



SemTab 3.0: Semantic Web Challenge on Tabular Data to KG Matching

Oktie Hassanzadeh and **Kavitha Srinivas**, IBM Research, USA

Jiaoyan Chen, University of Oxford, UK

Vasilis Efthymiou, ICS-FORTH, Greece

Ernesto Jiménez-Ruiz, City, University of London, UK

Juan Sequeda, data.world, USA

Vincenzo Cutrona, University of Milano - Bicocca, Italy

Daniela Oliveira and **Catia Pesquita**, Universidade de Lisboa, Portugal

Nora Abdelmageed, University of Jena, Germany

Madelon Hulsebos, University of Amsterdam, Netherlands

SemTab Challenge Prizes

- Prizes sponsored by **IBM Research**

IBM Research

Applications Track Prize

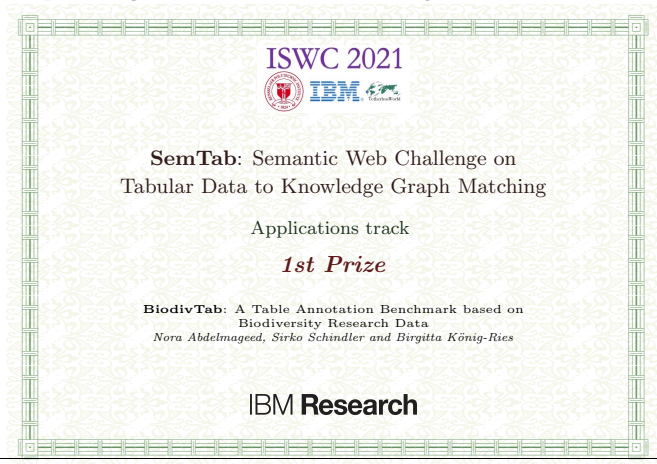
Applications Track Prize

- Honorary mention to **GitTables**

Madelon Hulsebos, Çağatay Demiralp and Paul Groth.
GitTables: A Large-Scale Corpus of Relational Tables.
<https://gittables.github.io/>

Applications Track Prize

- **BiodivTab**: <https://github.com/fusion-jena/BiodivTab>



Accuracy Prize

Accuracy Prize

- Honorary mention to **MTab**

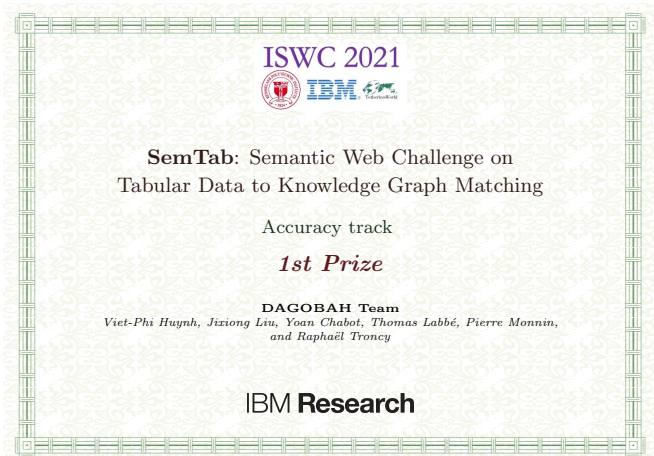
Phuc Nguyen, Ikuya Yamada, Natthawut Kertkeidkachorn,
Ryutaro Ichise and Hideaki Takeda.

SemTab 2021: Tabular Data Annotation with MTab Tool.

https://github.com/phucty/mtab_tool

Accuracy Track Prize

- **DAGOBDAH**: Table and Graph Contexts For Efficient Semantic Annotation Of Tabular Data



Usability Prize

Usability Prize

- Honorary mentions to **DAGOB AH**, **Magic** and **MantisTable**.
 - **DAGOB AH for the impressive GUI and framework.**
 - Viet-Phi Huynh, Jixiong Liu, Yoan Chabot, Frédéric Deuzé, Thomas Labbé, Pierre Monnin and Raphaël Troncy
 - **Magic for the lightweight app/GUI.**
 - Bram Steenwinckel, Filip De Turck and Femke Ongenaes
 - **MantisTable for an end-to-end framework including StilTool for evaluation.**
 - Roberto Avogadro and Marco Cremaschi

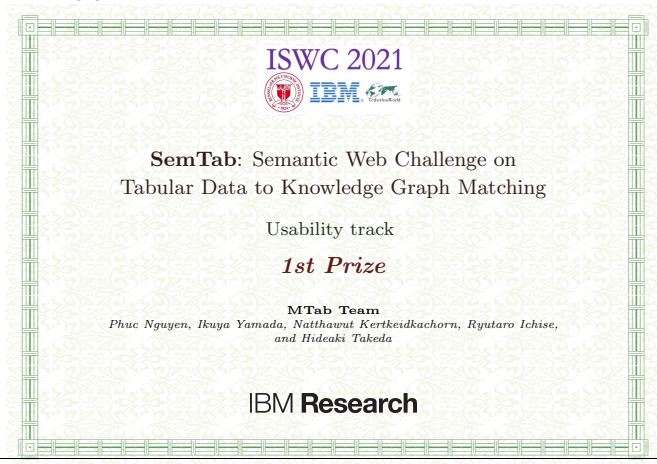
Usability Track (2nd Prize)

- **JenTab**: the only open-source participant with a friendly license



Usability Track (1st Prize)

- **MTab**: the mtab.app interface is the easiest to use overall



Conclusions

- SemTab targeted **3 KGs**: DBPedia, Wikidata and Schema.org
- **More dataset sources** this year (GitTables, BiodivTab, BioTables)
- The automatically generated datasets not challenging enough but a good baseline.
- **Tough tables** and **BiodivTab** are quite challenging
- **GitTables** very challenging.
 - Bring a new type of task combining CTA and CPA.
- **Usability** dimension considered for the SemTab 2021 prize.

Acknowledgements

- All participants
- Challenge organisers and their institutions
- AICrowd
- Our sponsor: IBM Research
- ISWC and OM organisers
- CEUR WS Proceedings

IBM Research



CITY
UNIVERSITY OF LONDON
EST 1894



INDE lab

ISWC 2021



Organisers



O. Hassanzadeh



J. Chen



V. Efthymiou



K. Srinivas



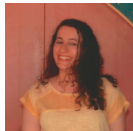
E. Jiménez-Ruiz



J. Sequeda



V. Cutrona



D. Oliveira



C. Pesquita



N. Abdelmageed



M. Hulsebos

Photo SemTab 2021

- Switch on your cameras!

Feedback and Discussion

- Comments on the rounds
- Proposals for next year?
 - New datasets/tasks
 - Systems as services
- **Extended discussion.** Poster sessions (wonder.me):
 - Today, Oct 27, 18:30-19:10 CET
 - Tomorrow, Oct 28, 15:00-15:30 CET