Press Release
December 2016

HOBBIT
Holistic Benchmarking of
Big Linked Data

H2020 Research and Innovation
Grant Agreement No. 688227

Website
http://project-hobbit.eu

Contact
Dr. Axel-Cyrille Ngonga Ngomo
Institute for Applied Informatics
Hainstraße 11, 04109 Leipzig
Germany
Phone: +49 341 97 32362
ngonga@informatik.uni-leipzig.de

Hobbit so far. Hobbit is now one year old!

HOBBIT during its first year was very active in several areas: members of the consortium attended a number of academic and industry-related events and conferences and published a number of posts on several topics related to benchmarking. Furthermore, members of the HOBBIT community were asked to complete a Survey that allowed HOBBIT partners to clarify the requirements for the HOBBIT benchmarks and platform. Our next immediate steps include the release of the first version of the HOBBIT platform and the establishment of the HOBBIT association.

Attendance to conferences and events

HOBBIT has attended several academic and industry driven conferences and events. We are proud to inform you that HOBBIT was present at the following academic events:

- The International Conference on Big Data 2016 (May 3 - 4, Alicante, Spain) where the project coordinator Axel Ngonga Ngomo (InfAI) presented the project,
- The ESWC 2016 conference (May 29 - June 2, Anissaras, Crete, Greece), where Tzanina Saveta (FORTH) and Irini Fundulaki (FORTH) presented the tutorials Instance Matching Benchmarks for Linked Data and Assessing the performance of RDF Engines respectively. Axel Ngonga Ngomo (InfAI) presented the following two papers in the conference: Detecting Similar Linked Datasets Using Topic Modelling and The Lazy Traveling Salesman - Memory Management for Large-Scale Link Discovery. HOBBIT had also a half-day session (EU HOBBIT Workshop) where Michael Röder (InfAI) gave an overview of the current design of the architecture underlying the HOBBIT platform and Ruben Taelman (iMinds) presented the most important results of the Survey that was distributed to the HOBBIT Community.
- The ACL 2016 (August 7-12, Berlin), where Anastasia Krithara (NCSR) gave a brief introduction of HOBBIT.
- The ISWC 2016 (October 17-21, Kobe Japan), where two tutorials were presented: the Link Discovery – Algorithms, Approaches and Benchmarks by Axel Ngonga Ngomo, Mohamed Ahmed Sherif (InfAI) and Irini Fundulaki (FORTH) and the SPARQL Querying Benchmarks by Axel Ngonga Ngomo, Muhammad Saleem and Ricardo Usbeck (InfAI). In addition to the tutorials, Axel Ngonga Ngomo (InfAI) and Irini Fundulaki (FORTH) were the two organizers of the 1st International Workshop on Benchmarking Linked Data (BLINK) who also presented three papers at the workshop.

- The EKAW 2016 (November 19-23, Bologna, Italy), where the work entitled TAIPAN: Automatic Property Mapping for Tabular Data was presented by Ivan Ermilov (InfAI).

HOBBIT was present at the following industry-driven events:
- The post EDF 2106 event (July 1, Eindhoven, Netherlands), that was co-organized by HOBBIT, BigDataEurope and BYTE projects. Axel Ngonga Ngomo (InfAI) gave a presentation of HOBBIT and with Irini Fundulaki (FORTH) and Frank Salliau (iMinds) discussed with the attendees their requirements regarding the benchmarks for Linked Data.
- Finally, HOBBIT was present at the ApacheCon Europe (November 15-18, Sevilla, Spain), where Axel Nonga Ngomo (InfAI) presented the project.
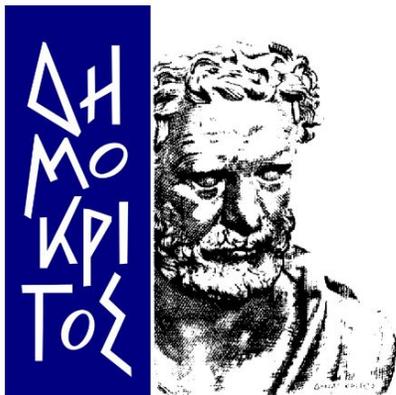
BlogPosts

During the same period, several interesting posts were also published at the HOBBIT website (project-hobbit.eu/) including a post on HOBBIT's initial data management plan, a mini survey on benchmarking RDF Query Engines, the approaches and benchmarks on versioning for Big Linked Data, HOBBIT's potential for enabling the business value of IoT, the industrial perspecitves of HOBBIT and IT Management, a post on linking benchmark for spatial data, and finally a post on HOBBIT's future. Furthermore several posts were published to present the HOBBIT's participation in the conferences presented previously.

Survey

We developed a survey that was distributed to the HOBBIT community with the purpose of gathering the requirements for the HOBBIT platform and benchmarks. The aim of this fully anonymized survey was for the HOBBIT consortium to determine how a participant's' company or organization could benefit from the HOBBIT platform. The survey participants were evenly divided among the different profiles, namely solution providers, technology users and scientific community (with most of them having two or more of these profiles). The survey supported our belief that HOBBIT platform will prove itself very useful for both the academia and industry since it will provide useful tools to the people involved.

More specifically 61 people (representing their organization) participated in the survey: 48 solution providers, 46 technology users and finally 47 people from the scientific community (note that participants may be associated with multiple profiles). From the results of the survey we note that Solution Providers are very interested in Storage and Querying solutions, equally but less interested in Interlinking, Classification/Enrichment and Discovery tools. Technology Users are mostly interested in Storage/Querying, followed by Classification/Enrichment and then Interlinking and members of the Scientific Community are very interested in Storage/Querying, followed by Classification/Enrichment and Discovery followed by Interlinking, Extraction, and Reasoning.

Participants were also asked the Key Performance Indicators (KPIs) of interest. Most prominent are Correctness, Accuracy followed by Runtime / Speed.

Learn about our Challenges

Join HOBBIT Community

Scalability, Memory and CPU usage are also important KPIs for the Survey participants.

Moving forward

Next steps

HOBBIT's next big step will be the release of the first version of the HOBBIT platform by February 2017. In addition, we are organizing the first HOBBIT evaluation campaigns that will take place in 2017: the DEBS Grand Challenge (Holistic Benchmarking of Big Linked Data),the Instance Matching Benchmark Challenge for Spatial Data at the OAEI workshop of ISWC, and the Mighty Storage Challenge that will be held in conjunction with ESWC 2017 More details can be found at the project's website https://project-hobbit.eu/challenges/.

Furthermore, an alpha version of the HOBBIT Data Generators is ready: currently we can generate Twitter, molding machine data as well as transportation data using those generators.

HOBBIT association

Finally, we are planning to establish the HOBBIT association by December 2017. We believe that it will play a key role in disseminating and extending the HOBBIT benchmarking platform (including benchmarks, the KPIs, use cases and datasets). Towards this direction, we will initiate a new survey to gather information from the industry and other potential partners on frameworks used to evaluate their software. The goal of the HOBBIT association will be to provide industry-grade benchmarks for Big Linked Data. The HOBBIT association intends to offer various types of membership depending on companies' specific needs, will work, hand in hand with other European benchmarking associations such as LDBC, provide periodic open reports comparing system performances for each of the Big Linked Data steps and financed via member fees, donations and the implementation of dedicated benchmarks for companies. These financial resources will be used to maintain the platform and organize further challenges for Big Linked Data even after the completion of the HOBBIT project.

Join our community to remain informed at project-hobbit.eu/get-involved. If you have any questions, do not hesitate to contact us as explained at project-hobbit.eu/contacts.