

Collaborative Project

## Holistic Benchmarking of Big Linked Data

Project Number: 688227

Start Date of Project: 2015/12/01

Duration: 36 months

# Deliverable 7.1.3 Second Preparations Report

<b>Dissemination Level</b>	Public
<b>Due Date of Deliverable</b>	Month 30, 31/05/2018
<b>Actual Submission Date</b>	Month 30, 31/05/2018
<b>Work Package</b>	WP7 - Evaluation Campaigns
<b>Task</b>	T7.1 & T7.2
<b>Type</b>	Report
<b>Approval Status</b>	Final
<b>Version</b>	1.0
<b>Number of Pages</b>	19
<b>Filename</b>	D7.1.3_Second_Preparations_Report.pdf

**Abstract:** This deliverable presents the state of the preparations for the organization of the second series of evaluation campaigns and corresponding workshops by the HOBBIT project.

---

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	09/05/2018	First draft created	Vassiliki Rentoumi & Grigorios Tzortzis (NCSR-D)
0.2	16/05/2018	Peer Reviewed	Filipe Teixeira (TomTom)
0.3	21/05/2018	Content Revised	Vassiliki Rentoumi & Grigorios Tzortzis (NCSR-D)
1.0	22/05/2018	Final version created	Vassiliki Rentoumi & Grigorios Tzortzis (NCSR-D)

## Author List

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

---

## Executive Summary

This deliverable presents the state of the preparations for the organization of the second series of evaluation campaigns and the second series of workshops by the HOBBIT project.

HOBBIT is organizing five benchmarking challenges (i.e. evaluation campaigns) which will evaluate participating systems using the eight benchmarks developed by HOBBIT over the HOBBIT platform. Each challenge is accompanied by a corresponding workshop, where the challenge results will be presented to the public. All challenges are organized in conjunction with renowned conferences in the Semantic Web domain. In 2018, HOBBIT has booked the following challenges:

- 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 ACM International Conference on Distributed and Event-Based Systems (DEBS 2018)
- the HOBBIT Link Discovery Task as part of the Ontology Alignment Evaluation Initiative (OAEI) at the International Semantic Web Conference (ISWC 2018)

The first four challenges will run in June 2018, while the last challenge will take place in October 2018.

In addition, HOBBIT launched four open challenges (Open MOCHA, Open OKE, Open SQA and StreamML open challenge) in the last quarter of 2017, where potential participants could submit their systems and join the challenge at virtually any point in time, with periodic cutoffs regulating the announcement of winners. The open challenges aimed at i) boosting participation through the absence of hard deadlines and ii) starting the dissemination of MOCHA, OKE and SQA @ ESWC 2018, as well as DEBS Grand Challenge @ DEBS 2018, early. These open challenges have been terminated as no systems were submitted in time for the first cutoff date and interested teams were directed to the “non-open” counterparts of the challenges.

---

## Contents

<b>Contents</b>	<b>3</b>
<b>List of Tables</b>	<b>5</b>
<b>1 Introduction</b>	<b>6</b>
<b>2 Benchmarks Description</b>	<b>6</b>
<b>3 MOCHA Challenge @ ESWC 2018</b>	<b>7</b>
3.1 Challenge Motivation . . . . .	7
3.2 Challenge Overview . . . . .	8
3.3 Challenge Tasks . . . . .	8
3.4 Open MOCHA Challenge . . . . .	9
<b>4 OKE Challenge @ ESWC 2018</b>	<b>10</b>
4.1 Challenge Motivation . . . . .	10
4.2 Challenge Overview . . . . .	10
4.3 Challenge Tasks . . . . .	11
4.4 Open OKE Challenge . . . . .	11
<b>5 SQA Challenge @ ESWC 2018</b>	<b>12</b>
5.1 Challenge Motivation . . . . .	12
5.2 Challenge Overview . . . . .	12
5.3 Challenge Tasks . . . . .	13
5.4 Open SQA Challenge . . . . .	13
<b>6 Grand Challenge @ DEBS 2018</b>	<b>14</b>
6.1 Challenge Motivation . . . . .	14
6.2 Challenge Overview . . . . .	14
6.3 Challenge Tasks . . . . .	14
6.4 StreamML Open Challenge . . . . .	15
<b>7 HOBBIT Link Discovery Task at OAEI 2017.5/2018 at ISWC 2018</b>	<b>15</b>
7.1 Challenge Motivation . . . . .	15
7.2 Challenge Overview . . . . .	16

---

7.3	Challenge Tasks . . . . .	16
<b>8</b>	<b>Milestones for the Organization of Challenges</b>	<b>17</b>
8.1	Current Status of Challenges . . . . .	18

---

## List of Tables

1	Overview of HOBBIT Benchmarks. . . . .	7
2	Mapping between HOBBIT Challenges/Benchmarks. . . . .	8

---

## 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. To achieve these goals, HOBBIT has defined eight benchmarks and is currently organizing five benchmarking challenges (i.e. evaluation campaigns) at renowned conferences, where participants are invited to submit systems that tackle the HOBBIT benchmarks. This year's challenges are the continuation of the 2017 challenges, where an updated (second) version of the benchmarks will be employed. The same venues as in 2017 were targeted to host the challenges.

Specifically, benchmarks about (i) the acquisition of data, (ii) knowledge extraction, (iii) link discovery, (iv) structured machine learning, (v) RDF data storage, (vi) data versioning, (vii) question answering and (viii) faceted browsing, have been developed. HOBBIT has booked the following challenges, covering all eight benchmarks:

- 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 as part of the Ontology Alignment Evaluation Initiative (OAEI) at ISWC 2018

In addition, HOBBIT launched four open challenges (Open MOCHA, Open OKE, Open SQA and StreamML open challenge) in the last quarter of 2017, where potential participants could submit their systems and join the challenge at virtually any point in time, with periodic cutoffs regulating the announcement of winners. The open challenges aimed at i) boosting participation through the absence of hard deadlines and ii) starting the dissemination of MOCHA, OKE and SQA @ ESWC 2018, as well as DEBS Grand Challenge @ DEBS 2018, early. These open challenges have been terminated as no systems were submitted in time for the first cutoff date and interested teams were directed to the “non-open” counterparts of the challenges.

This deliverable reports on the state of the preparations for the organization of the HOBBIT challenges in 2018. In particular, in Section 2 the eight benchmarks are presented, along with their correspondence to the challenges. Sections 3–7, describe the five HOBBIT challenges. The motivation and an overview of each challenges is given, while the tasks comprising each challenge, are analyzed. Finally, Section 8 outlines the main actions that are required in order to organize and run the challenges and presents the current status of each challenge in terms of the actions completed.

## 2 Benchmarks Description

Eight benchmarks have been developed within the HOBBIT project. A summary of these benchmarks is given in Table 1. These comprise the second version of the benchmarks which will be employed in the five HOBBIT challenges. Table 2 shows a mapping between the HOBBIT benchmarks and the

Table 1: Overview of HOBBIT Benchmarks.

---

<b>Benchmark</b>	<b>Short Description</b>
Data Acquisition	Evaluate storage solutions that deal with the ingestion of streams of RDF data
Knowledge Extraction	Test the performance (runtime and accuracy) of entity recognition and linking frameworks over streams of unstructured data (text)
Link Discovery	Go beyond mere instance matching and check how well tools performs on other types of links (e.g., geospatial links) when faced with large amounts of data
Structured Machine Learning	This benchmark is related to maritime surveillance domain. The benchmark performs an evaluation of prediction accuracy over of algorithms for geo-positioning systems
Data Storage	Stress test storage solutions for RDF when faced with realistic scenarios such as being the backend of a social network
Versioning	Check how well storage solutions deal with storing evolving data available in several versions and performing queries on and across these different versions
Question Answering	Evaluate the performance of data access solutions that can answer questions in natural language as well as keyword queries on large amounts of data
Faceted Browsing	Test storage solutions w.r.t. their performance as backends of data browsers

---

challenges in which the latter participate. More information on the HOBBIT benchmarks are available at the project website<sup>1</sup>.

### 3 MOCHA Challenge @ ESWC 2018

The Mighty Storage Challenge (MOCHA)<sup>2</sup> was successfully organized last year in conjunction with ESWC 2017. Given the high impact and prestige of the ESWC conference, we decided to again target this conference to organize the 2018 edition of the challenge. In December 2017, a formal application was submitted and the challenge was accepted as part of ESWC 2018, which takes place from the 3rd to the 7th of June, 2018, in Heraklion, Crete, Greece.

#### 3.1 Challenge Motivation

Triple stores are the backbone of most applications based on Linked Data. Hence, devising systems that achieve an acceptable performance on real datasets and real loads is of central importance for

---

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

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



Table 2: Mapping between HOBBIT Challenges/Benchmarks.

Event	Benchmark
<b>MOCHA</b> @ESWC 2018 June 3rd to June 7th, 2018	Data Acquisition (INFAI)
	Data Storage (OpenLink)
	Versioning (FORTH)
	Faceted Browsing (IAIS)
<b>OKE</b> @ESWC 2018 June 3rd to June 7th, 2018	Knowledge Extraction (INFAI)
<b>SQA</b> @ESWC 2018 June 3rd to June 7th, 2018	Question Answering (IAIS)
<b>Grand Challenge</b> @DEBS 2018 June 25th to June 29th, 2017	Structured Machine Learning (AGT)
<b>Link Discovery Track</b> OAEI@ISWC 2018 October 8th to October 12th, 2018	Link Discovery (FORTH)

the practical applicability of Semantic Web technologies. So far, it is only partly known whether we have already passed this cap. With this challenge, we aim to provide objective measures for how well current systems perform on real tasks of industrial relevance and detect bottlenecks of existing systems to further push their development towards practical usage.

### 3.2 Challenge Overview

The aim of the MOCHA challenge is to test the performance of solutions for SPARQL processing in aspects that are relevant for modern applications. These include ingesting data, answering queries on large datasets and serving as backend for applications driven by Linked Data. The challenge will test the systems against data derived from real applications and with realistic loads. An emphasis will be put on dealing with changing data in form of streams or updates.

### 3.3 Challenge Tasks

In the context of the MOCHA challenge four benchmarking tasks are organized, which are the same tasks as in 2017. These tasks are:

- **RDF Data Ingestion:** The constant growth of the Linked Data Web in velocity and volume has increased the need for triple stores to ingest streams of data and perform queries on this data efficiently. The aim of this task is to measure the performance of SPARQL query processing systems when faced with streams of data from industrial machinery in terms of efficiency and completeness. The experimental setup will hence be as follows: we will increase the size and velocity of RDF data used in our benchmarks to evaluate how well can a system store streaming

.....

RDF data obtained from industry. The data will be generated from one or multiple resources in parallel and will be inserted using SPARQL INSERT queries. This facet of triple stores has (to the best of our knowledge) never been benchmarked before. SPARQL SELECT queries will be used to test the system's ingestion performance and storage abilities. The components of the benchmark for this task are implemented in Java.

- **Data Storage Benchmark:** The aim of this task is to measure how data stores perform with different types of queries. Therefore, this task consists of an RDF benchmark that measures how datastores perform with interactive, simple, read, SPARQL queries. Running the queries is accompanied with high insert rate of the data (SPARQL INSERT queries), in order to mimic the real use cases where READ and WRITE operations are bundled together. Typical bulk loading scenarios will also be supported. The queries and query mixes will be designed to stress the system under test in different choke-point areas, while being credible and realistic.
- **Versioning RDF Data:** The evolution of datasets often requires storing different versions of the same dataset, so that interlinked datasets can refer to older versions of an evolving dataset and upgrade at their own pace, if at all. Supporting the functionality of accessing and querying past versions of an evolving dataset is the main challenge for archiving/versioning systems. In this task we will propose the second version of the versioning benchmark that will be used to test the ability of versioning systems to efficiently manage evolving datasets, where triples are added or deleted, and queries evaluated across the multiple versions of said datasets.
- **Faceted Browsing Benchmark:** Faceted browsing stands for a session-based (state-dependent) interactive method for query formulation over a multidimensional information space. It provides a user with an effective way for exploration through a search space. After having defined the initial search space, i.e., the set of resources of interest to the user, a browsing scenario consists of applying (or removing) filter restrictions of object-valued properties or of changing the range of a number-valued properties. Using such operations, aimed at selecting resources with desired properties, the user browses from state to state, where a state consists of the currently chosen facets and facet values and the current set of instances satisfying all chosen constraints. The task on faceted browsing checks existing solutions for their capabilities of enabling faceted browsing through large-scale RDF datasets, that is, it analyses their efficiency in navigating through large datasets, where the navigation is driven by intelligent iterative restrictions. We aim to measure the performance relative to dataset characteristics, such as overall size and graph characteristics.

More information on the MOCHA tasks are available on the challenge website<sup>3</sup>.

### 3.4 Open MOCHA Challenge

The MOCHA challenge was also launched as an open challenge in September 2017<sup>4</sup>. Following the open challenges paradigm, Open MOCHA would enable potential participants to submit their systems and join the challenge at virtually any point in time. Periodic cutoff dates would determine when the submitted systems would be evaluated and winners announced. The first cutoff was set for February 2018.

The motivation behind the consortium's decision to launch an open version of the MOCHA challenge was twofold; i) the absence of hard deadlines would help towards an increased participation and

---

<sup>3</sup><https://project-hobbit.eu/challenges/mighty-storage-challenge2018/msc-tasks/>

<sup>4</sup><https://project-hobbit.eu/open-challenges/mocha-open-challenge/>

.....

ii) the Open MOCHA would allow to start the dissemination of MOCHA for 2018 very early, thus enabling us to attract more participants for the challenge hosted at ESWC 2018.

Open MOCHA consisted of the same tasks described in Section 3.3 and was disseminated as mentioned in Section 8. For Open MOCHA we introduced the concept of challenge leaderboards, where a leaderboard of the participating systems was implemented as part of the HOBBIT platform<sup>5</sup> so that participants could continuously monitor their systems' performance against other competing systems. Moreover, we introduced on the challenge's website the "Express your Interest" button, where people interested in the challenge could register by simply providing their email address. This feature would allow us to better track potential participants and inform them of important aspects of the challenge, such as deadlines, release of benchmark data, leaderboard results etc.

Despite the dissemination efforts of the project and the fact that 6 people registered via the "Express your Interest" button, no system was submitted to the open challenge in time for the first cutoff in February 2018. Together with the fact that MOCHA was accepted at ESWC 2018, the consortium decided to discontinue the open challenge and invite the registered participants to join MOCHA @ ESWC 2018.

## 4 OKE Challenge @ ESWC 2018

The Open Knowledge Extraction (OKE) challenge<sup>6</sup> was successfully organized last year in conjunction with ESWC 2017. Given the high impact and prestige of the ESWC conference, we decided to again target this conference to organize the 2018 edition of the challenge. In December 2017, a formal application was submitted and the challenge was accepted as part of ESWC 2018, which takes place from the 3rd to the 7th of June, 2018, in Heraklion, Crete, Greece.

### 4.1 Challenge Motivation

Most of the Web content consists of natural language text, e.g., websites, news, blogs, micro-posts, etc., hence a main challenge is to extract as much relevant knowledge as possible from this content and publish it in the form of Semantic Web triples. There is huge work on knowledge extraction and knowledge discovery contributing to address this problem. In fact, results of knowledge extraction systems are usually evaluated against tasks that do not focus on specific Semantic Web goals. For example, tasks such as named entity recognition, named entity disambiguation and relation extraction are certainly of importance for the Semantic Web, but in most cases such tasks are designed without considering the output design and formalization in the form of Linked Data and OWL ontologies. This makes results of existing methods often not directly reusable for populating the Semantic Web, until a translation from linguistic semantics to formal semantics is performed.

### 4.2 Challenge Overview

The OKE challenge has the ambition to provide a reference framework for research on "Knowledge Extraction from text for the Semantic Web" by redefining a number of tasks, typically from information and knowledge extraction, by taking into account specific Semantic Web requirements. The goal of the OKE challenge is to test the performance of knowledge extraction systems with respect to the

---

<sup>5</sup><https://master.project-hobbit.eu/challenges>

<sup>6</sup><https://project-hobbit.eu/challenges/oke2018-challenge-eswc-2018/>

Semantic Web. The challenge makes use of small gold standard datasets that consist of manually curated documents and large silver standard datasets that consist of automatically generated synthetic documents.

### 4.3 Challenge Tasks

In the context of the OKE challenge four benchmarking tasks are organized. The challenge comprises of two tasks from 2017 and two new tasks (Relation Extraction and Knowledge Extraction tasks). These tasks are:

- **Focused Named Entity Identification and Linking:** The task comprises the identification of named entities in sentences and the disambiguation of the identified entities to the DBpedia knowledge base. The task is limited to three DBpedia ontology classes (Person, Place, Organisation) and their associated subclasses. A competing system is expected to identify elements in a given text by its start and end index, further to generate an RDF formalizing the linking of the identified entities to the DBpedia knowledge base.
- **Broader Named Entity Identification and Linking:** This task extends the above task in terms of the entity classes considered from DBpedia. Beside the three types in the first task, a competing system has to identify additional types of entities in this case.
- **Relation Extraction:** Given documents with sentences, the DBpedia knowledge base, a target entity type hierarchy and a target relation type hierarchy, the aim of this task is to find binary relations in sentences. Furthermore, annotations of the sentences in a document are given, i.e. entity mentions with positions, the linking of each entity and its type.
- **Knowledge Extraction:** This task is a combination of the first (Focused Named Entity Identification and Linking) and third (Relation Extraction) tasks mentioned above. The data consists of documents with sentences. A participating system is expected to respond with named entities found in the documents, with the linking to DBpedia as well as with relations between the entities.

More information on the OKE tasks are available on the challenge website<sup>7</sup>.

### 4.4 Open OKE Challenge

The OKE challenge was also launched as an open challenge in September 2017<sup>8</sup>. Following the open challenges paradigm, Open OKE would enable potential participants to submit their systems and join the challenge at virtually any point in time. Periodic cutoff dates would determine when the submitted systems would be evaluated and winners announced. The first cutoff was set for February 2018.

The motivation behind the consortium's decision to launch an open version of the OKE challenge was twofold; i) the absence of hard deadlines would help towards an increased participation and ii) the Open OKE would allow to start the dissemination of OKE for 2018 very early, thus enabling us to attract more participants for the challenge hosted at ESWC 2018.

<sup>7</sup><https://project-hobbit.eu/challenges/oke2018-challenge-eswc-2018/tasks/>

<sup>8</sup><https://project-hobbit.eu/open-challenges/oke-open-challenge/>

Open OKE consisted of three tasks; i) Focused Named Entity Identification and Linking, ii) Broader Named Entity Identification and Linking and iii) Focused Musical Named Entity Recognition and Linking. These tasks were part of the OKE challenge organized at ESWC 2017. The open challenge was disseminated as mentioned in Section 8. For Open OKE we introduced the concept of challenge leaderboards, where a leaderboard of the participating systems was implemented as part of the HOBBIT platform<sup>9</sup> so that participants could continuously monitor their systems' performance against other competing systems. Moreover, we introduced on the challenge's website the "Express your Interest" button, where people interested in the challenge could register by simply providing their email address. This feature would allow us to better track potential participants and inform them of important aspects of the challenge, such as deadlines, release of benchmark data, leaderboard results etc.

Despite the dissemination efforts of the project and the fact that 16 people registered via the "Express your Interest" button, no system was submitted to the open challenge in time for the first cutoff in February 2018. Together with the fact that OKE was accepted at ESWC 2018, the consortium decided to discontinue the open challenge and invite the registered participants to join OKE @ ESWC 2018.

## 5 SQA Challenge @ ESWC 2018

The Scalable Question Answering (SQA) challenge<sup>10</sup> is an offspring of the QALD (Question Answering over Linked Data) challenge that was successfully organized last year by HOBBIT, in conjunction both with ESWC 2017 and ISWC 2017. Given the high impact and prestige of the ESWC conference, we decided to target this conference to organize SQA. In December 2017, a formal application was submitted and the challenge was accepted as part of ESWC 2018, which takes place from the 3rd to the 7th of June, 2018, in Heraklion, Crete, Greece.

### 5.1 Challenge Motivation

The past years have seen a growing amount of research on question answering (QA) over Semantic Web data, shaping an interaction paradigm that allows end users to profit from the expressive power of Semantic Web standards while at the same time hiding their complexity behind an intuitive and easy-to-use interface. At the same time the growing amount of data has led to a heterogeneous data landscape where QA systems struggle to keep up with the volume, variety and veracity of the underlying knowledge. The SQA challenge aims at providing an up-to-date benchmark for assessing and comparing state-of-the-art-systems that mediate between a user, expressing his or her information need in natural language, and RDF data. The main goal is to gain insights into the strengths and shortcomings of different approaches and into possible solutions for coping with the large, heterogeneous and distributed nature of Semantic Web data.

### 5.2 Challenge Overview

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 the SQA challenge is

<sup>9</sup><https://master.project-hobbit.eu/challenges>

<sup>10</sup><https://project-hobbit.eu/challenges/sqa-challenge-eswc-2018/>

.....

the following: given an RDF dataset and a large volume of natural language questions or keywords, return the correct answers or the SPARQL queries that retrieve those answers.

### 5.3 Challenge Tasks

In the context of the SQA challenge a single benchmarking task is organized. This task is the continuation of the third task of the 2017 QALD challenge, namely Large-Scale Question Answering over RDF<sup>11</sup>. The SQA task is:

- **Scalable Question Answering over Linked Data:** The focus of this task is towards large-scale question answering, where given an RDF dataset and a large volume of natural language questions or keywords, the correct answers (or SPARQL queries that retrieve those answers) must be returned. The ranking of participating systems is based on the usual KPIs (precision, recall and F-measure) plus a “response power” measure, which is also taking into account the ability of the systems to cope with high volume demand without failure. The response power is the harmonic mean of three measures: precision, recall and the ratio between processed questions (an empty answer is considered as processed, a missing answer is considered as unprocessed) and total number of questions sent to the system.

More information on the SQA task is available on the challenge website<sup>12</sup>.

### 5.4 Open SQA Challenge

The SQA challenge was also launched as an open challenge in December 2017<sup>13</sup>. Following the open challenges paradigm, Open SQA would enable potential participants to submit their systems and join the challenge at virtually any point in time. Periodic cutoff dates would determine when the submitted systems would be evaluated and winners announced. The first cutoff was set for May 2018.

The motivation behind the consortium’s decision to launch an open version of the SQA challenge was twofold; i) the absence of hard deadlines would help towards an increased participation and ii) the Open SQA would allow to start the dissemination of SQA for 2018 early, thus enabling us to attract more participants for the challenge hosted at ESWC 2018.

Open SQA consisted of the same tasks described in Section 5.3 and was disseminated as mentioned in Section 8. For Open SQA we introduced the concept of challenge leaderboards, where a leaderboard of the participating systems was implemented as part of the HOBBIT platform<sup>14</sup> so that participants could continuously monitor their systems’ performance against other competing systems. Moreover, we introduced on the challenge’s website the “Express your Interest” button, where people interested in the challenge could register by simply providing their email address. This feature would allow us to better track potential participants and inform them of important aspects of the challenge, such as deadlines, release of benchmark data, leaderboard results etc.

Despite the dissemination efforts of the project and the fact that 17 people registered via the “Express your Interest” button, no system was submitted to the open challenge till March 2018. Together with the fact that SQA was accepted at ESWC 2018, the consortium decided to discontinue the open challenge and invite the registered participants to join SQA @ ESWC 2018.

<sup>11</sup><https://project-hobbit.eu/challenges/qald2017/qald2017-challenge-tasks/#task3>

<sup>12</sup><https://project-hobbit.eu/challenges/sqa-challenge-eswc-2018/#tasks>

<sup>13</sup><https://project-hobbit.eu/open-challenges/sqa-open-challenge/>

<sup>14</sup><https://master.project-hobbit.eu/challenges>

.....

---

## 6 Grand Challenge @ DEBS 2018

### 6.1 Challenge Motivation

The 2018 ACM DEBS Grand Challenge is the eighth in a series of challenges which seek to provide a common ground and uniform evaluation criteria for a competition aimed at both research and industrial event-based systems. Every year, the DEBS Grand Challenge participants have a chance to explore a new data set and a new problem and can compare their results based on the common evaluation criteria.

This year's competition focuses 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' destinations and arrival times.

### 6.2 Challenge Overview

The focus of the 2018 Grand Challenge is 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' 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. The Grand Challenge data is provided by the MarineTraffic company and hosted by the Big Data Ocean project (an EU Horizon 2020 project). The evaluation platform is provided by the HOBBIT project represented by AGT International<sup>15</sup>.

The 2018 edition of the Grand Challenge is part of DEBS 2018, which takes place from June 25th to June 29th, 2018, in Hamilton, New Zealand and is co-organized by ACM DEBS conference and HOBBIT.

### 6.3 Challenge Tasks

In the context of the 2018 Grand Challenge a benchmarking task is organized including two queries pertaining to the Structured Machine Learning benchmark. The overall goal of the challenge task is to make the naval transportation industry more reliable by providing predictions for vessels' 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.

More specifically query 1 concerns the prediction of the correct destination of a vessel. The prediction problem is to generate a continuous stream of predictions for the destination port of any vessel given the following information: (1) name of the port of origin, (2) unique ID of the vessel, (3) position of the vessel, (4) time stamp, and (5) vessel's draught. The above data is provided as a continuous stream of tuples and the goal of the system is to provide for every input tuple one output tuple containing the name of the destination port. A solution is considered correct at time stamp T if for a tuple with this timestamp as well as for all subsequent tuples the predicted destination port matches the actual destination port. The goal of any solution is not only to predict a correct destination port but also to predict it as soon as possible counting from the moment when a new port of origin appears for a given vessel. Moreover the goal of query 2 is the prediction of arrival times of ships. In particular there is a set of ports defined by respective bounding boxes of coordinates. Once

---

<sup>15</sup><http://www.agtinternational.com/>

.....

a ship leaves a port (i.e. the respective bounding box), the task is to predict the arrival time at its destination port (i.e. when the next defined bounding boxes will be entered). Also for this query, after port departure and until arrival, the solution must emit one prediction per position update.

More information on the Grand Challenge queries are available on the challenge website<sup>16</sup>.

## 6.4 StreamML Open Challenge

The StreamML (Stream Machine Learning) Open Challenge was launched as an open challenge in February 2018. Following the open challenges paradigm, the StreamML Open Challenge would enable potential participants to submit their systems and join the challenge at virtually any point in time. Periodic cutoff dates would determine when the submitted systems would be evaluated and winners announced. The first cutoff was set for May 2018.

StreamML Open Challenge comprised the offspring of DEBS GC 2017 as it reused the data set and the task description of DEBS GC 2017. In contrast to DEBS GC in StreamML Open Challenge systems would be continuously evaluated and the results would be shown on a leaderboard. The goal of the StreamML Open Challenge competition was to evaluate event-based systems for real-time analytics over high velocity and high volume data streams generated by manufacturing equipment. For more details regarding the tasks and training Data pertaining to StreamML Open Challenge please see <https://project-hobbit.eu/open-challenges/streamml-open-challenge/#custom-tab-0-76c2cee4b6fcaa61923e29550a021814>.

StreamML Open Challenge was disseminated as mentioned in Section 8. For StreamML Open Challenge as mentioned before we also introduced the concept of challenge leaderboards, where a leaderboard of the participating systems was implemented as part of the HOBBIT platform so that participants could continuously monitor their systems' performance against other competing systems. Moreover, we introduced on the challenge's website the "Express your Interest" button, where people interested in the challenge could register by simply providing their email address. This feature would allow us to better track potential participants and inform them of important aspects of the challenge, such as deadlines, release of benchmark data, leaderboard results etc.

Despite the dissemination efforts of the project and the fact that 10 people registered via the "Express your Interest" button, no system was submitted to the open challenge till March 2018. Therefore, the consortium decided to discontinue the open challenge and invite the registered participants to join DEBS GC 2018.

## 7 HOBBIT Link Discovery Task at OAEI 2017.5/2018 at ISWC 2018

### 7.1 Challenge Motivation

The number of datasets published in the Web of Data as part of the Linked Data Cloud is constantly increasing. The Linked Data paradigm is based on the unconstrained publication of information by different publishers, and the interlinking of Web resources across knowledge bases. In most cases, the cross-dataset links are not explicit in the dataset and must be automatically determined using Instance Matching (IM) and Link Discovery tools amongst others. The large variety of techniques requires their comparative evaluation to determine which one is best suited for a given context. Performing such an

<sup>16</sup><https://project-hobbit.eu/challenges/debs2018-grand-challenge/>



.....

assessment generally requires well-defined and widely accepted benchmarks to determine the weak and strong points of the proposed techniques and/or tools.

## 7.2 Challenge Overview

Although a number of real and synthetic benchmarks that address different data linking challenges have been proposed for evaluating the performance of such systems. So far, only a limited number of link discovery benchmarks target the problem of linking geo-spatial entities.

However, some of the largest knowledge bases on the Linked Open Data Web are geospatial knowledge bases (e.g., LinkedGeoData with more than 30 billion triples). Linking spatial resources requires techniques that differ from the classical mostly string-based approaches. In particular, considering the topology of the spatial resources and the topological relations between them is of central importance to systems driven by spatial data.

We believe that due to the large amount of available geospatial datasets employed in Linked Data and in several domains, it is critical that benchmarks for geospatial link discovery are developed.

The HOBBIT Link Discovery tasks at OAEI 2017.5 and at OAEI 2018 have been accepted at ISWC 2018, which takes place from October 8th to October 12th, 2018, in Monterey, California. In particular HOBBIT Link Discovery task 2017.5 comprises a preliminary campaign which will provide the first results which will be presented together with the HOBBIT Link Discovery task results at OAEI 2018 in ISWC 2018. Most importantly in the HOBBIT Link Discovery Tasks 2017.5 and 2018 the HOBBIT platform will be used for evaluation.

## 7.3 Challenge Tasks

In the context of the OAEI 2017.5 and OAEI 2018 Challenges, two benchmarking tasks namely Linking (Task 1) and Spatial (Task 2) are organized. In both tasks TomTom datasets are used for the creation of the appropriate benchmarks. TomTom datasets contain representations of traces (GPS fixes). Each trace consists of a number of points. Each point has time stamp, longitude, latitude and speed (value and metric). The points are sorted by timestamp of the corresponding GPS fix (ascending).

In particular, Task 1 (Linking), will measure how well the systems can match traces that have been altered using string-based approaches along with addition and deletion of intermediate points. As the TomTom dataset only contains coordinates and in order to apply string-based modifications based on LANCE<sup>17</sup> we have replaced a number of those with labels retrieved from Google Maps Api, Foursquare Api and Nominatim Openstreetmap Api. This task also contains changes on date format and changes on coordinate formats.

Moreover, Task 2 (Spatial) measures how well the systems can identify DE-9IM (Dimensionally Extended nine-Intersection Model) topological relations. The supported spatial relations are the following: Equals, Disjoint, Touches, Contains/Within, Covers/CoveredBy, Intersects, Crosses, Overlaps and the traces are represented in Well-known text (WKT) format. For each relation, a different pair of source and target dataset will be given to the participants.

---

<sup>17</sup>Saveta, E. Daskalaki, G. Flouris, I. Fundulaki, and A. Ngonga-Ngomo. LANCE: Piercing to the Heart of Instance Matching Tools. In ISWC, 2015

---

## 8 Milestones for the Organization of Challenges

In order to reassure the smooth preparation and organization of the selected events/workshops we undertook several actions:

1. Website Preparation: We proceeded to the website preparation pertaining to each challenge and open challenge in cooperation with the challenge organizers and FORTH. The full details for each challenge can be found on the HOBBIT website: <https://project-hobbit.eu/challenges/>.
2. Website for each challenge launched with “Express your Interest” button. This button allow us to easily track potential participants to the challenges.
3. Task Definition: We finalized the benchmark tasks descriptions that participate in each challenge.
4. Datasets & HOBBIT Platform: We provided the detailed description of the training & test data sets. We also ensured the timely preparation and delivery of training & test data pertaining to each task and of the HOBBIT platform for evaluation/testing. The detailed descriptions of data sets participating in each challenge task can be found on the website of each challenge under the tab “Tasks & Training Data”.
5. Specification of Program and Organizing Committees: We defined the organizing and the program committees for each challenge. Towards this action we sent several thousand email invitations to specific persons and to lists of persons with scientific expertise relevant to the tasks pertaining to each challenge. This action has been performed in cooperation with FORTH.
6. Preparation of Call for Participation (CFP): We structured the call for participation for each challenge and open challenge.
7. Publication of Call for Participation (CFP): The CFP for each challenge/open challenge has been published to various channels as follows:
  - to WikiCFP<sup>18</sup> and the EasyChair Smart CFP<sup>19</sup>
  - to 95 mailing lists related to semantic web (e.g. [semantic-web@w3.org](mailto:semantic-web@w3.org)).
  - The CFPs together with the “Stay Tuned and Participate” emails have were sent repeatedly to the following lists which consisting of emails of individual persons:
    - (a) the HOBBIT community list containing 116 subscribers. This list contains emails of people from all the organizations and universities participating in the HOBBIT consortium.
    - (b) the HOBBIT contact list which contains 182 members. This list consists of emails of members who expressed their interest and subscribed into the HOBBIT project and are not members of the HOBBIT consortium.
    - (c) the MOCHA, OKE, and QALD, challenges mailing lists: these lists comprise emails of persons who expressed their interest for MOCHA, OKE and QALD challenges, respectively, through the following googlegroup mailing lists:
      - [mightystoragechallenge-contact@googlegroups.com](mailto:mightystoragechallenge-contact@googlegroups.com)
      - [OKE-contact@googlegroups.com](mailto:OKE-contact@googlegroups.com)
      - [qald-contact@googlegroups.com](mailto:qald-contact@googlegroups.com)

---

<sup>18</sup><http://www.wikicfp.com/cfp/>

<sup>19</sup><https://easychair.org/cfp/>

- .....
- (d) the authors' list comprising 1027 authors who participated in previous editions of OKE, MOCHA, QALD and DEBS challenges
  - (e) the subscribed participants of the BioASQ project<sup>20</sup>
  - (f) the "Express your Interest" registered users
  - (g) the HOBBIT Twitter account
- Regarding HOBBIT Open challenges MOCHA, OKE and StreamML leaderboards launched and advertised to above channels
8. In summary, for each of the booked challenges we undertook the following actions with regards to various deadlines:
- (a) Publication of Call for Papers
  - (b) Publication of training data for each challenge task
  - (c) Submission of challenge papers by participants
  - (d) Reviewing of challenge papers
  - (e) Notification of acceptance/rejection of challenge papers
  - (f) Publication of test data for each challenge task
  - (g) Submission of camera-ready papers by participants
  - (h) Submission of systems by participants
  - (i) Release of evaluation results for participating systems
  - (j) Proclamation of winners
  - (k) Proceedings preparation

All the deadlines foreseen for points a-k for each of the upcoming challenges have been published in the corresponding challenges' websites, under the tab "important dates".

All the above actions (1-8) have been accomplished in cooperation with the HOBBIT partners responsible for the organization of each challenge.

## 8.1 Current Status of Challenges

The three challenges that are hosted by ESWC 2018 from June 3rd to June 7th, 2018, namely MOCHA, OKE and SQA, are at an advanced stage of preparation. Actions 1-7 have been completed, while for Action 8, the deadlines for points a-h have already passed and these points have been successfully completed. The MOCHA challenge has received 2 participating systems-papers each, SQA has received 5 participating systems-papers, while the OKE challenge has received 1 system-paper. The camera-ready papers describing the participating systems will be included in the challenges' post proceedings, which will be published by Springer as a Lecture Notes in Computer Science (LNCS) volume after the ESWC conference.

The Grand Challenge that is hosted by DEBS 2018 from June 25th to June 29th, 2018, is also at an advanced stage of preparation. Actions 1-7 have been completed, while for Action 8, the deadlines for points a-h have already passed and these points have been successfully completed. The Grand

---

<sup>20</sup><http://participants-area.bioasq.org/>

.....

Challenge has received 12 participating systems. The Grand Challenge proceedings will be published by ACM.

Finally, the HOBBIT Link Discovery Task will be part of the OAEI 2017.5 and OAEI 2018 will be held at the Ontology Matching workshop at ISWC 2018 from October 8th to October 12th. Regarding the Link Discovery Task of the OAEI 2017.5 campaign Actions 1-7 have been completed while for Action 8 the deadlines for points a-h have already passed and these points have been successfully completed. Task 1 (Linking) has received 3 systems, while Task 2 (Spatial) has received 4 systems. The Link Discovery task for OAEI 2018 is in a very early stage. Currently, the website and the call for participation are being prepared, while immediately after the program committee and the benchmarking tasks will be finalized.