

Contents[Publications](#)[Downloads](#)[Wraetlic tools](#)[Instances in WordNet 1.7](#)[Similarity and](#)[Relatedness in](#)[WordSim353](#)

WordSim353 - Similarity and Relatedness

[WordSim353](#) is a test collection for measuring word similarity or relatedness, developed and maintained by E. Gabrilovich.

This page contains a split of the test set into two subsets, one for evaluating *similarity*, and the other for evaluating *relatedness*, according to the procedure described in the following paper:

Eneko Agirre, Enrique Alfonseca, Keith Hall, Jana Kravalova, Marius Pasca, Aitor Soroa, A Study on Similarity and Relatedness Using Distributional and WordNet-based Approaches, In *Proceedings of NAACL-HLT 2009*.

If you publish results based on this dataset, please reference this paper.

The dataset is available for download [here](#). It contains the following files:

- **wordsim353_annotator1.txt**: the classification of the pairs according to the first annotator.
 - **wordsim353_annotator2.txt**: the classification of the pairs according to the second annotator.
 - **wordsim353_agreed.txt**: the classification of the pairs after agreement was reached.
 - **wordsim_relatedness_goldstandard.txt**: the final goldstandard for measuring relatedness, in the same format as the WordSim353 distribution.
 - **wordsim_similarity_goldstandard.txt**: the final goldstandard for measuring similarity, in the same format as the WordSim353 distribution.
-