the way for learning ontologies that integrate contextual information from a variety of sources and types.

INTRODUCTION

The need for efficient food systems to support food security (Tscharntke et al., 2012; Alexander et al., 2017), food production and distribution (Moe, 1998; Dabbene and Gay, 2011), and nutrition (Lemay et al., 2007; Kretsera et al., 2015; Berners-Lee et al., 2018; Barabási et al., 2019) to serve a growing planet is now more evident than ever (Guyomard et al., 2012). When it comes to food production and composition, various initiatives have proposed data repositories and ontologies regarding ingredients, processes, and final food products. Some examples of food compositional databases are USDA's FDC (US Department of Agriculture, Agricultural Research Service, 2019) which provides nutrient composition data for approximately 300,000 food entries and FooDB (Wishart, 2018) which provides quantitative chemical composition data in foods covering 80,000 chemicals in 800 foods. Other databases highlight non-ontological aspects, for instance, the GPC database (GS1, 2018) that contains barcodes for food products and the EFSA database (EFSA, 2015), which is a 32-feature categorization system. Concomitantly, there are multiple ontologies in various stages of development and usage (Dooley et al., 2018; Eftimov et al., 2019), with an ontology defined as the body of formally represented knowledge in some area of interest expressed by objects and concepts, and the relationships that hold among them (Genesereth and Nilsson, 2012). A notable example is FoodOn (Dooley et al., 2018), an open-source and formal food ontology curated by the FoodOn

FIGURE 1 | Overview of the LOVE ontology population framework. The hierarchical structure of the ontology is organized as a directed acyclic graph, where a class connects to its parent classes through directed edges. Target class is the parent class of the food instances. Note that some classes are part of the hierarchical ontological structure and do not contain any instances. All class and instance labels are used to query the Wikipedia corpus, which is then used to train food word embeddings. The mapping function then uses the word embeddings to map the candidate instances to the target classes. All relations between the instances and classes are of type "is a." We compare the predicted ontology to the ground truth ontology and report the performance using precision (more information in the "*Methods*" section).

consortium, which represents a food item by its properties and adheres to the FAIR standards (Wilkinson et al., 2016). As we move towards a detailed atlas of chemical food composition (Barabási et al., 2019), there is a current and present need for tools and frameworks that are data-driven and automated to support the creation and/or extension of evidence-based, detailed ontologies at scale.

The structure of an ontology is based on the triple of *subject*, *predicate*, and *object* which is similar to that of knowledge graphs (World Wide Web Consortium, 2011), yet there exist subtle distinctions. Ontologies are usually smaller in size, are domain-specific, capture complex relationships between the classes and instances, and can enforce their structure by applying sets of restrictions and rules (Benslimane et al., 2006; Ehrlinger and Wöß, 2016). Moreover, compared to the multi-relational knowledge graphs where different types of predicates can exist, ontologies connect concepts predominantly through subsumption or hypernymy relationships. Nonetheless, due to their structural similarities, several methods developed for the knowledge graph can also be applied to the area of ontology learning which includes tasks ranging from creating ontologies to extending and populating existing ontologies. However, in practice, the choice of embedding depends on the available corpus, and the method is specific to the task at hand. A task commonly seen in knowledge graphs is link prediction, where the starting state is a knowledge graph and the end result is a more accurate and/or more complete knowledge graph. Link prediction uses methods that explain the triples using the latent features such as Poincaré embeddings (Nickel and Kiela, 2017) or extract triples using contextual patterns from some text data. In the area of ontology learning, word embeddings created from text data are used to create and populate an ontology in an one-shot fashion using unsupervised methods such as clustering (Mahmoud et al., 2018) or to populate a skeleton knowledge graph initialized with seed instances in an iterative fashion (Jayawardana et al., 2017; Mitchell, 2018).

Here, we address the challenge of how to populate new instances into an existing ontological structure. We introduce LOVE (learning ontologies via embeddings), a semi-supervised framework for the automated ontology population (**Figure 1**), which uses word embeddings trained on a corpus obtained from Wikipedia. The required memory and computational time of the proposed method scale linearly with increasing number of instances. LOVE was applied on the FoodOn dataset to create the first food ontology using word embeddings. We evaluate the predicted ontology against FoodOn and achieve an increased precision of 89.7% when compared to the best alternate non-embedding-based method that uses Hamming distance (0.34 vs. 0.18, respectively, with a baseline precision of 4.7×10^{-4}).

METHODS

Data Preprocessing and Training of Word Embeddings

There are a total of 2,764 classes and 10,865 instances in FoodOn. Every class or food instance is identified by its label. For example, "cow milk cheese" is a class label, and "Brie cheese food product" is a food instance label. These labels are constructed using 4,139 unique words (e.g., "cow," "milk," "cheese," "Brie," "food," and "product"). We searched both the labels and their unique constituent words to obtain corresponding Wikipedia pages (**Figure 1**), which we refer to as Wikipedia corpus. We preprocessed the corpus as follows: lower-case conversion, synonym mapping, punctuation stripping, white-space stripping, numeric stripping, stop-words removal, short words stripping, and lemmatization. Note that the Wikipedia corpus consists of 142,948 unique words. For their training, we used the gensim (Rehurek and Sojka, 2010) implementation of the word2vec skip-gram model (Mikolov et al., 2013). Default settings of the gensim word2vec model were used except for the following parameters: number of epochs of 100, window size

TABLE 1 | Comparison of word similarity methods and their performance.

Method		Training corpus	Average precision	Average distance (hops)	Running time (s)
–	Random	n/a	4.6×10^{-4}	8.23	65.1
Non-embedding	Jaccard (Niwattanakul et al., 2013)	n/a	0.097	6.45	169.9
	Hamming (Hamming, 1950)		0.181	5.16	111.7
GloVe (Pennington et al., 2014)	50 d	Wikipedia 2014 + Gigaword 5	0.192	4.29	201.8
	100 d		0.228	3.87	
	200 d		0.261	3.53	
	300 d		0.297	3.32	
fastText (Joulin et al., 2016; Bojanowski et al., 2017)	Wikinews	Wikipedia 2017 + UMBC webbase + statmt.org	0.313	2.98	
	Crawl	Common Crawl	0.317	3.00	
Word2vec (Mikolov et al., 2013)	50 d	Subset of Wikipedia 2020	0.262	3.32	
	100 d		0.295	3.09	
	200 d		0.318	2.99	
	300 d		**0.344**	**2.91**	

Average distance denotes average path distance between the predicted class and the ground truth class among 100 multiple randomly selected seed instances. All methods were run parallel on 8 core CPU (16 threads) with 32 GB of memory. Note that the running time excludes the time used for training the word embeddings. Entries in bold correspond to the best performing method.

of 5, and minimum count of 1. We trained four different dimensions of word embeddings for word2vec: 50, 100, 200, and 300 d. In addition to word2vec, we also tested using the pre-trained word embeddings trained with GloVe (Pennington et al., 2014) and fastText (Joulin et al., 2016; Bojanowski et al., 2017). For GloVe, we downloaded pre-trained word embeddings of dimensions 50, 100, 200, and 300 d known as glove.6B. For fastText word embeddings, we used two different versions of word embeddings of size 300 d that have been trained using different training corpora. Refer to **Table 1** for complete information.

Ontology Population

As illustrated in **Figure 1**, our algorithm aims to map a food instance (e.g., "plum") through an "is a" relationship to its parent (e.g., "fruit," ideally), which we refer to as its target class. If we let i be a food instance and c be a target class, then $i \in I$ and $c \in C$, where I is the group of all food instances we seek to map and C is the group of target classes to which we map the food instance. We also define I_c to be all the food instances within a class c. To map the instance to its appropriate target class, we propose an approach based on the similarity of word embeddings. Our criteria for optimal population consider a linear combination of two scores:

$$\text{score}(c; i) = \alpha \cdot \text{score}_{\text{siblings}}(c; i) + (1 - \alpha) \cdot \text{score}_{\text{parent}}(c; i),$$

where α controls the ratio of the two terms. $\text{score}_{\text{siblings}}$ is the similarity of the food instance i with the seed instances in I_c:

$$\text{score}_{\text{siblings}}(c; i) = \text{sim}\left(\vec{\imath}, \sum_{i' \in I_c} \frac{\vec{\imath'}}{|I_c|} \right),$$
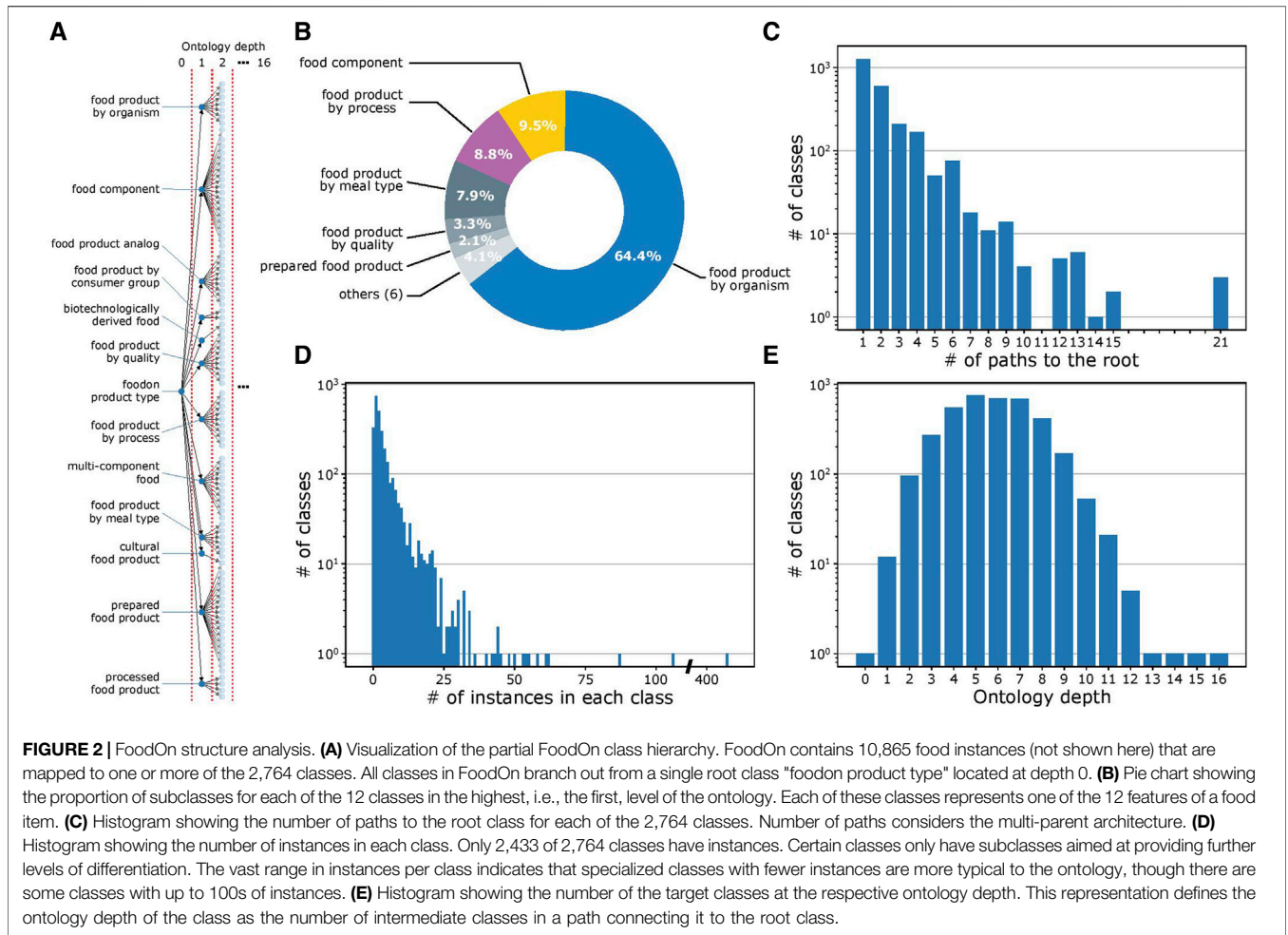
where $\vec{\imath}$ is the word embedding vector created by taking the average of the constituent word embeddings, $|I_c|$ is the number of all the seed instances in I_c, and sim () is the measure of similarity between the two word embedding vectors. $\text{score}_{\text{parent}}$ is the similarity of the food instance i with the target class c:

$$\text{score}_{\text{parent}}(c; i) = \text{sim}\left(\vec{\imath}, \vec{c} \right).$$

Finally, predicting which target class $\vec{c}$ the food instance i will get mapped to can be formulated as follows:

$$\bar{c} = \arg\max_{c} \text{score}(c; i).$$

For the scope of this work, we map the food instance to a single target class even if it was originally mapped to multiple classes. For the case of FoodOn, we observed that the precision of ontology learning increases as the number of seed food instances per class (n_{seed}) increases (**Supplementary Figure S1**) as a class is better represented as the number of seed instances increases. For sim (), we used Euclidean distance and cosine similarity, with the latter having better performance and used throughout this work (**Supplementary Figure S2**). We empirically set $\alpha = 0.8$ after testing all values between 0.0 and 1.0 with an interval of 0.1 (**Supplementary Figure S3**).

FIGURE 2 | FoodOn structure analysis. **(A)** Visualization of the partial FoodOn class hierarchy. FoodOn contains 10,865 food instances (not shown here) that are mapped to one or more of the 2,764 classes. All classes in FoodOn branch out from a single root class "foodon product type" located at depth 0. **(B)** Pie chart showing the proportion of subclasses for each of the 12 classes in the highest, i.e., the first, level of the ontology. Each of these classes represents one of the 12 features of a food item. **(C)** Histogram showing the number of paths to the root class for each of the 2,764 classes. Number of paths considers the multi-parent architecture. **(D)** Histogram showing the number of instances in each class. Only 2,433 of 2,764 classes have instances. Certain classes only have subclasses aimed at providing further levels of differentiation. The vast range in instances per class indicates that specialized classes with fewer instances are more typical to the ontology, though there are some classes with up to 100s of instances. **(E)** Histogram showing the number of the target classes at the respective ontology depth. This representation defines the ontology depth of the class as the number of intermediate classes in a path connecting it to the root class.

Evaluation Metrics of the Ontology Structure

The granularity and cohesiveness metrics have to do with fundamental design questions of ontologies such as the optimum number of classes and whether a class is overspecified or underspecified (Whetzel et al., 2011). *Granularity* is semantically defined as the ability to represent different levels of detail in data (Keet, 2008). In our work, we quantitatively define granularity of a certain ontology superclass c_A as

$$\text{granularity } (c_A) = \frac{|I_{C^A}|}{|C^A|},$$

where $C^A \subseteq C$ is the set of all the classes that have c_A as their superclass, I_{C^A} is the set of all food instances belonging to C^A, and c_A is a superclass of c_B if every instance of c_A is also an instance of c_B (Noy and McGuinness, 2011). *Cohesiveness* of a superclass is a measure of subclass semantic relevance, and by corollary, the degree of its subclasses has the same relation to each other (Gangemi et al., 2005). Here, we quantitatively define the cohesiveness of a certain ontology superclass c_A as
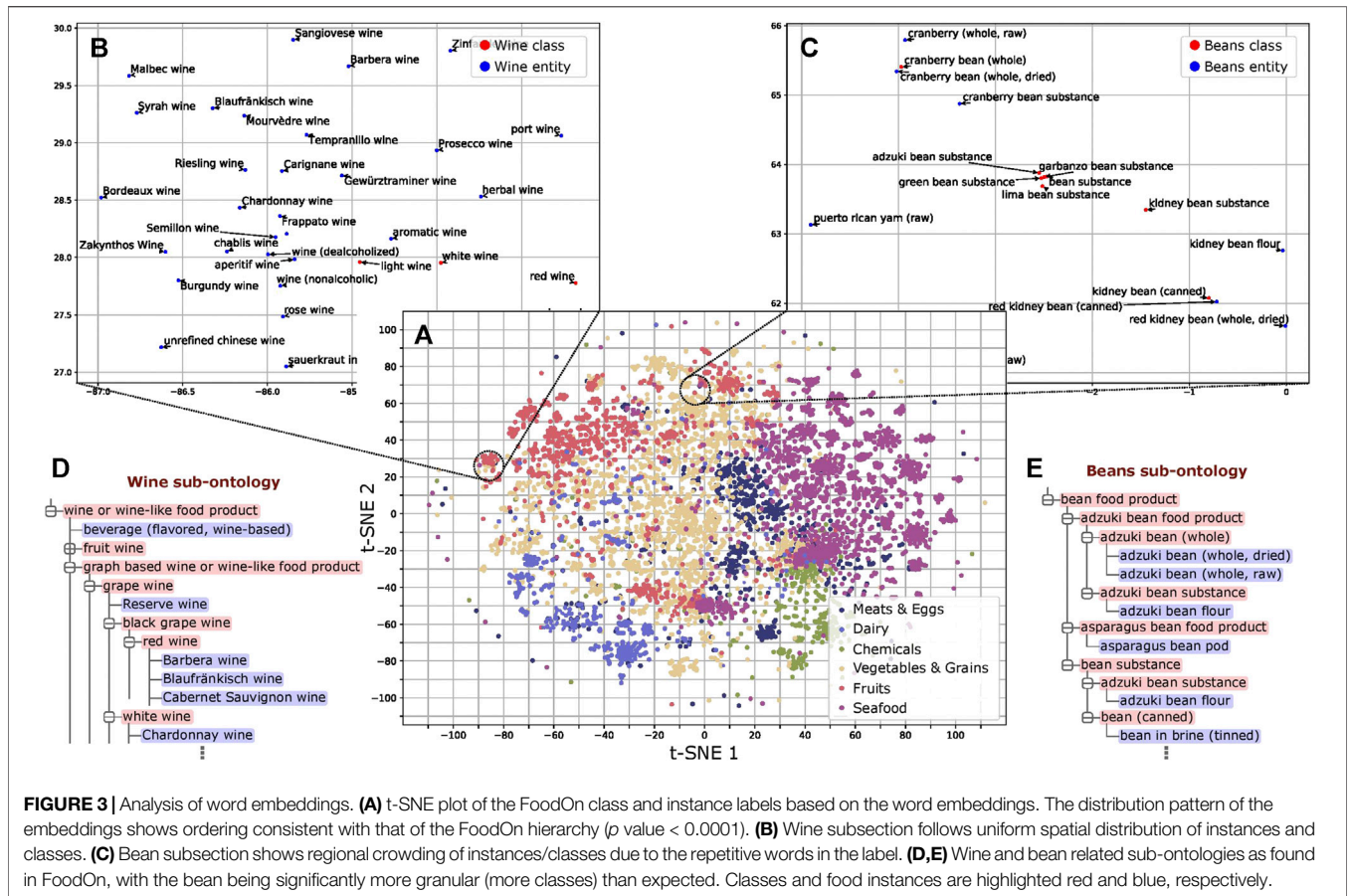
$$\text{cohesiveness} (c_A) = \frac{|C'^A|}{|C^A|},$$

where C'^A is the set of all correct subclasses within the superclass c_A. For example, in the superclass "cheese food product by organism" in FoodOn, the subclasses "cow cheese," "goat cheese," "sheep cheese," and "buffalo milk cheese" are correct, while the subclass "blue cheese" is not since it describes a method/process and not the point of origin. In this case, the cohesiveness value would be 4/5 = 0.8. Another example is in the case of the bean superclass where the subclasses that are bean varieties are correct and subclasses for processed forms of beans such as "bean flour" are not. The cohesiveness of the cheese superclass is 0.52, implying that only half of the subclasses are correct, and the bean superclass has a higher cohesiveness of 0.93 (**Supplementary File 2**).

Success Metric of the Ontology Population

We use precision to assess the performance of the ontology population and define it as follows:

$$\text{precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})},$$

FIGURE 3 | Analysis of word embeddings. **(A)** t-SNE plot of the FoodOn class and instance labels based on the word embeddings. The distribution pattern of the embeddings shows ordering consistent with that of the FoodOn hierarchy (*p* value < 0.0001). **(B)** Wine subsection follows uniform spatial distribution of instances and classes. **(C)** Bean subsection shows regional crowding of instances/classes due to the repetitive words in the label. **(D,E)** Wine and bean related sub-ontologies as found in FoodOn, with the bean being significantly more granular (more classes) than expected. Classes and food instances are highlighted red and blue, respectively.

where a food instance $i \in I_c$ is considered a TP if and only if the mapping function correctly placed i under c, and FP, otherwise. In addition, we define the path distance to be the shortest distance (hops) between the predicted class and the ground truth class, where a perfect ontology population algorithm would have a path distance of 0.

RESULTS

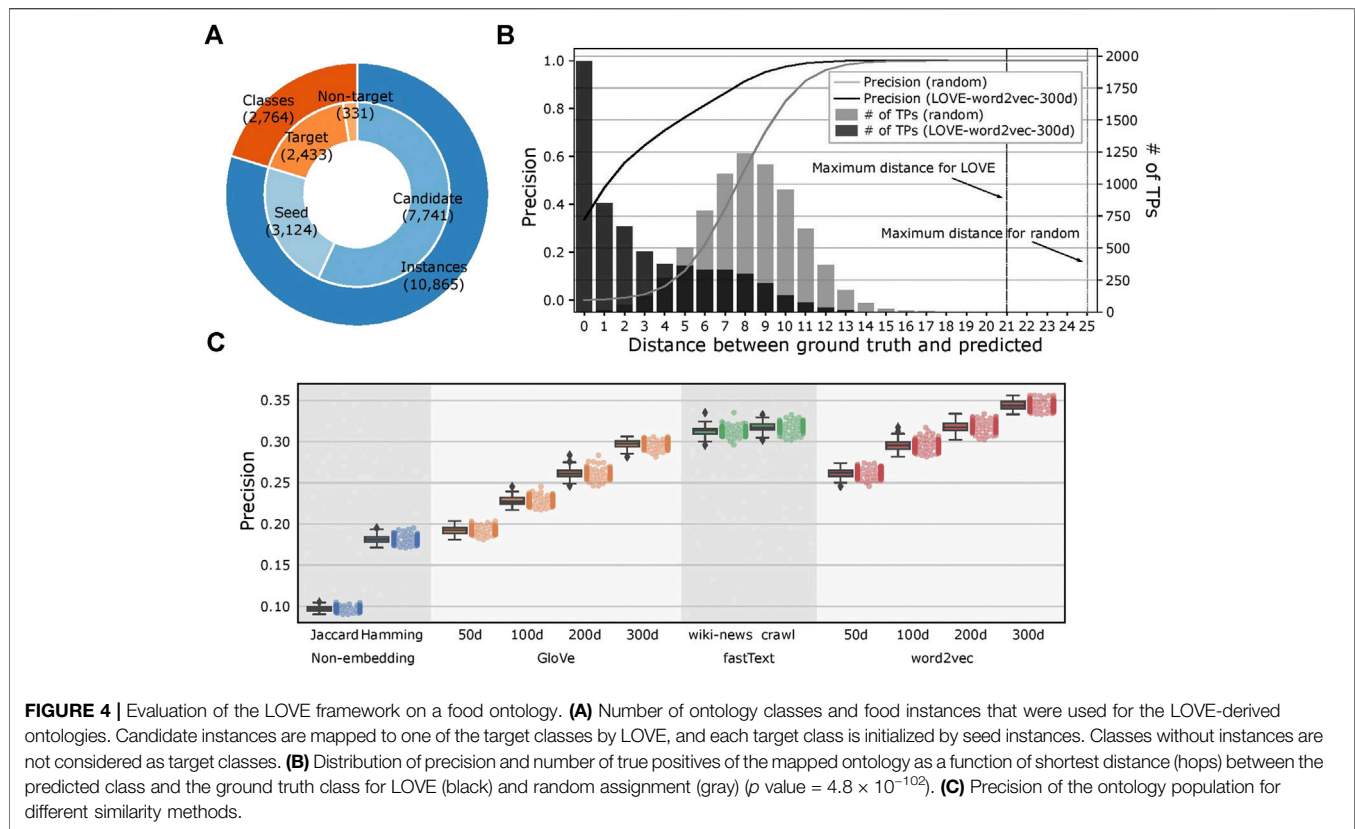### Structural Topology of FoodOn

Figure 2A provides a visualization of the ontology structure for the 2,764 classes in FoodOn. At the highest level of the ontology, every food item is described by various features, which minimally include its source organism and up to 11 other features, with each feature represented as a class (processes and material quality, among others; **Figure 2B**). A complete examination indicates that the ontology structure is heterogeneous in its granularity, with some classes having many subclasses and interconnectivity, while others have only one subclass. In a similar trend, while some classes have hundreds of instances, other classes have only one (**Figures 2C and D**). **Figure 2E** illustrates the variation in ontology depth for a given class, which is defined as the number of intermediate classes present in a given path that connects it to the root (Blanchard et al., 2005). Considering all the factors mentioned above, the FoodOn ontology is highly granular with an average of 3.15 food instances per class.

Granularity and Cohesiveness Impair Precision of Automated Methods

We trained word embeddings for the 13,629 instances and class labels in FoodOn to use in our method. These embeddings capture latent information of the food type as revealed by dimensionality reduction (Maaten and Hinton, 2008) and subsequent analysis (**Figure 3A** and **Supplementary File 1**). Regarding the structure of FoodOn, the granularity differs substantially as shown in **Figures 3B–E**, where we compare the superclasses "wine" and "beans," with granularity 5.64 vs. 1.96, respectively. We also noticed inconsistencies in the further classification of each superclass which we quantify by the cohesiveness. Relevant to our work of ontology learning, we found that both the cohesiveness and the granularity are positively associated with better ontology population performance (PCC of 0.56 and 0.51, respectively; p value = 2.5×10^{-2} and 4.5×10^{-2}, respectively) (**Supplementary Figure 4**; **Supplementary File 2**).

Learning Ontology via Embeddings Leads to Substantially Better Performance

We kept the ontological structure of FoodOn unchanged with 2,433 target ontology classes and created 100 different seeded-skeleton ontologies to test the statistical significance of the methods by selecting two random seeds for each target class. This process resulted in 3,124 food instances used as seeds from a total of

FIGURE 4 | Evaluation of the LOVE framework on a food ontology. **(A)** Number of ontology classes and food instances that were used for the LOVE-derived ontologies. Candidate instances are mapped to one of the target classes by LOVE, and each target class is initialized by seed instances. Classes without instances are not considered as target classes. **(B)** Distribution of precision and number of true positives of the mapped ontology as a function of shortest distance (hops) between the predicted class and the ground truth class for LOVE (black) and random assignment (gray) (p value = 4.8×10^{-102}). **(C)** Precision of the ontology population for different similarity methods.

10,865 instances, and the task was to map the remaining 7,741 food instances to the target classes (**Figure 4A**; **Supplementary File 3**). The LOVE-generated ontology, which uses the word embeddings of size 300 d trained using the Wikipedia corpus, had a significantly reduced path distance from what is expected from random chance (p value = 4.8×10^{-102}; **Figure 4B**). Moreover, ontology population methods based on the word embeddings performed better when compared to the traditional text similarity methods regardless of the embedding size or the training algorithm, with an 89.7% increased precision (0.34 vs. 0.18, respectively, p value = 2.6×10^{-138}) and a 43.6% shorter path distance (2.91 vs. 5.16, respectively, p value= 4.7×10^{-84}; **Figure 4C**; **Table 1**).

DISCUSSION

As shown in **Figure 3A**, there is an alignment of the word embeddings and the FoodOn classes at a high level. However, through deeper analysis of the ontology structure and the results of automated ontology learning, we discovered the causes for discrepancies between the user-defined ontology and the ontology representation from the corpus. The granularity and cohesiveness issues impacting the precision have to do with a well-known and fundamental design question of how many classes are too few or too many (Noy and McGuinness, 2011). The classes with lower than average granularity of 4 combine several features of a food item such as its source, process, and organoleptic quality. However, the

nomenclature is not consistent as it varies from a long and precise class name to less-precise representations. This is not a scalable approach to a data-driven automated ontology since it will require manually curated classes when mapping foods of yet unknown features such as sources and processes. Moreover, it will lead to errors in mapping class-class and class-instance relations if done manually, as the ontology grows. To avoid these issues, an extension would be for every variety-specific subclass to contain a flat list of instances. For example, in **Figure 3E**, the food instance "adzuki bean flour" is mapped to two parent classes in the bean superclass. Instead, the "product by process" class at a depth of one can have a subclass of "milled food" which aggregates all the flour variants and notably the "bean flour" class. This also addresses the problem of cohesiveness described in the "*Methods*" section. The ontology learning function can then be applied on each of the 12 highest parent classes (**Figure 2B**).

Taking into account the structural similarity between the ontology and the knowledge graph, we considered applying observable and latent feature-based link prediction models (Toutanova et al., 2015; Grover and Leskovec, 2016; Lao and Cohen, 2010) to populate the ontology. However, such models either are dependent on external data or require at least one pre-existing path connecting the candidate instance to the target class. A possible extension to our work is to train the word embeddings using other related corpora such as food-related literature and databases, for example, the FDC database (US Department of Agriculture, Agricultural Research Service, 2019). Moreover, the pertinent information can be extended to chemical composition,

phenotypic effects, and association with health states. Another natural extension would be to train methods that encode the hierarchical structure of the knowledge graphs, such as Poincaré embeddings (Nickel and Kiela, 2017), with hierarchical food domain data (Haussmann et al., 2019) for the ontology population task. Along with an optimally designed skeleton ontology, we expect that these improvements would lead to much improved accuracy of the automatically generated ontology.

CONCLUSION

In this work, we applied the learning ontologies via embeddings (LOVE) framework, which takes advantage of the semantic similarity of the word embeddings to the field of food ontologies. The automated method we proposed here is a solution to the manual burden of populating an ontology with continuous influx of new data. Therefore, the desired automation would be a semi-supervised method that yields high precision, with minimal manual intervention. Although the importance of automated ontology learning has been discussed before (Drumond and Girardi, 2008), to the best of our knowledge, this is the first time a solution is applied to an existing ontology in the food domain. We believe that our work is a step towards the fully automated ontologies.

REFERENCES

Alexander, P., Brown, C., Arneth, A., Finnigan, J., Moran, D., and Rounsevell, M. D. A. (2017). Losses, inefficiencies and waste in the global food system. *Agric. Syst.* 153, 190–200. doi:10.1016/j.agsy.2017.01.014

Barabási, A.-L., Menichetti, G., and Loscalzo, J. (2019). The unmapped chemical complexity of our diet. *Nat. Food* 1, 33–37. doi:10.1038/s43016-019-0005-1

Benslimane, D., Arara, A., Falquet, G., Maamar, Z., Thiran, P., and Gargouri, F. (2006). "Contextual ontologies motivations, challenges, and solutions," in Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), Vol. 4243 (Berlin: Springer-Verlag), LNCS 168–176.

Berners-Lee, M., Kennelly, C., Watson, R., and Hewitt, C. N. (2018). Current global food production is sufficient to meet human nutritional needs in 2050 provided there is radical societal adaptation. *Elem. Sci. Anth.* 6, 52. doi:10.1525/elementa. 310

Blanchard, E., Harzallah, M., Briand, H., and Kuntz, P. (2005). A typology of ontology-based semantic measures. *EMOI-INTEROP* 160, 3–11. doi:10.1007/3-540-34416-0_1

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). *Enriching word vectors with subword information*. Cambridge, MA: MIT Press.

Dabbene, F., and Gay, P. (2011). Food traceability systems: performance evaluation and optimization. *Comput. Electron. Agric.* 75, 139–146. doi:10.1016/j.compag. 2010.10.009

Dooley, D. M., Griffiths, E. J., Gosal, G. S., Buttigieg, P. L., and Hoehndorf, R., Lange, M. C., et al. (2018). FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. *Npj Sci. Food* 2, 1–10. doi:10.1038/s41538-018-0032-6

Drumond, L., and Girardi, R. (2008). *A survey of ontology learning procedures*. researchgate.net. Available at: http://gesec.deinf.ufma.br (Accessed July 18, 2020). doi:10.1145/1774088.1774379

EFSA. (2015). The food classification and description system FoodEx 2 (revision 2). *EFSA Support. Publ.* 12, 804E. doi:10.2903/sp.efsa.2015.en-804

DATA AVAILABILITY STATEMENT

All data, code and instructions on how to reproduce the results are available at https://github.com/IBPA/LOVE.

AUTHOR CONTRIBUTIONS

TN and JY contributed equally in preparation for the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frai.2020.584784/full#supplementary-material

Eftimov, T., Ispirova, G., Potočnik, D., Ogrinc, N., and Seljak, B. K. (2019). ISO-FOOD ontology: a formal representation of the knowledge within the domain of isotopes for food science. *Food Chem.* 277, 382–390. doi:10.1016/j.foodchem. 2018.10.118

Ehrlinger, L., and Wöß, W. (2016). Toward a definition of knowledge graphs. researchgate.net. Available at: https://www.researchgate.net/publication/323316736 (Accessed July 18, 2020).

Gangemi, A., Catenacci, C., Ciaramita, M., and Lehmann, J. (2005). Ontology evaluation and validation: an integrated formal model for the quality diagnostic task. Available at: http//www.loa-cnr.it/Files/OntoEval4OntoDev_Final. pdf (Accessed July 18, 2020).

Genesereth, M. R., and Nilsson, N. J. (2012). *Logical foundations of artificial intelligence*. Burlington, MA: Morgan Kaufmann.

Grover, A., and Leskovec, J. (2016). node2vec: scalable feature learning for networks. dl.acm.org 13–17 August 2016, 855–864.

GS1. (2018). GPC.

Guyomard, H., Darcy-Vrillon, B., Esnouf, C., Marin, M., Russel, M., and Guillou, M. (2012). Eating patterns and food systems: critical knowledge requirements for policy design and implementation. *Agric. Food Secur.* 1, 13. doi:10.1186/2048-7010-1-13

Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell Syst. Tech. J.* 29, 147–160. doi:10.1002/j.1538-7305.1950.tb00463.x

Haussmann, S., Seneviratne, O., Chen, Y., Ne'eman, Y., Codella, J., et al. (2019). "FoodKG: a semantics-driven knowledge graph for food recommendation," in Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), Vol. 11779 (Berlin, Germany: Springer), LNCS 146–162. doi:10.1007/978-3-030-30796-7_10

Jayawardana, V., Lakmal, D., de Silva, N., Perera, A. S., Sugathadasa, K., Ayesha, B., et al. (2017). "Semi-supervised instance population of an ontology using word vector embedding," in 2017 Seventeenth international conference on advances in ICT for emerging regions (ICTer), Colombo, Sri Lanka, 6–9 September, 2017 (Piscataway, NJ: IEEE), 1–7. doi:10.1109/icter.2017.8257822

Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2016). Bag of tricks for efficient text classification. arXiv Preprint arXiv1607.01759.

Keet, C. M. (2008). *A formal theory of granularity*. Bolzano, Italy: Free University of Bozen-Bolzano.

Kretsera, A., Murphya, D., Finleyb, J., and Brennerc, R. (2015). A partnership for public health: branded food products database. *Proc. Food Sci.* 4, 18–26. doi:10. 1016/j.profoo.2015.06.005

Lao, N., and Cohen, W. W. (2010). Relational retrieval using a combination of path-constrained random walks. *Mach. Learn.* 81, 53–67 (2010). doi:10.1007/ s10994-010-5205-8

Lemay, D. G., Zivkovic, A. M., and German, J. B. (2007). Building the bridges to bioinformatics in nutrition research. *Am. J. Clin. Nutr.* 86, 1261–1269. doi:10. 1093/ajcn/86.5.1261

Maaten, L. v. d., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605. doi:10.1007/s10994-011-5273-4

Mahmoud, N., Elbeh, H., and Abdlkader, H. M. (2018). "Ontology learning based on word embeddings for text big data extraction," in 2018 14th international computer engineering conference (ICENCO), Cairo, Egypt, 29–30 December 2018 (IEEE), 183–188. doi:10.1109/icenco.2018.8636154

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Adv. Neural Inf. Process. Syst.* 2, 3111–3119.

Mitchell, T. (2018). Never-ending learning. *Commun. ACM* 61, 103–115. doi:10. 1037/e660332010-001

Moe, T. (1998). Perspectives on traceability in food manufacture. *Trends Food Sci. Technol.* 9, 211–214. doi:10.1016/s0924-2244(98)00037-5

Nickel, M., and Kiela, D. (2017). Poincaré embeddings for learning hierarchical representations. papers.nips.cc. http://papers.nips.cc/paper/7213-poincare-embeddings-for-learning-hierarchical-represent.

Niwattanakul, S., Singthongchai, J., Naenudorn, E., and Wanapu, S. (2013). "Using of Jaccard coefficient for keywords similarity," in Proceedings of the international multiconference of engineers and computer scientists, Hong Kong, 13–15 March 2013, Vol. 1, 380–384.

Noy, N. F., and McGuinness, D. L. (2001). *Ontology development 101: a guide to creating your first ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.

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), Doha, Qatar (Stroudsburg, PA: Association for Computational Linguistics) 1532–1543.

Rehurek, R., and Sojka, P. (2010). "Software, Framework for topic modeling with large corpora," in Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks (ELRA), Paris, France, 45–50.

Toutanova, K., Chen, D., Pantel, P., Poon, H., Choudhury, P., Gamon, M., et al. (2015). Representing text for joint embedding of text and knowledge bases. aclweb.org. http://lemurproject.org/clueweb12/ (Accessed July 18, 2020).

Tscharntke, T., Clough, Y., Wanger, T. C., Jackson, L., Motzke, I., Perfecto, I., et al. (2012). Global food security, biodiversity conservation and the future of agricultural intensification. *Biol. Conserv.* 151, 53–59. doi:10.1016/j.biocon.2012.01.068

US Department of Agriculture, Agricultural Research Service (2019). *FoodData central*.

Whetzel, P., Noy, N., and Shah, N. H. (2011). BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. *Nucleic Acids Res.* 39 (Suppl. 2), W541–W545. doi:10.1093/nar/gkr469

Wilkinson, M. D., Dumontier, M., and Aalbersberg, I. J. J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 1–9. 10.1038/sdata.2016.18

Wishart, D. S. (2018). *FooDB: the food database*. doi:10.1007/978-1-59259-335-4_27

World Wide Web Consortium. (2011). *RDF 1.1 concepts and abstract syntax*, W3C recommendation 25, No. 02.