

# Supplementary material for: Semantic Web and Zero-Shot Learning of Large Scale Visual Classes

## Appendix A: Linked Open Data

### A.1 Linking process

Figure 1. presents the interlinking of the different knowledge bases used in the article. Blue ellipses represent knowledge bases. Under the name of the knowledge base figures the type and number of entities of the knowledge base mapped to Imagenet classes by our process. Blue wide arrows represent the percentage of resources matched between two knowledge bases. We can see that the main bottleneck of our linking process is between Babel synsets and DBPedia resources as only 58.0% of Babel synsets are linked to a DBPedia resource.

Green rectangles correspond to the vector representations extracted from the knowledge bases. A second bottleneck concerns sense embeddings. Sense embeddings are obtained by first performing word-sense disambiguation on the English Wikipedia corpus using BabelNet. Then a word2vec model is learned on the disambiguated corpus. However, word-senses appearing less than 5 times in the whole corpus are discarded. Due to this limitation, only 18.627 (85.3%) Babel synsets find a sense embedding representation. In total, we gathered 11.609 classes for which we generated both sense embeddings and Wikipedia mappings.

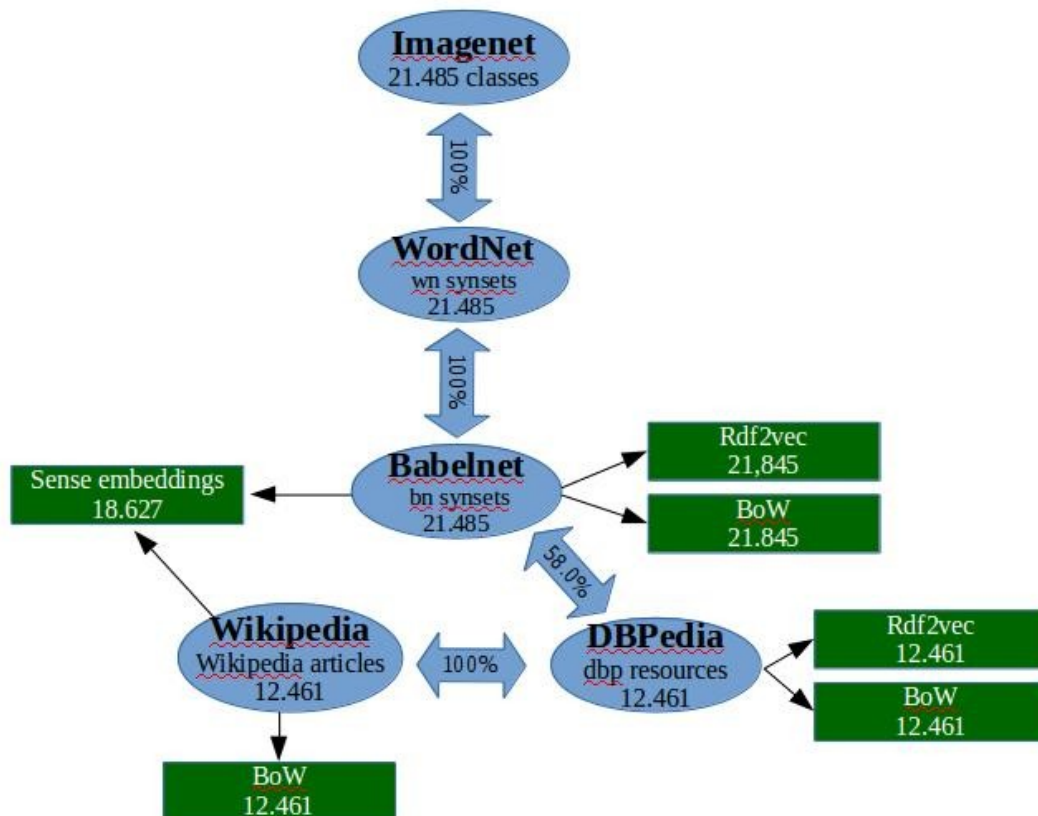


Figure 1. Schema of the linking process

## A.2 Wordnet/Wikipedia mapping evaluation

To assess the quality of the mapping we generated between Wordnet synsets and Wikipedia articles, we manually inspect the Wikipedia articles associated to 100 Wordnet synsets. To do so, we compare the gloss of Wordnet synsets to the first introduction lines of Wikipedia articles. We subjectively assess the quality of the pair correspondence as perfect match/imperfect match and complete miss. Table 1. below illustrates examples of associations.

Quality	Wordnet synset	Wordnet gloss	Wikipedia title	Wikipedia intro
Perfect Match	n12514592	<b>Tagasaste, Chamaecytisus pamensis, Cytesis proliferus</b> , (shrub of Canary Islands having bristle-tipped oblanceolate leaves; used as cattle fodder)	<i>Cytisus proliferus</i>	<b><i>Cytisus proliferus</i>, tagasaste or tree lucerne</b> , is a small spreading evergreen tree that grows 3-4m high. It is a well known fertilizer tree. It is a member of the Fabaceae (pea) family and is indigenous to the dry volcanic slopes of the Canary islands.
Perfect Match	n03304197	<b>Exocet</b> (a guided missile developed by the French government for use against ships)	Exocet	The <b>Exocet</b> (French for “flying fish”) is a French-built anti-ship missile whose various versions can be launched from surface vessels, submarines, helicopters and fixed-wing aircraft.
...				
Imperfect Match	n08495908	<b>aphelion</b> (apoapsis in solar orbit; the point in the orbit of a planet or comet that is at the greatest distance from the sun)	Perihelion and aphelion	The <b>perihelion</b> is the point in the <a href="#">orbit</a> of a <a href="#">celestial body</a> where it is nearest to its orbital focus, generally a star. It is the opposite of <b>aphelion</b> , which is the point in the orbit where the celestial body is farthest from its focus
Imperfect Match	n04104500	<b>Roman building</b> (a building constructed by the ancient Romans)	Ancient Roman architecture	<b>Ancient Roman architecture</b> adopted the external language of classical <a href="#">Greek architecture</a> for the purposes of the <a href="#">ancient Romans</a> , but grew so different from Greek buildings as to become a new <a href="#">architectural</a> style
Miss	n03238131	<b>dressing room</b> (a room in which you can change clothes)	Parts of a theater	No Intro
Miss	n02934168	<b>cable, line, transmission line</b> (a conductor for transmitting electrical or optical signals or electric power)	Cable (foreign exchange)	<b>Cable</b> (or <b>the cable</b> ) is a <a href="#">foreign exchange</a> term used for the <a href="#">GBP/USD currency pair</a> rate ( <a href="#">British pound</a> vs the <a href="#">US dollar</a> ).
Miss	n02934168	<b>Fauve, fauvist</b> (a member of a group of French	Fauve	<b>Fauve</b> collective, sometimes stylized

		painters who followed fauvism)	(collective)	as FAUVE, is a French arts collective of music and videography established in 2010 in <a href="#">Paris</a>
Miss	n10292316	<b>manufacturer</b> , <a href="#">producer</a> (someone who manufactures something)	Hip hop production	<b>Hip hop production</b> is the creation of <a href="#">hip hop music</a> . aspects of hip hop

Table 1. Manual inspection of Wordnet synsets/Wikipedia articles matching

Out of the 100 Wordnet synsets/Wikipedia articles pair randomly selected, we found 4 complete miss-associations and 2 slightly imperfect matches. All incorrect associations are reported in Table 1. The three last misses seem to be caused by homonyms while the reason for the first miss is less clear. The imperfect matches reveal an interesting biases: Wordnet defines semantic concepts in a formal description. Hence perihelion and aphelion are represented by two distinct resources in Wordnet. However, Wikipedia presents encyclopedic information for users to read. Hence, it is not constrained by Wordnet formalism and presents the related concepts of perihelion and aphelion in the same article. Regarding the second imperfect match, Wordnet synset refers to Roman buildings, i.e. building built following a Roman architecture while Wikipedia refers to the general concept of Roman architecture.

### B. DBPedia datasets

To generate the DBPedia semantic vectors, we used the following DBPedia datasets:

1. Page Links

Dataset containing internal links between DBpedia instances. The dataset was created from the internal links between Wikipedia articles

2. Instance types

Contains triples of the form dbp:ressource, rdf:type, dbp:type from the mapping-based extraction

3. Article categories

Gather links from concepts (DBpedia resources) to categories using the SKOS vocabulary.

4. Mapping based object

High-quality data extracted from Wikipedia pages Infoboxes using the mapping-based extraction (Object properties only).

### C. Model Architecture

Figure 2. represents the model architecture presented in the main article. It is made of two branches, the visual core model is represented in blue and the semantic branch is represented in red. The ZSL part, represented in gray combines both branches to perform classification of unseen classes.

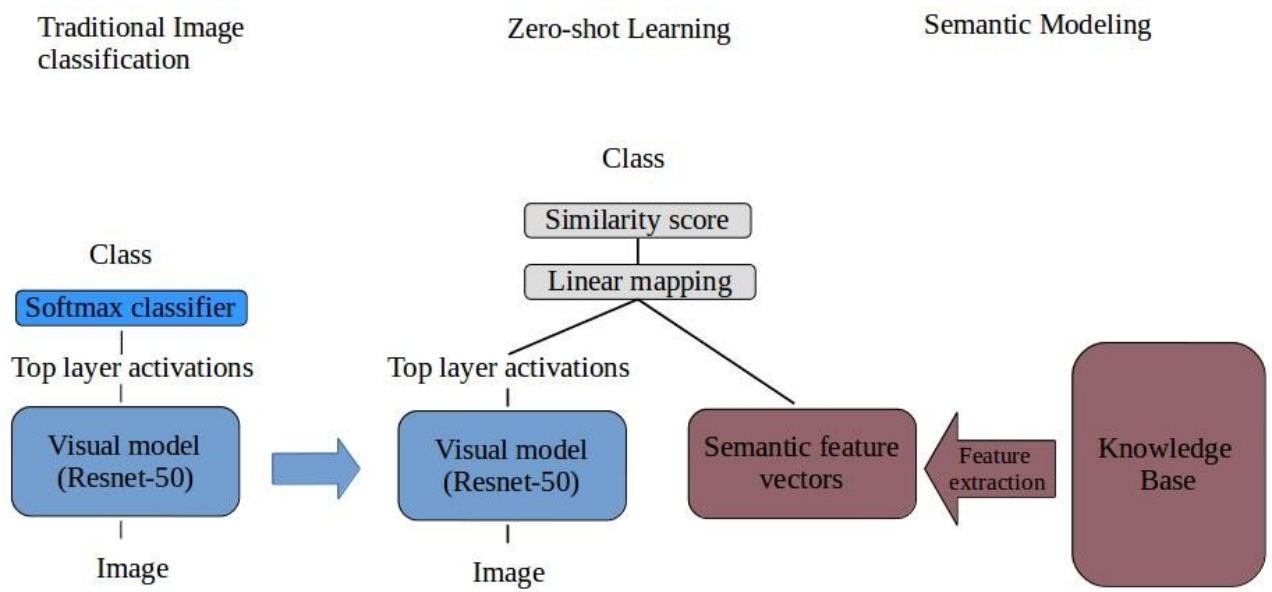


Figure 2. Architecture of our model