

Results of SemTab 2021^{*}

Vincenzo Cutrona¹, Jiaoyan Chen², Vasilis Efthymiou³, Otkie Hassanzadeh⁴,
Ernesto Jiménez-Ruiz^{5,6}, Juan Sequeda⁷, Kavitha Srinivas⁴, Nora Abdelmageed⁸,
Madelon Hulsebos⁹, Daniela Oliveira¹⁰, and Catia Pesquita¹⁰

¹ SUPSI, Switzerland. vincenzo.cutrona@supsi.ch

² University of Oxford, UK. jiaoyan.chen@cs.ox.ac.uk

³ FORTH-ICS, Greece. vefthym@ics.forth.gr

⁴ IBM Research, USA. hassanzadeh@us.ibm.com, kavitha.srinivas@ibm.com

⁵ City, University of London, UK. ernesto.jimenez-ruiz@city.ac.uk

⁶ SIRIUS, University of Oslo, Norway. ernestoj@uio.no

⁷ data.world, US. juan@data.world

⁸ University of Jena, Germany. nora.abdelmageed@uni-jena.de

⁹ University of Amsterdam, The Netherlands. m.hulsebos@uva.nl

¹⁰ LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal.

dpoliveira@fc.ul.pt, clpesquita@fc.ul.pt

Abstract. SemTab 2021 was the third edition of the Semantic Web Challenge on Tabular Data to Knowledge Graph Matching, successfully collocated with the 20th International Semantic Web Conference (ISWC) and the 16th Ontology Matching (OM) Workshop. SemTab provides a common framework to conduct a systematic evaluation of state-of-the-art systems.

Keywords: Tabular data · Knowledge Graphs · Matching · SemTab · Semantic Web Challenge · Semantic Table Interpretation

1 Motivation

Data in tabular format are the most frequent input to data analytics pipeline, thanks to their high storage and processing efficiency. Also, the tabular format allows users to represent the information in a compacted way, by exploiting the clear data structure defined by rows and columns. However, such clear structure does not imply a clear understanding of the semantic structure (*e.g.*, relationships between columns), as well as the meaning of the content (*e.g.*, if data are about a specific topic). The lack of understanding hinders data analytics processes, requiring additional effort to properly understand the data first. Gaining the semantic understanding is valuable for many applications, including data cleaning, data mining, data integration, data analysis and machine learning, and knowledge discovery. For example, the semantic understanding can help in assessing what kind of transformations are more appropriate for a dataset, or which datasets can be integrated to enable new analytics (*e.g.*, marketing analysis) [10].

^{*} Copyright ©2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In addition to their efficiency, the huge availability of tabular data on the Web makes Web tables a valuable source to consider for data miners (*e.g.*, open data CSV files). Adding semantic information to Web tables is useful for a wide range of applications, including web search, question answering, and knowledge base construction.

Tabular data to Knowledge Graph (KG) matching is the process of clarifying the semantic meaning of a table by mapping its elements (*i.e.*, cells, columns, rows) to semantic tags (*i.e.*, entities, classes, properties) from KGs (*e.g.*, Wikidata, DBpedia). The task difficulty increases when table metadata (*e.g.*, table captions, table description, or column names) being missing, incomplete or ambiguous.

The tabular data to KG matching process is typically broken down into the following tasks: (*i*) cell to KG entity matching (CEA task), (*ii*) column to KG class matching (CTA task), and (*iii*) column pair to KG property matching (CPA task).

Over the last decade several approaches made advances in addressing one or several of above tasks, also constructing benchmark datasets ([18, 22, 17, 11]). The creation of SemTab¹ [15, 16] aimed at putting this significant amount of work into a common framework, enabling the systematic evaluation of state-of-the-art systems. The ambition is to make SemTab becoming the reference challenge in the Semantic Web community, in the same way the OAEI² is for the Ontology Matching community.³

2 The Challenge

The SemTab 2021 challenge has been organised into 3 different tracks: the **Accuracy Track**, which is the standard track proposed in previous editions; the **Usability Track**, a new track addressing the lack of publicly available, easy-to-use and generic solutions; and the **Applications Track**, which focuses on applications in real-world settings where the output of matching systems can contribute. The application track was also open to the submission of novel benchmark datasets.

2.1 Accuracy Track

The Accuracy Track included 3 rounds, running from June 30 to October 15. Different target KGs were used across rounds (see Table 1):

- DBpedia [3]: [http://downloads.dbpedia.org/wiki-archive/\(version 2016-10\)](http://downloads.dbpedia.org/wiki-archive/(version%202016-10))
- Wikidata [24]: <https://zenodo.org/record/6153449>
- Schema.org [12]: https://gittables.github.io/downloads/schema_20210528.pkl

The different rounds of SemTab 2021 have been organised to evaluate participating systems on different datasets with variable difficulty. All the rounds were run with the support of Alcrowd;⁴ SemTab 2021 also used the STILTool system [8, 5] for getting additional insights about the submitted solutions.

¹ <http://www.cs.ox.ac.uk/isg/challenges/sem-tab/>

² <http://oaei.ontologymatching.org/>

³ <http://ontologymatching.org/>

⁴ <https://www.aicrowd.com/>

Table 1: Datasets used across SemTab 2021 rounds.

	Rounds			Tasks			Target KGs		
	R1	R2	R3	CTA	CPA	CEA	DBpedia	Wikidata	Schema.org
2T [9]	✓			✓		✓	✓	✓	
BioTable [20]		✓		✓	✓	✓		✓	
AG [15]		✓	✓	✓	✓	✓		✓	
BiodivTab [2]			✓	✓		✓		✓	
GitTables [13]			✓	✓			✓		✓

Datasets The different datasets used to run SemTab 2021 rounds are reported in Table 1, with some statistics available in Table 2. All the datasets are available in Zenodo:

- Tough Tables (2T): a dataset featuring high-quality manually-curated tables with non-obviously linkable cells, *i.e.*, where values are ambiguous names, typos, and misspelled entity names. These challenges are particularly relevant for the annotation of structured legacy sources to existing KGs.
Link: <https://doi.org/10.5281/zenodo.6211551>
- BioTable: a dataset focused on molecular biology data covering different entities. It has the largest number of rows per table in the challenge.
Link: <https://doi.org/10.5281/zenodo.5606585>
- Automatically Generated (AG):⁵ a synthetic dataset with tables generated automatically by means of SPARQL queries. AG is the largest dataset used in SemTab.
Link: <https://zenodo.org/record/6154708>
- BiodivTab: a dataset with tables from real-world biodiversity research datasets. Original tables have been adapted for the SemTab challenge.
Link: <https://doi.org/10.5281/zenodo.5584180>
- GitTables: a large-scale corpus of relational tables extracted from CSV files in GitHub. The main purpose of this dataset is to facilitate learning table representation models and applications in *e.g.*, data management. A subset of tables has been curated for benchmarking column type detection methods in SemTab.
Link: <https://doi.org/10.5281/zenodo.5706316>

Table 3 shows the participation per round. Compared with previous editions, we had 11 participants (vs 28 in 2020) submitting to at least one round.⁶ We identified 6 core participants (vs 8 in 2020), which completed ~ 14 tasks on average (out of 17 tasks). Seven participants submitted a system paper to the challenge: *MTab* [19], *MAGIC* [23], *MantisTable V* [4], *JenTab* [1], *GBMTab* [26], *Kepler-aSI* [6], and *DAGOBAH* [14].

Evaluation measures As per the previous editions, systems have been evaluated on a single annotation for each provided target, for all the tasks; *i.e.*, in CEA, target cells are to be annotated with a single entity from the target KG; in CTA, target columns are to be annotated with a single type from the target KG (as fine-grained as possible).

⁵ In SemTab 2021, also referred to as *Hard Tables*.

⁶ Alcrowd leaderboard scores 23 participants because of test submissions.

Table 2: Statistics of the datasets in each SemTab 2021 round. For target values: W =Wikidata; D =DBpedia; S =Schema.org.

	AG		2T	BioTables	BiodivTab	GitTables
	Round 2	Round 3	Round 1	Round 2	Round 3	Round 3
Tables #	1,750.00	7,207.00	180.00	110.00	50.00	1,101.00
Avg. Rows # (total)	16.73	8.18	1,080.21	2,449.08	259.06	58.20
Avg. Cols # (total)	3.19	2.48	4.46	5.97	23.96	15.87
Avg. Rows # (target CEA)	16.73 _W	8.18 _W	1,080.19 _D 1,080.21 _W	2,449.08 _W	258.28 _W	
Avg. Cols # (target CEA)	1.65 _W	1.00 _W	3.00 _D 3.00 _W	5.97 _W	13.60 _W	
Avg. Cols # (target CTA)	1.25 _W	1.00 _W	3.00 _D 3.00 _W	5.97 _W	12.28 _W	3.08 _D 2.62 _S
Avg. Cols # (target CPA)	3.19 _W	2.48 _W		5.97 _W		

Table 3: Participation in the SemTab 2021 challenge.

	Round 1	Round 2		Round 3		
	2T	BioTable	AG	AG	BiodivTab	GitTables
CEA	5 _D 7 _W	6	6	5	5	-
CTA	3 _D 7 _W	7	6	6	6	4 _D 2 _S
CPA	-	6	6	5	-	-
Total	11	7	6	6	6	4

The evaluation measures for CEA, CPA and CTA (DBpedia and Schema.org) are the standard Precision, Recall and F1-score, as defined in Equation 1:

$$P = \frac{|\text{Correct Annotations}|}{|\text{System Annotations}|}, \quad R = \frac{|\text{Correct Annotations}|}{|\text{Target Annotations}|}, \quad F1 = \frac{2 \times P \times R}{P + R} \quad (1)$$

where target annotations refer to the target cells for CEA, the target columns for CTA, and the target column pairs for CPA. We consider an annotation as *correct* when it is included within the ground truth set (a target cell usually has multiple annotations in the ground truth, because of redirect and same-as links in KGs).

Given the fine-grained type hierarchy in Wikidata, we adopted approximations of Precision and Recall in the CTA evaluation. Approximations adapt their numerators to consider partially correct annotations, *i.e.*, annotations that are ancestors or descendants of the ground truth (GT) classes. The correctness score *cscore* of a CTA annotation α considers the distance between the annotation and the GT classes in the type hierarchy, and it is defined as

$$\text{cscore}(\alpha) = \begin{cases} 0.8^{d(\alpha)}, & \text{if } \alpha \text{ is in GT, or an ancestor of the GT, with } d(\alpha) \leq 5 \\ 0.7^{d(\alpha)}, & \text{if } \alpha \text{ is a descendant of the GT, with } d(\alpha) \leq 3 \\ 0, & \text{otherwise;} \end{cases} \quad (2)$$

Table 4: Average F1-score consider the 11 participating systems. We included MTab results for 2T from SemTab 2020.

	Round 1	Round 2		Round 3		
	2T	BioTable	AG	AG	BiodivTab	GitTables
CEA	0.51 _D 0.52 _W	0.82	0.91	0.90	0.41	-
CTA	0.35 _D 0.53 _W	0.78	0.91	0.80	0.23	0.04 _D 0.19 _S
CPA	-	0.88	0.96	0.95	-	-

where $d(\alpha)$ is the shortest distance to one of the GT classes (as for CEA, also CTA GT columns may have multiple classes). For example, $d(\alpha) = 0$ if α is a class in the ground truth ($cscore(\alpha) = 1$), and $d(\alpha) = 2$ if α is a grandchild of a class in the ground truth ($cscore(\alpha) = 0.49$). Types in the higher level(s) of the KG type hierarchy are not considered in the GT (e.g., Q35120 [entity] in Wikidata). Given the correctness score $cscore$, approximated Precision (AP), Recall (AR), and F1-score (AF1) for the CTA evaluation are as follows:

$$AP = \frac{\sum cscore(\alpha)}{|\text{System Annotations}|}, \quad AR = \frac{\sum cscore(\alpha)}{|\text{Target Annotations}|}, \quad AF1 = \frac{2 \times AP \times AR}{AP + AR} \quad (3)$$

Results Table 4 contains the average F1-score achieved by the 11 participating systems. The Tough Tables dataset still represent a challenge for almost all the systems, specially considering the fact that the the dataset is the same as in SemTab 2020. The BiodivTab and GitTables datasets brought additional complexity in Round 3, highlighting that real-world tables are challenging.

CEA task. Results for the CEA task are reported in Figure 1 for all the datasets. The Round 1 used the same 2T tables from last year edition,⁷ raising the difficulty bar at the very beginning. Most of the systems faced important challenges when dealing with 2T tables, with only 2 systems managing to achieve an F1-score over 0.8 and several of them participating in only one of the tasks. It is worth noting the work of the DAGOBAN team, which improved their system over the last year, being able to achieve higher scores on 2T this year. Starting from Round 2, systems have been evaluated on datasets never seen before. The AG datasets aimed at bringing new challenges in each round, and we can observe than only the best systems managed to maintain almost the same score on the two different versions of this dataset. Concerning bio-related datasets, performance in Round 2 were positive (slightly below 0.9 on average), confirming that tables with many rows (~2,500 on average) do not represent a problem for most of all the systems. Instead, the complexity brought by the (relatively small) tables in the BiodivTab dataset represented a new problem to solve, showing significantly reduced

⁷ The Wikidata targets have been updated to the current Wikidata live version.

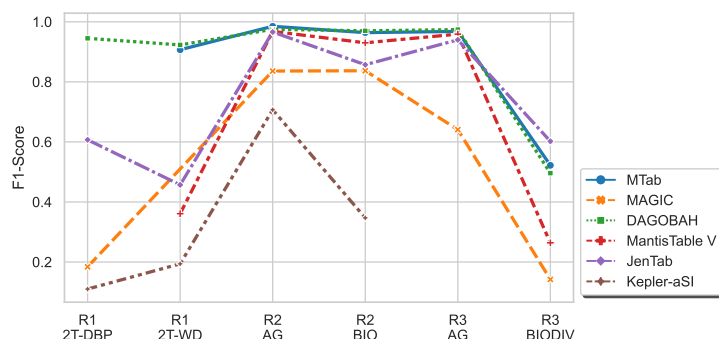


Fig. 1: Results in the CEA task for the core participants. MTab results on 2T are from 2020.

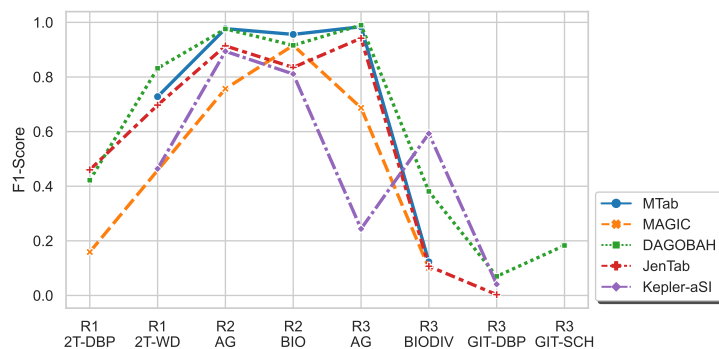


Fig. 2: Results in the CTA task for the core participants.

performance (none of the systems scored over 0.6). The JenTab system ranked 1st over a very difficult dataset. It is worth noting, however, that members of the JenTab team are also the providers of the BiodivTab dataset.

CTA task. As shown in Figure 2, the results in the CTA tasks resemble the trend already seen from the CEA results. This is an indicator that most of the systems solve the CTA tasks based on annotations found in the CEA. Additional challenges have been included in Round 3 with the GitTables dataset, where we can see a critical performance drop for all the involved systems. It is worth emphasising that, given the general picture provided by the results in CTA, more research is needed to make existing systems able to deal with real-world tables, where the cells may be missing a correspondence to the target KG.

CPA task. Results for the CPA tasks are plotted in Figure 3. Currently, only BioTables and the AG datasets provide a GT for CPA. Results are overall positive for all the tasks, with a general improvement from Round 2 to Round 3 for all the involved systems, except for MAGIC, whose performance dropped a bit during the last round.

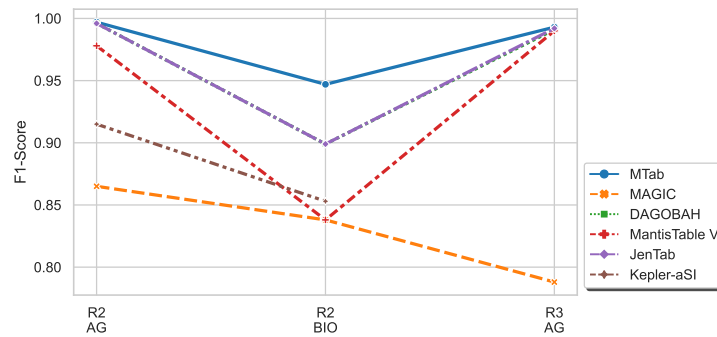


Fig. 3: Results in the CPA task for the core participants.

2.2 Usability Track

Starting from SemTab 2021, the organisation committee agreed to include a new track focusing on system usability. The main goal of this track is to mitigate a pain point in the community: the lack of publicly available, easy-to-use, and generic solution that will address the needs of a variety of applications and settings.

Evaluation measures Deeply evaluating the usability of a system requires user studies to monitor different parameters [21]. Within the SemTab scope, we decided to simply verify the overall usability of tools as judged by a review panel. Participants' solutions were examined for the following criteria:

- *Open source*: open-source solutions make a great contribution to the community, especially when released with a permissive license. Publicly available resources can be used as a starting point for new tools or research investigations, and make experiments easily reproducible.
- *System dependencies*: some tools may require specific platforms to be executed on premises, or have a huge resource consumption that may affect the use in common settings. For example, requiring many indexes/databases may prevent the usage of a tool by users with limited access to hardware.
- *Model generality*: a tool may be considered general when it applies to different (and new) applications/domains, requiring near-zero adaptations; for example, tools employing machine learning techniques should not require extensive training and tuning to be adapted to different contexts.
- *Availability*: tools may not be released as open source, but offered as a publicly available services. In this case, a tool served as a public service supports further research activities, and represent a big contribution to the community.
- *User experience*: the purpose of a tool is to help people in solving a task; for this reason, semantic table to graph matching tools should come with a well-designed user interface that makes the tool usable also by practitioner with a limited experience in semantic matching. That is, the tool should not require an extensive training to be mastered.

Table 5: Usability evaluation details.

	Open source	Availability as a Service	User Experience (GUI)
MTab		✓	✓
MAGIC	✓		✓
DAGOBAAH			✓
MantisTable V	✓		✓
JenTab	✓		
Kepler-aSI			

Results Almost all the core participants obtained good results in this track, by performing well on one or more of the above evaluation criteria. Evaluation details are reported in Table 5. We exclude system dependencies and model generality because of the insufficient available evidence, which resulted in these two criteria not impacting the overall assessment strongly. Indeed, available data about system performance (*i.e.*, accuracy) with reference to the different datasets and target KGs used in SemTab rounds do not allow us to draw any consistent conclusions. For example, it is not clear if tools were customized or tweaked (*e.g.*, changing the lookup function for noisy data) to increase their accuracy in different rounds; we are not able to assess how hard a system adapts to a different context (*e.g.*, changing the target KG).

The evaluation panel concluded that most of the tools are pre-configured and can potentially be used out of the box: for example, JenTab has been packaged in Docker containers to ease the deployment and execution of the tool on local premises. In general, tools requirements vary in complexity, but they are reasonable overall (*e.g.*, pre-processing required, like creating new indexes or embeddings).

Considering the other criteria, JenTab is the only system released as open source under a permissive license (Apache 2.0). The MTab tool has been made publicly available as a Web service, free to use (MIT license); but the back-end application has not been disclosed. However, having a public API enables MTab serving third-party application (with no rate limit), and this was a key point in declaring MTab the most usable tool. Systems like DAGOBAAH and MantisTable delivered a framework with impressive GUIs, while others (*e.g.*, MAGIC) opted for a lightweight application.

2.3 Applications Track

This new track aims at addressing applications in real-world settings that take advantage of the output of the matching systems. Challenging dataset proposals have also been accepted and included within the SemTab 2021 rounds.

Results A specific application has been identified within the biological domain, where new data are constantly produced thanks to the advances in the field. The domain is particularly challenging from the semantics standpoint because of the the complexity of the biological relations between entities. Within SemTab, the data representation significantly impact the systems performance since entities are usually represented by

codes (e.g., chemical formulas or gene names). Two different datasets have been submitted related to the biological domain; the first one, BioTables, is a dataset focused on molecular biology data; the second, BiodivTab, is a dataset focused on biodiversity research data and data augmentation.

Along side the above domain, a different dataset has been submitted to this track and also included in Round 3, GitTables. This dataset includes relational tables extracted from CSV files hosted at GitHub, and it comes with a peculiarity: the GT for CTA uses a mixture of classes and properties to annotate columns (both for the DBpedia and Schema.org versions).

The three datasets brought new complexity and contributed to increment the data diversity among the SemTab benchmark datasets.

2.4 Prizes

As in previous editions, IBM Research⁸ sponsored SemTab 2021 and awarded the best systems in each track with the following prizes:

- **Accuracy Track:** DAGOBAN (1st prize) was the top system in most of the tasks, showing appreciable improvements over the last years. Honorary mention to MTab
- **Usability Track:** MTab team (1st prize), for providing the easy-to-use MTab tool⁹ along with Web services to lookup entities and annotate tables; JenTab (2st prize), for being the only open-source system with a permissive license. Honorary mentions to DAGOBAN, MAGIC and MantisTable.
- **Applications Track:** BiodivTab dataset (1st prize), for having brought new challenges in CEA and CTA tasks. Honorary mention to GitTables.

3 Lessons Learned and Future Work

Avoiding over-fitting to AG. We have been using the same automated dataset generation process, with some variations that make it more challenging, since the first SemTab challenge. This may be resulting to participating systems that explicitly target datasets with characteristics similar to those of the AG datasets. This becomes evident from the almost perfect results shown in Table 4. For that reason, this year we have introduced several new datasets, while we are also planning to use as much as possible real data, rather than synthetic, in the future versions of the challenge.

System generalizability beyond KGs. Many systems currently rely on matching table values to entities in KGs. In this version of SemTab, we challenged the participating systems on their ability to detect the semantic types of table columns even when their values are not linkable to KG entities. We conclude that most systems do not generalize well in this scenario as indicated by the performance drop on the CTA task for GitTables (see Section 2.1). Improving systems to this end would make them useful for expanding KG coverage by matching tables from novel data sources to KGs in order to populate the

⁸ <https://www.research.ibm.com/>

⁹ <https://github.com/phucty/mtab.tool>

“unknown unknowns” [25]. This generalizability would also benefit the applicability of the systems in offline databases. We plan to encourage and evaluate systems on their generalizability towards novel data sources in future versions of SemTab.

CTA vs CPA: the case of GitTables. Since the first edition of SemTab, we are used to consider CTA and CPA as two separated tasks, the first focuses on ontology classes, and the latter is dedicated to properties. However, GitTables annotations for CTA includes also properties from DBpedia and Schema.org. The rationale behind this choice stands in the relational nature of the considered tables: columns typically correspond to the attributes of an entity, which are reflected by properties in DBpedia and Schema.org, for example. Also, this choice is very useful when annotating *literal columns* (i.e., columns not containing mentions of entities), avoiding annotations based on datatypes (e.g., xsd:string). Therefore, GitTables introduced a new technical challenge, which potentially contributed to the complexity observed from the results in Figure 2. The case of GitTables may result in a new task to accomplish in the future, given that it enables table-to-KG matching with tables from alternative data sources and contexts (e.g., database dumps from industry).

Usability track. We believe that the introduction of the usability track has contributed to making participating systems publicly accessible. Our goal was exactly to encourage this, despite the competitive nature that a challenge may have. Thus, we consider this new track to be a very important one and we are planning to keep it in the next challenges. Next SemTab editions may consider to improve the evaluation of this track, for example by adopting the *System Usability Scale* (SUS) [7] to score the overall user experience. In particular, developing a systematic way to evaluate systems’ generality and dependencies would definitely improve the evaluation of this track.

Applications track. We believe that the call of the application track has grasped more attention from the community by introducing their own datasets. Contributions from the community like BiodivTab, BioTable and GitTables help in extending the SemTab benchmark with new real-world challenges that are hard to reproduce in synthetic datasets as AG. Thus, this new track has been an important addition to SemTab.

Acknowledgements

We would like to thank the challenge participants, the ISWC & OM organisers, the Alcrowd team, and our sponsor IBM Research that played a key role in the success of SemTab. We also thank Paul Groth and Çağatay Demiralp for their contributions to GitTables. Moreover, we would like to thank Sirko Schindler and Birgitta König-Ries for their contribution to BiodivTab. This work was also supported by the SIRIUS Centre for Scalable Data Access (Research Council of Norway), Samsung Research UK, the EPSRC projects UK FIRES and ConCur, and the HFRI project ResponsibleER (No 969). DO and CP were supported by FCT through LASIGE (UIDB/00408/2020 and UIDP/00408/2020). We would also like to acknowledge that the work of the challenge organisers was greatly simplified by using the EasyChair conference management system and the CEUR-WS.org open-access publication service.

References

1. N. Abdelmageed and S. Schindler. JenTab Meets SemTab 2021's New Challenges. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
2. N. Abdelmageed, S. Schindler, and B. König-Ries. BiodivTab: A Tabular Benchmark based on Biodiversity Research Data. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
3. S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. DBpedia: A Nucleus for a Web of Open Data. In *The Semantic Web*, pages 722–735. Springer Berlin Heidelberg, 2007.
4. R. Avogadro and M. Cremaschi. MantisTable V: A novel and efficient approach to Semantic Table Interpretation. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
5. R. Avogadro, M. Cremaschi, E. Jiménez-Ruiz, and A. Rula. A Framework for Quality Assessment of Semantic Annotations of Tabular Data. In *20th International Semantic Web Conference (ISWC)*, pages 528–545, 2021.
6. W. Baazouzi, M. Kachroudi, and S. Faiz. Kepler-aSI at SemTab 2021. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
7. J. Brooke. SUS: a 'quick and dirty' usability scale. *Usability evaluation in industry*, 189(3), 1996.
8. M. Cremaschi, A. Siano, R. Avogadro, E. Jiménez-Ruiz, and A. Maurino. STILTool: A Semantic Table Interpretation evaluation Tool. In *ESWC 2020 Satellite Events*, pages 61–66, 2020.
9. V. Cutrona, F. Bianchi, E. Jiménez-Ruiz, and M. Palmonari. Tough Tables: Carefully Evaluating Entity Linking for Tabular Data. In *19th International Semantic Web Conference (ISWC)*, pages 328–343, 2020.
10. V. Cutrona, F. D. Paoli, A. Košmerlj, N. Nikolov, M. Palmonari, F. Perales, and D. Roman. Semantically-Enabled Optimization of Digital Marketing Campaigns. In *International Semantic Web Conference (ISWC)*, pages 345–362. Springer, 2019.
11. V. Efthymiou, O. Hassanzadeh, M. Rodriguez-Muro, and V. Christophides. Matching Web Tables with Knowledge Base Entities: From Entity Lookups to Entity Embeddings. In *ISWC*, volume 10587, pages 260–277. Springer, 2017.
12. R. V. Guha, D. Brickley, and S. Macbeth. Schema.Org: Evolution of Structured Data on the Web. *Commun. ACM*, 59(2):44–51, jan 2016.
13. M. Hulsebos, Ç. Demiralp, and P. Groth. GitTables: A Large-Scale Corpus of Relational Tables. *CoRR*, abs/2106.07258, 2021.
14. V.-P. Huynh, J. Liu, Y. Chabot, F. Deuzé, T. Labbé, P. Monnin, and R. Troncy. DAGOBAN: Table and Graph Contexts For Efficient Semantic Annotation Of Tabular Data. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
15. E. Jimenez-Ruiz, O. Hassanzadeh, V. Efthymiou, J. Chen, and K. Srinivas. SemTab 2019: Resources to Benchmark Tabular Data to Knowledge Graph Matching Systems. In *The Semantic Web: ESWC*. Springer International Publishing, 2020.
16. E. Jiménez-Ruiz, O. Hassanzadeh, V. Efthymiou, J. Chen, K. Srinivas, and V. Cutrona. Results of SemTab 2020. In *Proceedings of the Semantic Web Challenge on Tabular Data to Knowledge Graph Matching co-located with the 19th International Semantic Web Conference (ISWC 2020)*, pages 1–8, 2020.
17. O. Lehmborg, D. Ritze, R. Meusel, and C. Bizer. A large public corpus of web tables containing time and context metadata. In *WWW*, 2016.

18. G. Limaye, S. Sarawagi, and S. Chakrabarti. Annotating and searching web tables using entities, types and relationships. *VLDB Endowment*, 3(1-2):1338–1347, 2010.
19. P. Nguyen, I. Yamada, N. Kertkeidkachorn, R. Ichise, and H. Takeda. SemTab 2021: Tabular Data Annotation with MTab Tool. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
20. D. Oliveira and C. Pesquita. SemTab 2021 BioTable Dataset. doi:10.5281/zenodo.5606585, Oct. 2021.
21. C. Pesquita, V. Ivanova, S. Lohmann, and P. Lambrix. A framework to conduct and report on empirical user studies in semantic web contexts. In *European Knowledge Acquisition Workshop*, pages 567–583. Springer, 2018.
22. D. Ritze, O. Lehmborg, and C. Bizer. Matching HTML Tables to DBpedia. In *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics, WIMS*, pages 10:1–10:6. ACM, 2015.
23. B. Steenwinckel, F. D. Turck, and F. Ongenaes. MAGIC: Mining an Augmented Graph using INK, starting from a CSV. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.
24. D. Vrandečić and M. Krötzsch. Wikidata: a free collaborative knowledge base. *Commun. ACM*, 57(10):78–85, 2014.
25. G. Weikum. Knowledge Graphs 2021: A Data Odyssey. *Proc. VLDB Endow.*, 14(12):3233–3238, 2021.
26. L. Yang, S. Shen, J. Ding, and J. Jin. GBMTab: A Graph-Based Method for Interpreting Semantic Table to Knowledge Graph. In *Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab)*. CEUR-WS.org, 2021.