

Collaborative Project

Holistic Benchmarking of Big Linked Data

Project Number: 688227

Start Date of Project: 2015/12/01

Duration: 36 months

Deliverable 7.1.1 First Workshop Proceedings

Dissemination Level	Public
Due Date of Deliverable	Month 26, 31/01/2018
Actual Submission Date	Month 26, 31/01/2018
Work Package	WP7 - Evaluation Campaigns
Task	T7.2
Type	Report
Approval Status	Final
Version	1.0
Number of Pages	9
Filename	D7.1.1_First_Workshop_Proceedings.pdf

Abstract: This deliverable includes the workshop proceedings pertaining to the first series of evaluation campaigns organized by HOBBIT.

The information in this document reflects only the author's views and the European Commission is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688227.

History

Version	Date	Reason	Revised by
0.1	05/01/2018	First draft created	Grigorios Tzortzis & Vassiliki Rentoumi (NCSR-D)
0.2	07/01/2018	Peer Reviewed	Irini Fundulaki (FORTH)
0.3	12/01/2018	Content Revised	Grigorios Tzortzis & Vassiliki Rentoumi (NCSR-D)
1.0	15/01/2018	Final version created	Grigorios Tzortzis & Vassiliki Rentoumi (NCSR-D)

Author List

Organization	Name	Contact Information
NCSR-D	Grigorios Tzortzis	gtzortzi@iit.demokritos.gr
NCSR-D	Vassiliki Rentoumi	vrentoumi@iit.demokritos.gr

Executive Summary

This deliverable includes the workshop proceedings pertaining to the first series of evaluation campaigns (i.e. challenges) organized by HOBBIT. The proceedings are not reproduced in the deliverable due to copyright restrictions. Links to the websites where the proceedings volumes can be found are provided in this document.

HOBBIT has organized five benchmarking challenges in conjunction with established conferences in the Semantic Web domain:

- the Mighty Storage Challenge (MOCHA) at the European Semantic Web Conference (ESWC) 2017
- the Open Knowledge Extraction (OKE) challenge at the European Semantic Web Conference (ESWC) 2017
- the Question Answering over Linked Data (QALD) challenge at the European Semantic Web Conference (ESWC) 2017
- the DEBS Grand Challenge at the International Conference on Distributed and Event-Based Systems (DEBS) 2017
- the HOBBIT Link Discovery Task, as part of the OAEI 2017 campaign, at the Ontology Matching (OM) workshop at the International Semantic Web Conference (ISWC) 2017

Contents

Contents	3
1 Introduction	4
2 Proceedings of the MOCHA Challenge at ESWC 2017	4
3 Proceedings of the OKE Challenge at ESWC 2017	5
4 Proceedings of the QALD Challenge at ESWC 2017	6
5 Proceedings of the DEBS Grand Challenge at DEBS 2017	7
6 Proceedings of the HOBBIT Link Discovery Task at OAEI OM 2017 Workshop at ISWC 2017	8
7 Conclusions	9

1 Introduction

The main objectives of WP7 - Evaluation Campaigns are (i) to organize benchmarking campaigns in order to measure the fitness of implemented systems for processing Big Linked Data and (ii) to organize workshops in order to present and promote these systems and the results of the evaluation campaigns.

This deliverable reports on the proceedings of the workshops organized by HOBBIT to support the first series of the project's challenges (i.e. evaluation campaigns). Five challenges were arranged at renowned conferences, where participants were invited to submit systems that tackle the HOBBIT benchmarks. In particular, HOBBIT organized:

- the MOCHA challenge at ESWC 2017
- the OKE challenge at ESWC 2017
- the QALD challenge at ESWC 2017
- the DEBS Grand Challenge at DEBS 2017
- the HOBBIT Link Discovery Task at OAEI OM 2017 Workshop at ISWC 2017

Sections 2-6 provide a brief overview of each of the five challenges, followed by a list of the papers describing the systems that were accepted at each challenge. The proceedings are not reproduced in the deliverable due to copyright restrictions. Instead, links to the websites where the proceedings volumes can be found are given. Section 7 concludes this deliverable.

Additional information on the five HOBBIT challenges can be found on the project's website¹, as well as related deliverables D7.2.1 – First Workshop Organization Report, D7.3.1 – First Challenge Results Overview and D7.4.1 – First Challenge Evaluation. D7.2.1 reports on the organizational aspects of the challenges, D7.3.1 reports on the benchmarks, the challenges' tasks, and the participating systems and their results, and D7.4.1 reports on the quantitative and qualitative evaluation of the challenges.

2 Proceedings of the MOCHA Challenge at ESWC 2017

The Mighty Storage Challenge (MOCHA) was successfully held at ESWC 2017² and consisted of four tasks:

- Task 1: RDF data ingestion – This task measures how well systems can ingest streams of RDF data.
- Task 2: Data storage – This task measures how data stores perform with different types of queries.
- Task 3: Versioning – This task measures how well versioning and archiving systems for Linked Data perform when they store multiple versions of large datasets.
- Task 4: Browsing – This task checks existing solutions for how well they support applications that need browsing through large datasets.

¹<https://project-hobbit.eu/>

²<https://project-hobbit.eu/challenges/mighty-storage-challenge/>

.....

To support the tasks of the challenge, HOBBIT developed the Data Acquisition³, Data Storage⁴, Versioning⁵ and Faceted Browsing⁶ benchmarks.

The challenge proceedings consist of the following papers:

- Kleanthi Georgala, Mirko Spasic, Milos Jovanovik, Henning Petzka, Michael Roder and Axel-Cyrille Ngonga Ngomo, *MOCHA2017: The Mighty Storage Challenge at ESWC 2017*, pp. 3-15
- Alexander Potocki, Daniel Hladky and Martin Voigt, *Challenge Accepted: QUAD Meets MOCHA2017*, pp. 16-20
- Mirko Spasic and Milos Jovanovik, *MOCHA 2017 as a Challenge for Virtuoso*, pp. 21-32

The first paper was prepared by the MOCHA organizers and contains an overview of the challenge and baseline systems. The other papers present the systems that participated in the challenge (systems are described in D7.3.1 – First Challenge Results Overview).

The MOCHA challenge papers have been published by Springer on the proceedings volume *Dragoni M., Solanki M. and Blomqvist E. (eds), Semantic Web Challenges, Communications in Computer and Information Science, vol. 769, 2017*⁷. This volume contains the papers of all challenges that were organized at the ESWC 2017 conference.

3 Proceedings of the OKE Challenge at ESWC 2017

The Open Knowledge Extraction (OKE) challenge was successfully held at ESWC 2017⁸ and consisted of four tasks:

- Task 1: Focused Named Entity Identification and Linking
- Task 2: Broader Named Entity Identification and Linking
- Task 3: Focused Musical Named Entity Recognition and Linking
- Task 4: Knowledge Extraction

To support the tasks of the challenge, HOBBIT developed the Knowledge Extraction benchmark⁹.

The challenge proceedings consist of the following papers:

- Rene Speck, Michael Roder, Sergio Oramas, Luis Espinosa-Anke and Axel-Cyrille Ngonga Ngomo, *Open Knowledge Extraction Challenge 2017*, pp. 35-48
- Julien Plu, Raphael Troncy and Giuseppe Rizzo, *ADEL@OKE 2017: A Generic Method for Indexing Knowledge Bases for Entity Linking*, pp. 49-55

³<https://ckan.project-hobbit.eu/dataset/benchmark-for-sensor-data-odin>

⁴<https://ckan.project-hobbit.eu/dataset/data-storage-benchmark>

⁵<https://ckan.project-hobbit.eu/dataset/versioning-benchmark>

⁶<https://ckan.project-hobbit.eu/dataset/faceted-browsing-benchmark>

⁷<https://doi.org/10.1007/978-3-319-69146-6>

⁸<https://project-hobbit.eu/challenges/oke2017-challenge-eswc-2017/>

⁹<https://ckan.project-hobbit.eu/dataset/data-extraction-benchmark-for-unstructured-data>

.....

.....

The first paper has been prepared by the OKE organizers and contains an overview of the challenge and baseline systems, while the other papers present the systems that participated in the challenge (systems are described in D7.3.1 – First Challenge Results Overview).

The OKE challenge papers have been published by Springer on the proceedings volume *Dragoni M., Solanki M. and Blomqvist E. (eds), Semantic Web Challenges, Communications in Computer and Information Science, vol. 769, 2017*¹⁰. This volume contains the papers of all challenges that were organized at the ESWC 2017 conference.

4 Proceedings of the QALD Challenge at ESWC 2017

The Question Answering over Linked Data (QALD) challenge was successfully held at ESWC 2017¹¹ and consisted of four tasks:

- Task 1: Multilingual question answering over DBpedia
- Task 2: Hybrid question answering
- Task 3: Large-scale question answering over RDF
- Task 4: English question answering over Wikidata

To support the tasks of the challenge, HOBBIT developed the Question Answering benchmark¹².

The challenge proceedings consist of the following papers:

- Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Bastian Haarmann, Anastasia Krithara, Michael Roder and Giulio Napolitano, *7th Open Challenge on Question Answering over Linked Data (QALD-7)*, pp. 59-69
- Daniil Sorokin and Iryna Gurevych, *End-to-End Representation Learning for Question Answering with Weak Supervision*, pp. 70-83
- Dennis Diefenbach, Kamal Singh and Pierre Maret, *WDAqua-core0: A Question Answering Component for the Research Community*, pp. 84-89
- Nikolay Radoev, Mathieu Tremblay, Michel Gagnon and Amal Zouaq, *AMAL: Answering French Natural Language Questions using Dbpedia*, pp. 90-105

The first paper has been prepared by the QALD organizers and contains an overview of the challenge. The other papers present the systems that participated in the challenge (systems are described in D7.3.1 – First Challenge Results Overview).

The QALD challenge papers have been published by Springer on the proceedings volume *Dragoni M., Solanki M. and Blomqvist E. (eds), Semantic Web Challenges, Communications in Computer and Information Science, vol. 769, 2017*¹³. This volume contains the papers of all challenges that were organized at the ESWC 2017 conference.

¹⁰<https://doi.org/10.1007/978-3-319-69146-6>

¹¹<https://project-hobbit.eu/challenges/qald2017/>

¹²<https://ckan.project-hobbit.eu/dataset/question-answering-benchmark>

¹³<https://doi.org/10.1007/978-3-319-69146-6>

.....

5 Proceedings of the DEBS Grand Challenge at DEBS 2017

The ACM DEBS Grand Challenge is a popular challenge series organized by the DEBS conference. The 2017 edition of the DEBS Grand Challenge was organized in collaboration with the HOBBIT project¹⁴, where HOBBIT provided the dataset used for benchmarking the systems and the platform for the challenge. The challenge focused on the analysis of RDF streaming data generated by digital and analogue sensors embedded within manufacturing equipment. Specifically, the scenario focused on the the detection of anomalies in the behavior of the manufacturing equipment. To support the tasks of the challenge, HOBBIT developed the Structured Machine Learning benchmark¹⁵.

The challenge proceedings consist of the following papers:

- Vincenzo Gulisano, Zbigniew Jerzak, Roman Katerinenko, Martin Strohbach and Holger Ziekow, *The DEBS 2017 Grand Challenge*, pp. 271-273
- Nicolo Rivetti, Yann Busnel and Avigdor Gal, *FlinkMan: Anomaly Detection in Manufacturing Equipment with Apache Flink: Grand Challenge*, pp. 274-279
- Nihla Akram, Sachini Siriwardene, Malith Jayasinghe, Miyuru Dayarathna, Isuru Perera, Seshika Fernando, Srinath Perera, Upul Bandara and Sriskandarajah Suhothayan, *Anomaly Detection of Manufacturing Equipment via High Performance RDF Data Stream Processing: Grand Challenge*, pp. 280-285
- Ciprian Amariei, Paul Diac and Emanuel Onica, *Optimized Stage Processing for Anomaly Detection on Numerical Data Streams: Grand Challenge*, pp. 286-291
- Dimitrije Jankov, Sourav Sikdar, Rohan Mukherjee, Kia Teymourian and Chris Jermaine, *Real-time High Performance Anomaly Detection over Data Streams: Grand Challenge*, pp. 292-297
- Christian Mayer, Ruben Mayer and Majd Abdo, *StreamLearner: Distributed Incremental Machine Learning on Event Streams: Grand Challenge*, pp. 298-303
- Joong-Hyun Choi, Kang-Woo Lee, Hyungkun Jung and Eun-Sun Cho, *Runtime Anomaly Detection Method in Smart Factories using Machine Learning on RDF Event Streams: Grand Challenge*, pp. 304-309
- Tarek Zaarour, Niki Pavlopoulou, Souleiman Hasan, Umair ul Hassan and Edward Curry, *Automatic Anomaly Detection over Sliding Windows: Grand Challenge*, pp. 310-314

The first paper has been prepared by the challenge organizers and contains an overview of the challenge. The other papers present the systems that participated in the challenge (systems are described in D7.3.1 – First Challenge Results Overview).

The DEBS Grand Challenge papers have been published by ACM as part of the DEBS 2017 conference proceedings volume, *DEBS '17: Proceedings of the 11th ACM International Conference on Distributed and Event-based Systems*, ACM, New York, NY, USA, 2017¹⁶.

¹⁴<https://project-hobbit.eu/challenges/debs-grand-challenge/>

¹⁵<https://ckan.project-hobbit.eu/dataset/sml>

¹⁶<https://dl.acm.org/citation.cfm?id=3093742>

6 Proceedings of the HOBBIT Link Discovery Task at OAEI OM 2017 Workshop at ISWC 2017

As part of the OAEI (Ontology Alignment Evaluation Initiative) campaign at the 2017 Ontology Matching (OM) workshop which was held in conjunction with ISWC 2017, HOBBIT organized the Link Discover Task¹⁷. The Link Discovery Task consisted of two sub-tasks:

- Task 1: Linking – This task measures how well systems can match traces (GPS fixes) that have been altered using string-based approaches along with addition and deletion of intermediate points.
- Task 2: Spatial – This task measures how well systems can identify DE-9IM (Dimensionally Extended nine-Intersection Model) topological relations on GPS fixes.

To support the Link Discovery Task, HOBBIT developed the Link Discovery benchmark¹⁸.

In the OAEI campaign, each participating system was *optionally* accompanied by a non-peer reviewed “results” paper (systems are described in D7.3.1 – First Challenge Results Overview). Results papers serve the purpose of keeping track of the participants and the description of systems which took part in the campaign. These papers were published in the OM workshop’s website¹⁹.

In HOBBIT’s Link Discovery Task four systems participated in total and two of them published results papers:

- Daniel Faria, Booma S. Balasubramani, Vivek Shivaprabhu, Isabela Mott, Catia Pesquita, Francisco Couto and Isabel Cruz, *Results of AML in OAEI 2017*²⁰
- Abderrahmane Khiat and Maximilian Mackeprang, *I-Match and OntoIdea Results for OAEI 2017*²¹

The other two systems were published in previously held conferences:

- Mohamed Ahmed Sherif, Kevin Dreßler, Panayiotis Smeros, and Axel-Cyrille Ngonga Ngomo, *RADON - Rapid Discovery of Topological Relations*, 31st AAAI Conference on Artificial Intelligence (AAAI), 2017
- Panayiotis Smeros and Manolis Koubarakis, *Discovering Spatial and Temporal Links Among RDF Data*, Workshop on Linked Data on the Web (LDOW), 2016

Finally, an overview paper of OAEI has been prepared by the campaign organizers:

- Manel Achichi et al., *Results of the Ontology Alignment Evaluation Initiative 2017*²²

¹⁷<https://project-hobbit.eu/challenges/om2017/>

¹⁸<https://ckan.project-hobbit.eu/dataset/linkingbenchmark>

¹⁹<http://om2017.ontologymatching.org/>

²⁰http://www.dit.unitn.it/~pavel/om2017/papers/oaei17_paper2.pdf

²¹http://www.dit.unitn.it/~pavel/om2017/papers/oaei17_paper4.pdf

²²http://www.dit.unitn.it/~pavel/om2017/papers/oaei17_paper0.pdf

7 Conclusions

The HOBBIT project successfully organized its first series of evaluation campaigns which consisted of five challenges. The MOCHA, OKE and QALD challenges were organized in conjunction with the ESWC 2017 conference. Also, HOBBIT was responsible for the 2017 DEBS Grand Challenge that annually runs as part of the DEBS conference, as well as the Link Discovery Task at the 2017 OAEI campaign which was held under the Ontology Matching workshop at the ISWC 2017 conference. Participating systems were evaluated using the HOBBIT benchmarks and platform and described in papers published in proceedings volumes of the respective conferences.

Already, preparations for the organization of the second series of evaluation campaigns have started. HOBBIT will organize the MOCHA, OKE and SQA (Scalable Question Answering), an offspring of QALD, challenges at the ESWC 2018 conference. It will also contribute the Link Discovery task in OAEI 2017.5 challenge which will run at ESWC 2018. HOBBIT will again be responsible for the DEBS Grand Challenge in 2018 and plans to run the Link Discovery task as part of the 2018 OAEI campaign at ISWC 2018. Finally, HOBBIT has already launched the MOCHA, OKE, SQA and StreamML open challenges²³.

²³<https://project-hobbit.eu/open-challenges/>
