

Collaborative Project

## Holistic Benchmarking of Big Linked Data

Project Number: 688227

Start Date of Project: 2015/12/01

Duration: 36 months

# Deliverable 7.4 Second Workshop Proceedings

<b>Dissemination Level</b>	Public
<b>Due Date of Deliverable</b>	Month 36, 30/11/2018
<b>Actual Submission Date</b>	Month 36, 30/11/2018
<b>Work Package</b>	WP7 - Evaluation Campaigns
<b>Task</b>	T7.4
<b>Type</b>	Report
<b>Approval Status</b>	Final
<b>Version</b>	1.0
<b>Number of Pages</b>	9
<b>Filename</b>	D7.4_Second_Workshop_Proceedings.pdf

**Abstract:** This deliverable includes the workshop proceedings pertaining to the second 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	06/11/2018	First draft created	Georgios Katsimpras & Vassiliki Rentoumi (NCSR-D)
0.2	13/11/2018	Peer Reviewed	Carolin Walter (AGT)
0.3	14/11/2018	Content Revised	Georgios Katsimpras & Vassiliki Rentoumi (NCSR-D)
1.0	15/11/2018	Final version created	Georgios Katsimpras & Vassiliki Rentoumi (NCSR-D)

## Author List

Organization	Name	Contact Information
NCSR-D	Georgios Katsimpras	gkatsibras@iit.demokritos.gr
NCSR-D	Vassiliki Rentoumi	vrentoumi@iit.demokritos.gr

---

## Executive Summary

This deliverable includes the workshop proceedings pertaining to the second 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) 2018
- the Open Knowledge Extraction (OKE) challenge at the European Semantic Web Conference (ESWC) 2018
- the Scalable Question Answering (SQA) challenge at the European Semantic Web Conference (ESWC) 2018
- the DEBS Grand Challenge at the International Conference on Distributed and Event-Based Systems (DEBS) 2018
- the HOBBIT Link Discovery Task, as part of the OAEI 2018 campaign, at the Ontology Matching (OM) workshop at the International Semantic Web Conference (ISWC) 2018

---

## Contents

<b>Contents</b>	<b>3</b>
<b>1 Introduction</b>	<b>4</b>
<b>2 Proceedings of the MOCHA Challenge at ESWC 2018</b>	<b>4</b>
<b>3 Proceedings of the OKE Challenge at ESWC 2018</b>	<b>5</b>
<b>4 Proceedings of the SQA 2018 Workshop at ESWC 2018</b>	<b>6</b>
<b>5 Proceedings of the DEBS Grand Challenge at DEBS 2018</b>	<b>7</b>
<b>6 Proceedings of the HOBBIT Link Discovery Task at OAEI OM 2018 Workshop at ISWC 2018</b>	<b>8</b>
<b>7 Conclusions</b>	<b>9</b>

---

## 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 second 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 2018
- the OKE challenge at ESWC 2018
- the SQA challenge at ESWC 2018
- the DEBS Grand Challenge at DEBS 2018
- the HOBBIT Link Discovery Task at OAEI OM 2018 Workshop at ISWC 2018

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 website<sup>1</sup>, as well as related deliverables D7.2.2 – Second Workshop Organization Report, D7.3.2 – Second Challenge Results Overview and D7.4.2 – Second Challenge Evaluation. D7.2.2 reports on the organizational aspects of the challenges, D7.3.2 reports on the benchmarks, the challenges' tasks, and the participating systems and their results, and D7.4.2 reports on the quantitative and qualitative evaluation of the challenges.

## 2 Proceedings of the MOCHA Challenge at ESWC 2018

The Mighty Storage Challenge (MOCHA) was successfully held at ESWC 2018<sup>2</sup> 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.

---

<sup>1</sup><https://project-hobbit.eu/>

<sup>2</sup><https://project-hobbit.eu/challenges/mighty-storage-challenge2018/>

- 
- Task 4: Browsing – This task checks existing solutions for how well they support applications that need browsing through large datasets.

To support the tasks of the challenge, HOBBIT developed the Sensor Streams<sup>3</sup>, Data Storage<sup>4</sup>, Versioning<sup>5</sup> and Faceted Browsing<sup>6</sup> benchmarks.

The challenge proceedings consist of the following papers:

- Kleanthi Georgala, Mirko Spasic, Milos Jovanovik, Vassilis Papakonstantinou, Claus Stadler, Michael Röder, and Axel-Cyrille Ngonga Ngomo, *MOCHA2018: The Mighty Storage Challenge at ESWC 2018*, pp. 3-16
- Ruben Taelman, Miel Vander Sande and Ruben Verborgh, *Versioned Querying with OSTRICH and Comunica in MOCHA 2018*, pp. 17-23
- Milos Jovanovik and Mirko Spasic, *Benchmarking Virtuoso 8 at the Mighty Storage Challenge 2018: Training Results*, pp. 24-35

The papers present the systems that participated in the challenge (systems are described in D7.3.2 – Second Challenge Results Overview).

The MOCHA challenge papers have been published by Springer on the proceedings volume *Davide Buscaldi, Aldo Gangemi, Diego Reforgiato Recupero, Semantic Web Challenges, Communications in Computer and Information Science, vol. 927, 2018*<sup>7</sup>. This volume contains the papers of all challenges that were organized at the ESWC 2018 conference.

### 3 Proceedings of the OKE Challenge at ESWC 2018

The Open Knowledge Extraction (OKE) challenge was successfully held at ESWC 2018<sup>8</sup> and consisted of four tasks:

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

To support the tasks of the challenge, HOBBIT developed the Unstructured Sensor Streams benchmark<sup>9</sup>.

The challenge proceedings consist of the following papers:

---

<sup>3</sup>[https://project-hobbit.eu/wp-content/uploads/2018/06/D3.1.2\\_Data\\_Extraction\\_Benchmark\\_for\\_Sensor\\_Data.pdf](https://project-hobbit.eu/wp-content/uploads/2018/06/D3.1.2_Data_Extraction_Benchmark_for_Sensor_Data.pdf)

<sup>4</sup>[https://project-hobbit.eu/wp-content/uploads/2018/06/D5.1.2\\_Second\\_Version\\_of\\_the\\_Data\\_Storage.pdf](https://project-hobbit.eu/wp-content/uploads/2018/06/D5.1.2_Second_Version_of_the_Data_Storage.pdf)

<sup>5</sup>[https://project-hobbit.eu/wp-content/uploads/2018/06/D5.2.2\\_Second\\_version\\_of\\_the\\_Versioning\\_Benchmark.pdf](https://project-hobbit.eu/wp-content/uploads/2018/06/D5.2.2_Second_version_of_the_Versioning_Benchmark.pdf)

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

<sup>7</sup><https://www.springer.com/us/book/9783030000714>

<sup>8</sup><https://project-hobbit.eu/challenges/oke2018-challenge-eswc-2018/>

<sup>9</sup>[https://project-hobbit.eu/wp-content/uploads/2018/06/D3.2.2\\_Second\\_Version\\_of\\_the\\_Data\\_Extraction\\_Benchmark\\_for\\_unstructured\\_data.pdf](https://project-hobbit.eu/wp-content/uploads/2018/06/D3.2.2_Second_Version_of_the_Data_Extraction_Benchmark_for_unstructured_data.pdf)

---

- 
- Rene Speck, Michael Röder, Felix Conrads, Hyndavi Rebba, Catherine Camilla Romiyo, Gurudevi Salakki, Rutuja Suryawanshi, Danish Ahmed, Nikit Srivastava, Mohit Mahajan, and Axel-Cyrille Ngonga Ngomo, *Open Knowledge Extraction Challenge 2018*, pp. 39-51
  - Héctor Cerezo-Costas and Manuela Martín -Vicente, *Relation Extraction for Knowledge Base Completion: A supervised approach*, pp. 52-66

The papers present the systems that participated in the challenge (systems are described in D 7.3.2 – Second Challenge Results Overview).

The OKE challenge papers have been published by Springer on the proceedings volume *Davide Buscaldi, Aldo Gangemi, Diego Reforgiato Recupero, Semantic Web Challenges, Communications in Computer and Information Science, vol. 927, 2018*<sup>7</sup>. This volume contains the papers of all challenges that were organized at the ESWC 2018 conference.

## 4 Proceedings of the SQA 2018 Workshop at ESWC 2018

The key challenge for Scalable Question Answering over Linked Data is the need to translate a user's information request into such a form that it can be efficiently evaluated using standard Semantic Web query processing and inferencing techniques. Therefore, the main task of SQA Challenge was the following: given an RDF dataset and a large volume of natural language questions or keywords, return the correct answers (or SPARQL queries that retrieves those answers).

To support the SQA2018 task, HOBBIT developed the second version of the Question Answering benchmark<sup>7</sup>.

The challenge proceedings consist of the following papers:

- Giulio Napolitano, Ricardo Usbeck, and Axel-Cyrille Ngonga Ngomo, *The Scalable Question Answering over Linked Data (SQA) Challenge 2018*, pp. 69-75
- Dennis Diefenbach, Kamal Singh and Pierre Maret, *On the scalability of the QA system WDAqua-core1*, pp. 76-81
- Elizaveta Zimina, Jyrki Nummenmaa, Kalervo Jarvelin, Jaakko Peltonen, Kostas Stefanidis and Heikki Hyyro, *GQA: Grammatical Question Answering for RDF Data*, pp. 82-97
- Nikolay Radoev, Mathieu Tremblay, Amal Zouaq and Michel Gagnon, *LAMA: a Language Adaptive Method for Question Answering*, pp. 98-117

The papers present the systems that participated in the challenge (systems are described in D7.3.2 – Second Challenge Results Overview).

The SQA 2018 challenge papers have been published by Springer on the proceedings volume *Davide Buscaldi, Aldo Gangemi, Diego Reforgiato Recupero, Semantic Web Challenges, Communications in Computer and Information Science, vol. 927, 2018*<sup>7</sup>. This volume contains the papers of all challenges that were organized at the ESWC 2018 conference.

## 5 Proceedings of the DEBS Grand Challenge at DEBS 2018

The ACM DEBS Grand Challenge is a popular challenge series organized by the DEBS conference. The 2018 edition of the DEBS Grand Challenge was organized in collaboration with the HOBBIT project<sup>10</sup>, where HOBBIT provided the platform for the challenge and data was provided by the MarineTraffic company. The challenge focused on the application of machine learning to spatio-temporal streaming data. The goal of the challenge is to make the naval transportation industry more reliable by providing predictions for vessels's destinations and arrival times. Predicting both correct destinations and arrival times of vessels are relevant problems, that once solved, will boost the efficiency of the overall supply chain management. To support the tasks of the challenge, HOBBIT developed the Structured Machine Learning benchmark<sup>11</sup>.

The challenge proceedings consist of the following papers:

- Vincenzo Gulisano, Zbigniew Jerzak, Pavel Smirnov, Martin Strohbach, Holger Ziekow, Dimitris Zissis, *The DEBS 2018 Grand Challenge*, pp. 191-194
- Ciprian Amariei, Paul Diac, Emanuel Onica and Valentin Rosca, *Grand Challenge: Cell Grid Architecture for Maritime Route Prediction on AIS Data Streams*, pp. 202-204
- Moti Bachar, Gal Elimelech, Itai Gat, Gil Sobol, Nicolo Rivetti and Avigdor Gal. *Grand Challenge: Venelia, On-line Learning and Prediction of Vessel Destination*, pp. 209-212
- Abderrahmen Kammoun, Tanguy Raynaud, Syed Gillani, Kamal Singh, Jacques Fayolle and Frederique Laforest. *Grand Challenge: A Scalable Framework for Accelerating Situation Prediction over Spatio-temporal Event Streams*, pp. 221-223
- Duc-Duy Nguyen, Chan Le Van and Muhammad Intizar Ali. *Vessel Destination and Arrival Time Prediction using Sequence-to-Sequence Models over Spatial Grid*, pp. 217-220
- Hyungkun Jung, Kang-Woo Lee, Joong-Hyun Choi and Eun-Sun Cho. *Grand Challenge: Bayesian Estimation on Destination Ports and Arrival times of Vessels*, pp. 195-197
- Florian Schmidt, Oleh Bodunov, Andre Martin, Andrey Brito and Christof Fetzer, *Grand Challenge: Real-time Destination and ETA Prediction for Maritime Traffic*, pp. 198-201
- Chun-Xun Lin, Tsung-Wei Huang, Guannan Guo and Martin Wong, *MtDetector: A High-Performance Marine Traffic Detector at Stream Scale*, pp. 205-208
- Rim Moussa. *Scalable Maritime Traffic Map Inference and Real-time Prediction of Vessels' Future Locations on Apache Spark*, pp. 213-216
- Valentin Rosca, Emanuel Onica, Paul Diac and Ciprian Amariei, *Grand Challenge: Predicting Destinations by Nearest Neighbor Search on Training Vessel Routes*, pp. 224-225

The papers present the systems that participated in the challenge (systems are described in D7.3.2 – Second Challenge Results Overview).

The DEBS Grand Challenge papers have been published by ACM as part of the DEBS 2018 conference proceedings volume, *DEBS '18: Proceedings of the 12th ACM International Conference on Distributed and Event-based Systems*, ACM, Hamilton, NY, New Zealand, 2018<sup>12</sup>.

<sup>10</sup><https://project-hobbit.eu/challenges/debs2018-grand-challenge/>

<sup>11</sup>[https://project-hobbit.eu/wp-content/uploads/2018/06/D4.2.2\\_Second\\_Version\\_of\\_the\\_Data\\_Analytics\\_Benchmark.pdf](https://project-hobbit.eu/wp-content/uploads/2018/06/D4.2.2_Second_Version_of_the_Data_Analytics_Benchmark.pdf)

<sup>12</sup><https://dl.acm.org/citation.cfm?id=3210284>



---

## 6 Proceedings of the HOBBIT Link Discovery Task at OAEI OM 2018 Workshop at ISWC 2018

As part of the OAEI (Ontology Alignment Evaluation Initiative) campaign at the 2018 Ontology Matching (OM) workshop which was held in conjunction with ISWC 2018, HOBBIT organized the Link Discovery Task<sup>13</sup>. 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 benchmark<sup>14</sup>.

Further in OAEI 2018 the HOBBIT platform has been used for evaluation together with SEALS. The transition from SEALS platform to the HOBBIT platform has been supported by the OAEI 2017.5 pre-campaign. The experiences of this pre-campaign are discussed in the following paper: *Ernesto Jimerez-Ruiz et al, Introducing the HOBBIT platform into the ontology alignment evaluation campaign*<sup>15</sup> which has been published on the OM 2018 workshop website<sup>16</sup>.

In the OAEI 2018 campaign, each participating system was *optionally* accompanied by a non-peer reviewed “results” paper. 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 website<sup>16</sup>.

In HOBBIT Link Discovery Task three systems (*AML*, *RADON*, *SILK*) participated in total and one of them (i.e. *AML*) published a results’ paper on the OM 2018 workshop website:

- *AML*: Daniel Faria, Catia Pesquita, Booma S. Balasubramani, Teemo Tervo, David Carrio, Rodrigo Garilha, Francisco Couto and Isabel Cruz, *Results of AML participation in OAEI 2018*<sup>17</sup>

The other two systems *RADON*, *SILK* were published in previously held conferences:

- *RADON*: 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
- *SILK*: 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 2018 has been prepared by the campaign organizers:

- Pavel Shvaiko et al., *Results of the Ontology Alignment Evaluation Initiative 2018*<sup>18</sup>

---

<sup>13</sup><https://project-hobbit.eu/challenges/om2018/>

<sup>14</sup><https://ckan.project-hobbit.eu/dataset/linkingbenchmark>

<sup>15</sup>[http://disi.unitn.it/~pavel/om2018/papers/om2018\\_LTpaper5.pdf](http://disi.unitn.it/~pavel/om2018/papers/om2018_LTpaper5.pdf)

<sup>16</sup><http://om2018.ontologymatching.org/>

<sup>17</sup>[http://www.dit.unitn.it/~pavel/om2018/papers/oaei18\\_paper2.pdf](http://www.dit.unitn.it/~pavel/om2018/papers/oaei18_paper2.pdf)

<sup>18</sup><http://oaei.ontologymatching.org/2018/results/index.html>

---

---

## 7 Conclusions

The HOBBIT project successfully organized its second series of evaluation campaigns which consisted of five challenges. The MOCHA, OKE and SQA challenges were organized in conjunction with the ESWC 2018 conference. Also, HOBBIT was responsible for the 2018 DEBS Grand Challenge that annually runs as part of the DEBS conference, as well as the Link Discovery Task at the 2018 OAEI campaign which was held under the Ontology Matching workshop at the ISWC 2018 conference. Participating systems were evaluated using the HOBBIT benchmarks and platform and described in papers published in proceedings volumes of the respective conferences.