Word Sense Disambiguation (WSD) is a long-standing task in Natural Language Processing and Artificial Intelligence. While progress has been made in recent years, the evaluation of WSD models has been limited to a set of (mostly SemEval-based) standard datasets.

The SemDeep-6 workshop includes a challenge (shared task) based on a new multi-domain evaluation benchmark for WSD “Target Sense Verification for Words in Context” (WiC-TSV).

The main difference between WiC-TSV and common WSD task statement is that in WiC-TSV there is no standard sense inventory that systems need to model in full. Each instance in the dataset is associated with a target word and a sense, and therefore systems are not required to model all senses of the target word, but rather only a single sense. The task is to decide if the target word denotes the intended meaning associated with the corresponding sense or not; therefore, a binary classification task. The WiC-TSV task resembles the usage of automatic tagging in enterprise settings.

Training, development and test sets will be provided for the WiC-TSV challenge (training and development sets already available). For more information and instructions on how to participate, please visit <https://competitions.codalab.org/competitions/23683>.

Advantages of the new benchmark:

1. It provides a more realistic WSD setting in which a target ambiguous word is compared against its entry in a knowledge graph. Therefore, the task statement of WiC-TSV resembles the usage of enterprise knowledge graphs for entity linking.
2. The task is more targeted at word-level representation, as in one of the tasks (i.e. hypernymy task) the model is not provided with any contextual information and, therefore, needs to have a clear understanding of the word to be able to make correct judgements;
3. WiC-TSV as well as its predecessor [WiC](https://pilehvar.github.io/wic/) is also independent from external sense inventories and binary classification nature of the task.