OWL2Vec*: Embedding of OWL Ontologies

Jiaoyan Chen, Pan Hu, Ernesto Jimenez-Ruiz, Ole Magnus Holter,
Denvar Antonyrajah, Ian Horrocks

UNIVERSITY OF OXFORD

CITY UNIVERSITY LONDON

Background: OWL Ontology and Examples

TBox:
- Atomic concepts (classes)
- Atomic roles (object properties)
- Complex classes and properties (by logical constructors)
- Constraints, subsumption relationship, etc.

ABox:
- Individuals (instances)
- Membership assertions, role assertions (facts)
- OWL/RDF/RDFS Built-in and bespoke properties

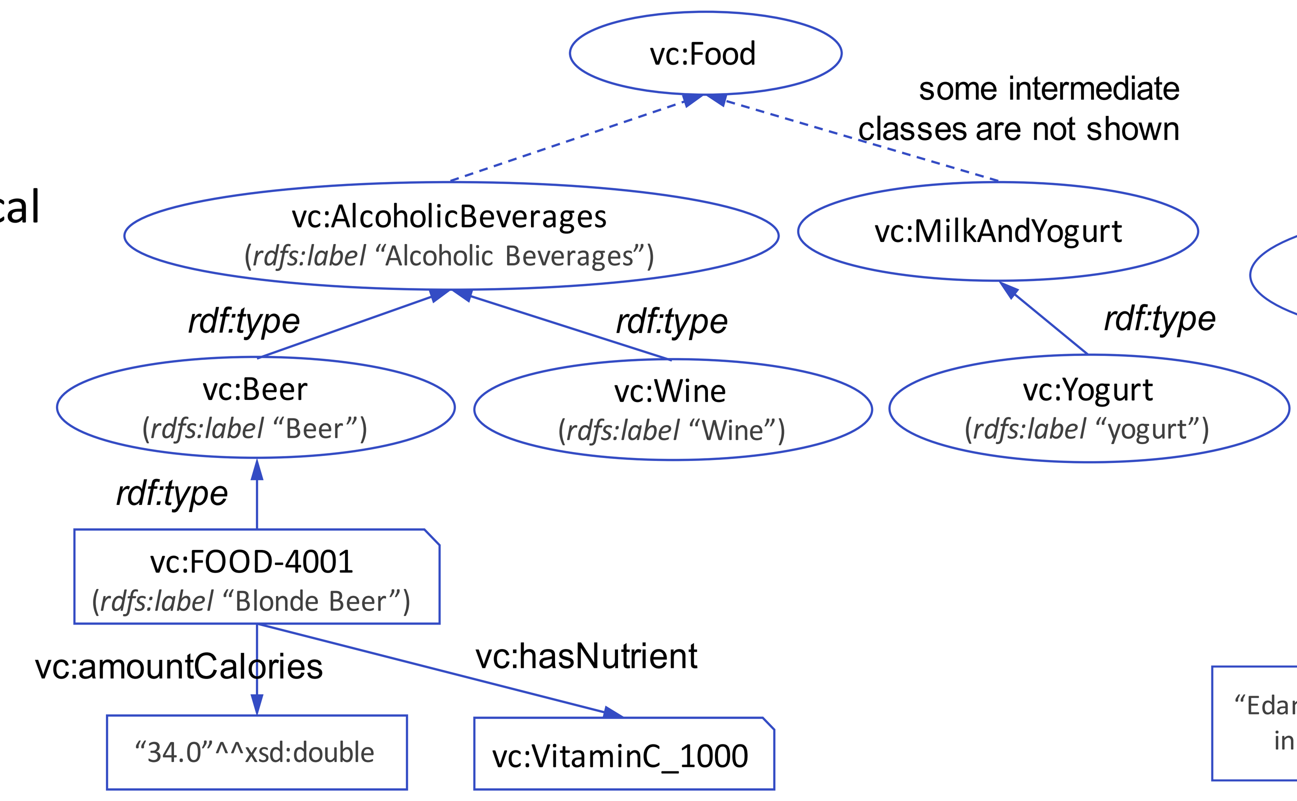


Fig. 1. The healthy lifestyle ontology named HeLis

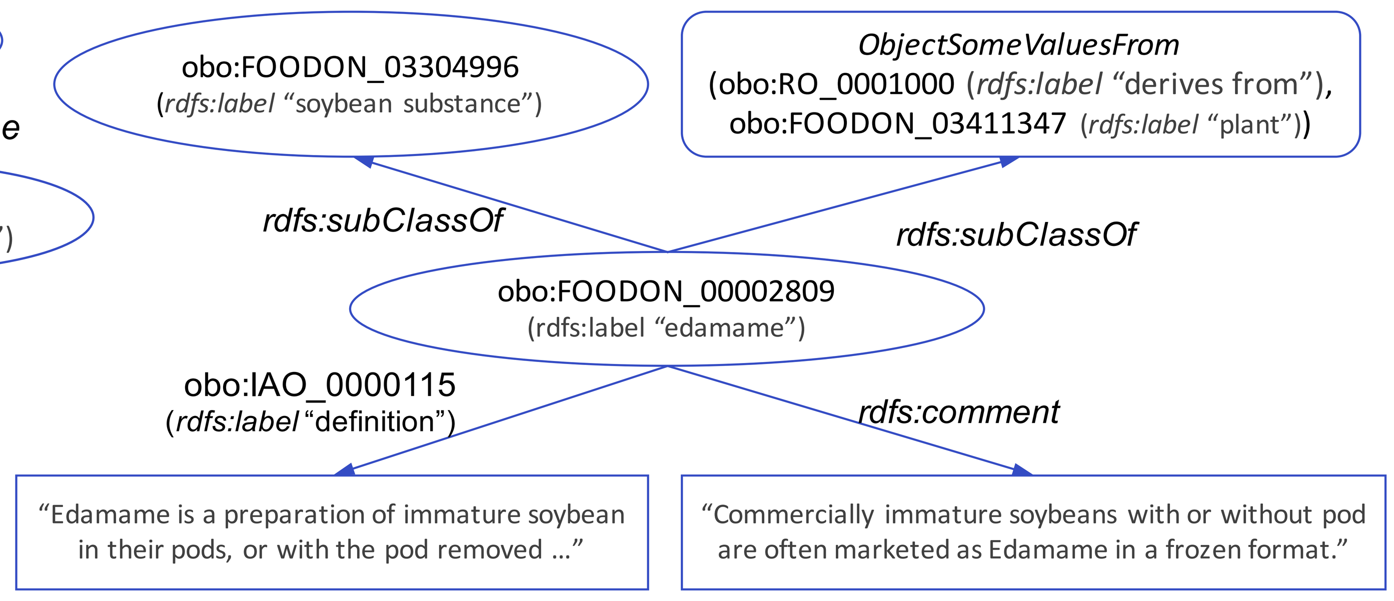
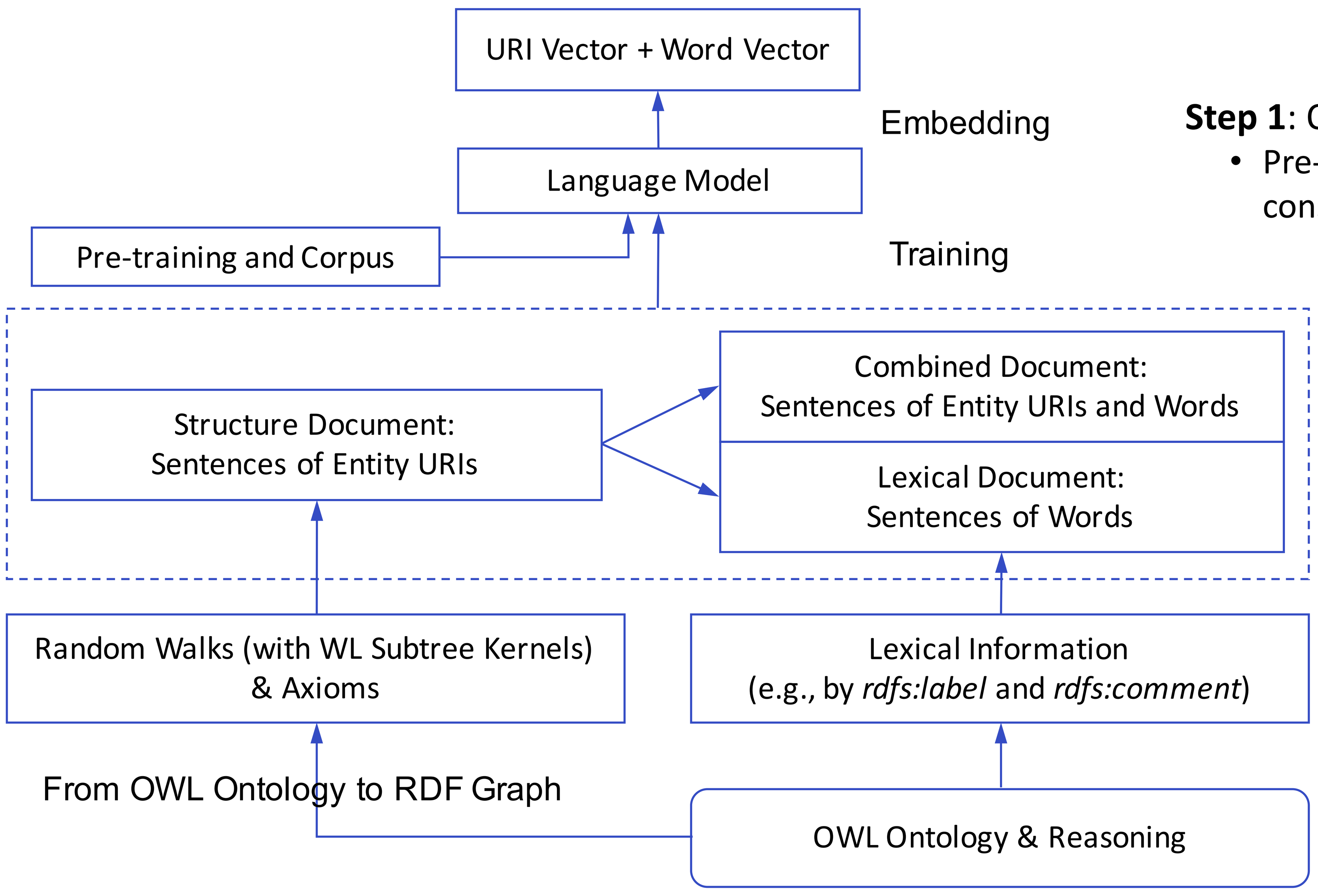


Fig. 2. The healthy lifestyle ontology named HeLis

Others:
- Meta data by annotation properties: label, synonym, class definition, comment, etc.
- Entity, URI (prefix + name/id), Literal

Methodology: Corpus Generation and Word2Vec Embedding



Step 1: OWL Ontology to RDF Graph
- Pre-reasoning by HermiT is considered and evaluated

Solution #1: W3C OWL to RDF Graph Mapping e.g.,
<obo:FOODON_00002809, rdfs:subClassOf, _:x>
<_:x, rdf:type, owl:Restriction>
<_:x, owl:OnProperty, obo:RO_0001000>
<_:x, owl:SomeValueFrom, obo:FOODON_03411347>

Solution #2: Projection Rules e.g.,
<obo:FOODON_00002809, rdfs:subClassOf,
obo:FOODON_03411347

Axiom of Condition 1	Axiom or Triple(s) of Condition 2	Projected Triple(s)
$A \sqsubseteq \square r.D$ or $\square r.D \sqsubseteq A$	$D \equiv B \mid B_1 \sqcup ... \sqcup B_n \mid B_1 \sqcap ... \sqcap B_n$	$\langle A, r, B\rangle$ or $\langle A, r, B_i\rangle$ for $i \in 1, ..., n$
$\exists r.\top \sqsubseteq A$ (domain)	$\top \sqsubseteq \forall r.B$ (range)	
$A \sqsubseteq \exists r.\{b\}$	$B(b)$	
$r \sqsubseteq r'$	$\langle A, r', B\rangle$ has been projected	
$r' \equiv r^-$	$\langle B, r', A\rangle$ has been projected	
$s_1 \circ ... \circ s_n \sqsubseteq r$	$\langle A, s_1, C_1\rangle ... \langle C_n, s_n, B\rangle$ have been projected	
$B \sqsubseteq A$	–	$\langle B, rdfs{:}subClassOf, A\rangle$ $\langle A, rdfs{:}subClassOf^-, B\rangle$
$A(a)$	–	$\langle a, rdf{:}type, A\rangle$ $\langle A, rdf{:}type^-, a\rangle$
$r(a, b)$	–	$\langle a, r, b\rangle$

Step 2: Structure Document -- Sequences of entities (IRIs)
- Random walk, Random walk + Weisfeiler-Lehman (WL) subtree kernel
- Axioms (OWL Manchester Syntax)

Step 3: Lexical Document -- Sequences of words
- Text from annotation properties (e.g., labels, comment and definition)
- From structure document (replacing the entities by their labels)

Step 4: Combined Document -- Sequences of words and IRIs
- Replace partial entities in every entity sequence by their words
- Strategies: random selection of an entity for replacing or traversal of entities for replacing

Step 5: Word Embedding Model -- CBOW architecture of Word2Vec
- Using the pre-trained Word2Vec model
- Using URI vector vs average word vector in applying the embeddings

(Potential) Applications and Discussion

- **Ontology completion (membership and subsumption prediction)**
  - Input: the OWL2Vec* embeddings of two entities (as learned features)
  - Classifiers e.g., Random Forest and Logistic Regression
  - Learn from the known axioms, predict the missing axioms (partitioning the axioms into train/valid/test sets in evaluation)
  - Output: a score in [0, 1]
- **Ontology Clustering [1]**
- **Ontology Alignment**
  - [2]: OWL2Vec* embeddings as features; confident mappings from LogMap and disjointness rules for training (distant supervision)
- **Neural-symbolic AI**
  - Ontology-guided Zero-shot Learning [3][4][5]: ontology can act as external knowledge for describing the relationship between "seen" classes and "unseen" classes; the key technique is ontology (with logical expressions, taxonomy, meta data) embedding where OWL2Vec* can be applied (TODO)
  - Ontology for machine learning sample shortage

[1] Ritchie, A., Chen, J., Castro, L. J., Rebholz-Schuhmann, D., & Jimenez-Ruiz, E. (2021, April). Ontology Clustering with OWL2Vec. DeepOntoNLP – ESWC'21

[2] Chen, J., Jiménez-Ruiz, E., Horrocks, I., Antonyrajah, D., Hadian, A., & Lee, J. (2021). Augmenting Ontology Alignment by Semantic Embedding and Distant Supervision. ESWC'21.

[3] Chen, J., Lécué, F., Geng, Y., Pan, J. Z., & Chen, H. (2020). Ontology-guided Semantic Composition for Zero-Shot Learning. KR'20.

[4] Geng, Y., Chen, J., Chen, Z., Pan, J. Z., Ye, Z., Yuan, Z., ... & Chen, H. (2021). OntoZSL: Ontology-enhanced Zero-shot Learning. WWW'21.

[5] Chen, J., Geng, Y., Chen, Z., Horrocks, I., Pan, J. Z., & Chen, H. (2021). Knowledge-aware Zero-Shot Learning: Survey and Perspective. IJCAI'21, Survey Track.

Results – Membership Prediction on HeLis and Subsumption Prediction on FoodOn and GO

Overall Results: Quantum Embedding, EL Embedding, Onto2Vec, TransE < RDF2Vec, TransR, DistMult, OWL2Vec, OPA2Vec < Transformer < Pre-trained Word2vec < **OWL2Vec***

	HeLis			
Method	MRR	Hits@1	Hits@5	Hits@10
Transformer (*label*)	0.657	0.515	0.824	0.897
Transformer (*all text*)	0.599	0.390	0.870	0.912
RDF2Vec	0.345	0.219	0.460	0.655
TransE	0.181	0.09	0.232	0.355
TransR	0.298	0.184	0.391	0.559
DistMult	0.253	0.166	0.304	0.437
Quantum Embeding	0.159	0.132	0.163	0.190
Onto2Vec	0.211	0.108	0.268	0.397
OPA2Vec	0.237	0.146	0.286	0.408
OWL2Vec	0.335	0.215	0.397	0.601
Pre-trained *Word2Vec*	0.899	0.877	0.923	0.933
OWL2Vec*	**0.953**	**0.932**	**0.978**	**0.987**

Ablation Study:
- HeLis: All three documents ≈ structural + literal documents; URI vector + word vector > word vector alone >> URI vector alone
- FoodOn and GO: Structural document + literal document + word vector is the best
- Walking depths, walking strategies, pre-training and reasoning have all ben evaluated

	FoodOn				GO			
Method	MRR	Hits@1	Hits@5	Hits@10	MRR	Hits@1	Hits@5	Hits@10
Transformer (*label*)	0.016	0.005	0.027	0.046	0.009	0.001	0.009	0.018
Transformer (*all text*)	0.022	0.011	0.032	0.050	0.014	0.008	0.017	0.019
RDF2Vec	0.078	0.053	0.097	0.119	0.043	0.017	0.057	0.087
TransE	0.029	0.011	0.044	0.065	0.015	0.005	0.018	0.030
TransR	0.072	0.044	0.093	0.130	0.048	0.016	0.076	0.113
DistMult	0.076	0.045	0.099	0.134	0.046	0.018	0.68	0.097
EL Embeding	0.040	0.014	0.067	0.099	0.018	0.005	0.021	0.036
Onto2Vec	0.034	0.014	0.047	0.064	0.024	0.008	0.031	0.053
OPA2Vec	0.093	0.058	0.117	0.156	0.075	0.032	0.106	0.157
OWL2Vec	0.091	0.052	0.121	0.152	0.031	0.012	0.040	0.067
Pre-trained *Word2Vec*	0.136	0.089	0.175	0.227	0.123	0.055	0.177	0.260
OWL2Vec*	**0.213**	**0.143**	**0.287**	**0.357**	**0.170**	**0.076**	**0.258**	**0.376**