

SemTab 2025

Semantic Web Challenge on Tabular Data to Knowledge Graph Matching

[Tracks](#)

[Join the forum](#) [↗](#)



About the Challenge

Tabular data in the form of **CSV files** is the common input format in a data analytics pipeline. However, a lack of understanding of the **semantic structure** and meaning of the content may hinder the data analytics process. Thus gaining this semantic understanding will be very valuable for **data integration**, **data cleaning**, **data mining**, **machine learning** and **knowledge discovery** tasks.

Tables on the Web may also be the source of highly valuable data. The addition of **semantic information** to Web tables may enhance a wide range of applications, such as **web search**, **question answering**, and **Knowledge Base (KB)** construction.

Tabular data to Knowledge Graph (KG) matching is the process of assigning semantic tags from KGs (e.g., **Wikidata** or **DBpedia**) to the elements of the table. This task however is often difficult in practice due to metadata (e.g., table and column names) being missing, incomplete or ambiguous.

The **SemTab challenge** aims at benchmarking systems dealing with the tabular data to KG matching problem, so as to facilitate their comparison on the same basis and the **reproducibility** of the results.

The 2025 edition of this challenge will be collocated with the [International Semantic Web Conference](#) [↗](#).

Challenge Tracks

MammoTab

TASK: CEA (Wikidata v. 20240720)

Participants will address the Semantic Table Interpretation challenges using a carefully selected subset of **670 tables** from the new version of the **MammoTab dataset** [↗](#), comprising a total of **84,907 cell annotations**. MammoTab is a large-scale benchmark designed to provide realistic and complex scenarios, including tables affected by typical challenges of web and Wikipedia data. The test set is not included in the dataset in order to preserve the impartiality of the final evaluation and to discourage ad-hoc solutions.

Only approaches based on Large Language Models are allowed, either:

- in fine-tuning settings, or
- using Retrieval-Augmented Generation strategies.

The evaluation will focus on the **Cell Entity Annotation (CEA)** task using the Wikidata KG (v. 20240720) [↗](#), but will also take into account the ability of the proposed approaches to effectively deal with the following key challenges:

Disambiguation	Homonymy	Alias resolution
Correctly linking ambiguous mentions to the intended entities.	Managing mentions referring to entities with identical or very similar names.	Recognizing entities referred by alternative names, acronyms, or nicknames.
NIL Detection	Noise Robustness	Collective Inference
Correctly identifying mentions that do not correspond to any entity in the Knowledge Graph.	Dealing with incomplete, noisy, or imprecise table contexts.	Leveraging inter-cell and inter-column signals to improve the consistency of annotations.

Participants are expected to demonstrate not only strong CEA performance, but also robustness and versatility across all these dimensions, which are critical for real-world table interpretation scenarios.

[Download](#)

secu-table

TASK: CEA, CTA, CPA (SEPSes)

TASK: CEA (Wikidata)

The secu-table dataset involved security data extracted from Common Vulnerability and Exposure (CVE) and Common Weakness Enumeration (CWE) data sources. The dataset is composed of **20% of tables without any errors** and **80% of tables containing errors** such as ambiguity, NIL, missing context, misspelt data. Secu-table dataset is composed of **1,554 tables**, divided into **78 tables** provided as ground truth and **1,478 tables** for testing.

The participant are invited to use **open source LLMs** to address the STI tasks which are: cell entity annotation (CEA), column type annotation (CTA), and column property annotation (CPA) using the **SEPSes Computer Security Knowledge Graph** [↗](#) and the CEA task using the **Wikidata KG** [↗](#).

The evaluation of the participants' results will consider the recall, precision, and F-scores. In addition to these scores, the participants are invited to provide the evaluation of the LLMs capabilities to make a prediction or to abstain (or to say "I don't know").

[Download](#)

Co-Chairs

Marco Cremaschi

University of Milan - Bicocca

[✉ marco.cremaschi@unimib.it](mailto:marco.cremaschi@unimib.it)

Fabio D'Adda

University of Milan - Bicocca

[✉ fabio.dadda@unimib.it](mailto:fabio.dadda@unimib.it)

Fidel Jiomekong Azanzi

University of Yaoundé, Cameroon

[✉ fidel.jiomekong@facsciences-uy1.cm](mailto:fidel.jiomekong@facsciences-uy1.cm)

Jean Petit Yvelos

University of Yaoundé, Cameroon

[✉ jeanpetityvelos@gmail.com](mailto:jeanpetityvelos@gmail.com)

Ernesto Jimenez-Ruiz

City St George's, University of London

[✉ ernesto.jimenez-ruiz@citystgeorges.ac.uk](mailto:ernesto.jimenez-ruiz@citystgeorges.ac.uk)

Oktie Hassanzadeh

IBM Research

[✉ hassanzadeh@us.ibm.com](mailto:hassanzadeh@us.ibm.com)

Acknowledgements

The challenge is currently supported by IBM Research and the ISWC 2025.

