

Δελτίο Τύπου

Ιανουάριος 2018

HOBBIT

Holistic Benchmarking of
Big Linked Data

H2020 Research and Innovation
Grant Agreement No. 688227

Ιστοσελίδα
<http://project-hobbit.eu>

Επικοινωνία

Dr. Axel-Cyrille Ngonga Ngomo
Institute for Applied Informatics
Hainstraße 11, 04109 Leipzig
Germany

Τηλ.: +49 341 97 32362
ngonga@informatik.uni-leipzig.de



HOBBIT
Holistic Benchmarking
of Big Linked Data

Το HOBBIT γιορτάζει δύο χρόνια!

Είμαστε στην ευχάριστη θέση να ανακοινώσουμε ότι το έργο HOBBIT έγινε πλέον δύο ετών! Τα μέλη της κοινοπραξίας υπήρξαν πολύ ενεργά κατά τη δεύτερη χρονιά του έργου και εργάστηκαν με πάθος για την επίτευξη των βασικών στόχων του έργου: την ανάπτυξη προτύπων αξιολόγησης (benchmarks), την ανάπτυξη της HOBBIT πλατφόρμας και την διοργάνωση εκστρατειών αξιολόγησης (challenges).

Το HOBBIT ολοκλήρωσε και παρέδωσε την πρώτη έκδοση των προτύπων αξιολόγησης *Παραγωγής και Συλλογής (Generation & Acquisition)*, *Ανάλυσης και Επεξεργασίας (Analysis & Processing)*, *Αποθήκευσης και Επιμέλειας (Storage & Curation)*, και *Οπτικοποίησης και Υπηρεσιών (Visualization & Services)*. Συγκεκριμένα:

- Τα πρότυπα αξιολόγησης *Παραγωγής και Συλλογής* μετρούν την επίδοση συστημάτων επεξεργασίας SPARQL ερωτήσεων, τα οποία δέχονται ροές δεδομένων από βιομηχανικά μηχανήματα, ως προς την αποδοτικότητα και την πληρότητα. Προκειμένου να αντικατοπτριστούν με ρεαλιστικό τρόπο τα φορτία που δέχονται τα συστήματα αποθήκευσης τριπλετών (triple stores) σε πραγματικές εφαρμογές, συλλέχθηκαν πραγματικά δεδομένα γύρω από τις δημόσιες συγκοινωνίες, το Twitter, την κυκλοφορία και τους αισθητήρες στις μονάδες έγχυσης πλαστικών της Weidmüller. Στο πλαίσιο αυτό το HOBBIT παρέχει επίσης πρότυπα αξιολόγησης για την μέτρηση της απόδοσης συστημάτων εξαγωγής σε μη δομημένες ροές δεδομένων φυσικής γλώσσας. Για την δημιουργία αυτών των προτύπων αξιολόγησης λάβαμε υπόψη τα πραγματικά μη δομημένα σύνολα δεδομένων που έχουν συλλεχθεί από εμπειρογνώμονες καθώς και τις μη δομημένες ροές δεδομένων του Bengal, μιας γενικού σκοπού γεννήτριας δεδομένων.
- Τα πρότυπα αξιολόγησης *Ανάλυσης και Επεξεργασίας* επικεντρώνονται στην αξιολόγηση *συστημάτων εντοπισμού συνδέσμων και μεθόδων μηχανικής μάθησης* (με και χωρίς επίβλεψη) για την ανάλυση δεδομένων. Πιο συγκεκριμένα, τα πρότυπα αξιολόγησης εντοπισμού συνδέσμων που έχει αναπτύξει το HOBBIT μπορούν να χρησιμοποιηθούν για την αξιολόγηση (α) εργαλείων αντιστοίχισης στιγμιότυπων που χρησιμοποιούν μεθόδους που βασίζονται σε συμβολοσειρές για τον εντοπισμό των οντοτήτων που ταιριάζουν και (β) συστημάτων που επικεντρώνονται στις τοπολογικές σχέσεις που έχουν προταθεί στο DE-9IM μοντέλο (Dimensionally Extended nine-Intersection Model). Το πρότυπο αξιολόγησης ανάλυσης χρησιμοποιείται για την αξιολόγηση της αποδοτικότητας και της αποτελεσματικότητας μεθόδων μηχανικής μάθησης (με και χωρίς επίβλεψη) σε δομημένα δεδομένα.





- Τα πρότυπα αξιολόγησης Αποθήκευσης και Επιμέλειας στοχεύουν στην αξιολόγηση της επίδοσης *συστημάτων αποθήκευσης δεδομένων* και *συστημάτων εκδόσεων* για Συνδεδεμένα Δεδομένα. Το πρότυπο αξιολόγησης αποθήκευσης δεδομένων επικεντρώνεται στις τυπικές προκλήσεις που αντιμετωπίζουν τα συστήματα αποθήκευσης συνδεδεμένων δεδομένων και βασίζεται στο πρότυπο αξιολόγησης Social Network Benchmark (SNB) που έχει αναπτυχθεί από το Ευρωπαϊκό FP7 έργο LDBC. Το πρότυπο αξιολόγησης εκδόσεων στοχεύει στην αξιολόγηση της ικανότητας των συστημάτων εκδόσεων να διαχειριστούν αποδοτικά Συνδεδεμένα Δεδομένα που εξελίσσονται καθώς και ερωτήσεις που απαντώνται κάνοντας χρήση πολλαπλών εκδόσεων των δεδομένων. Το πρότυπο αξιολόγησης εκδόσεων επεκτείνει το Semantic Publishing Benchmark (SPB) του έργου LDBC που είναι εμπνευσμένο από τον εκδοτικό τομέα.
- Τα πρότυπα αξιολόγησης Οπτικοποίησης και Υπηρεσιών στοχεύουν στην αξιολόγηση της επίδοσης *συστημάτων απάντησης ερωτήσεων* και *συστημάτων πολύπλευρης περιήγησης* (faceted browsing systems) σε Συνδεδεμένα Δεδομένα, χωρίς τη μεσολάβηση χρηστών. Τα πρότυπα αξιολόγησης που αναπτύχθηκαν δεν έχουν ως στόχο να αξιολογήσουν τις διεπαφές χρηστών, αλλά επικεντρώνονται στην παροχή μετρήσεων για την επίδοση και την ακρίβεια των μεθόδων που χρησιμοποιούνται στις διεπαφές χρηστών. Για τα πρότυπα αξιολόγησης αναπτύχθηκαν σενάρια περιήγησης που ανταποκρίνονται σε πραγματικά σενάρια χρήσης και τα οποία αξιολογούν τα συστήματα ως προς διάφορα επίπεδα δυσκολίας.



Όλα τα προαναφερθέντα πρότυπα αξιολόγησης είναι διαθέσιμα στο αποθετήριο CKAN του έργου HOBBIT στη διεύθυνση <https://ckan.project-hobbit.eu/dataset>, μαζί με τον πηγαίο κώδικα και τις σχετιζόμενες δημοσιεύσεις

Η HOBBIT πλατφόρμα έγινε διαθέσιμη στις αρχές του δεύτερου έτους του έργου. Πρόκειται για μία ανοικτού κώδικα κατανεμημένη πλατφόρμα συγκριτικής αξιολόγησης (benchmarking platform) για τον κύκλο ζωής συνδεδεμένων δεδομένων, η οποία ακολουθεί τις αρχές του μοντέλου FAIR και η οποία μπορεί να μεταφορτωθεί και να εκτελεστεί τοπικά. Η HOBBIT πλατφόρμα μπορεί επίσης να πορσπελαστεί διαδικτυακά και να χρησιμοποιηθεί α) για την οργάνωση εκστρατειών αξιολόγησης (benchmarking challenges) και β) από ανθρώπους οι οποίοι δεν διαθέτουν την απαραίτητη υποδομή και οι οποίοι επιθυμούν να χρησιμοποιήσουν τα προσφερόμενα πρότυπα αξιολόγησης.

Η HOBBIT πλατφόρμα είναι προσπελάσιμη διαδικτυακά στη διεύθυνση master.project-hobbit.eu και ο κώδικάς της είναι διαθέσιμος στη διεύθυνση <https://github.com/hobbit-project>.

Τα πρότυπα αξιολόγησης (benchmarks) και η πλατφόρμα χρησιμοποιήθηκαν εκτενώς στις εκστρατείες αξιολόγησης (challenges) που διοργάνωσε το έργο HOBBIT. Οι εκστρατείες αξιολόγησης είναι οι εξής: "Mighty Storage Challenge (MOCHA)", "Query Answering over Linked Data (QALD)", "Open Knowledge Extraction (OKE)" και "DEBS Grand Challenge". Οι εκστρατείες αξιολόγησης (MOCHA, QALD και OKE) οργανώθηκαν στο πλαίσιο του ESWC 2017, ενώ το DEBS Grand Challenge πραγματοποιήθηκε σε συνεργασία με τη συνέδριο DEBS 2017.

Μάθετε για τις εκστρατείες αξιολόγησης του HOBBIT



καθώς για τις ανοιχτές εκστρατείες αξιολόγησης



Γίνε μέλος της κοινότητας HOBBIT



Εκτός από τα προαναφερθέντα, το HOBBIT διοργάνωσε την εκστρατεία αξιολόγησης QALD-8 στο πλαίσιο του συνεδρίου ISWC 2017. Το QALD-8 ενσωματώθηκε στο Workshop “Natural Language Interfaces for Web of Data (NLIWoD)”. Το HOBBIT πρότεινε και συνδιοργάνωσε μια νέα ενότητα εργασίας (track) στο Workshop “Ontology Matching (OM) 2017”, το οποίο λειτουργεί υπό την αιγίδα του ΟΑΕΙ. Μετά από αυτήν την επιτυχημένη προσπάθεια, οι διοργανωτές του OM και τα μέλη της κοινοπραξίας HOBBIT αποφάσισαν να εργαστούν για την αντικατάσταση της πλατφόρμας SEALS, η οποία χρησιμοποιούνταν επί σειρά ετών στο πλαίσιο του OM, με την πλατφόρμα HOBBIT. Η εκστρατεία ΟΑΕΙ 2017.5 που θα διεξαχθεί σε συνεργασία με το ESWC 2018, έχει ως στόχο οι συμμετέχοντες να συγκρίνουν τα συστήματά τους χρησιμοποιώντας αποκλειστικά την πλατφόρμα HOBBIT. Περισσότερες πληροφορίες σχετικά με την εκστρατεία ΟΑΕΙ 2017.5 μπορείτε να βρείτε εδώ: <http://oaei.ontologymatching.org/2017.5>

Ο στόχος του Mighty Storage Challenge (MOCHA) ήταν να αξιολογήσει προσεγγίσεις για την επεξεργασία SPARQL σε πτυχές που σχετίζονται με σύγχρονες εφαρμογές. Αυτές περιλαμβάνουν τη λήψη δεδομένων, την απάντηση ερωτημάτων βάσει μεγάλων δεδομένων και την στήριξη (backend) εφαρμογών που βασίζονται στα δια-συνδεδεμένα δεδομένα. Τρία συστήματα συμμετείχαν στις εργασίες MOCHA: (α) Το OpenSource Edition Virtuoso 7.2, που αναπτύχθηκε από την OpenLink Software, το οποίο χρησίμευσε ως baseline σύστημα για όλες τα tasks του MOCHA 2017 (MOCHA Baseline), (β) QUAD, (γ) Virtuoso Commercial Edition 8.0 (beta), που αναπτύχθηκε από την OpenLink Software.

Η εκστρατεία αξιολόγησης «Question Answering over Linked Data (QALD)» είχε ως στόχο να προσφέρει ένα σύγχρονο πρότυπο αξιολόγησης (benchmark) για την αξιολόγηση και τη σύγκριση state-of-the-art συστημάτων που μεσολαβούν μεταξύ ενός χρήστη που εκφράζεται σε φυσική γλώσσα και των RDF δεδομένων. Η εκστρατεία αξιολόγησης αποτελούνταν από τέσσερα tasks: τις πολυγλωσσικές ερωταπαντήσεις στην DBpedia, τις υβριδικές ερωταπαντήσεις, τις ερωταπαντήσεις μεγάλης κλίμακας σε RDF και τέλος, τις ερωταπαντήσεις σε Wikidata. Τρία συστήματα συμμετείχαν στο task: το WDAqua, το AMAL και το ganswer.

Η εκστρατεία αξιολόγησης (ΟΚΕ) διοργάνωσε τέσσερα διαφορετικά tasks: α) την Στοχευμένη Αναγνώριση Ονομάτων Οντοτήτων, β) την Ευρύτερη Αναγνώριση Ονομάτων Οντοτήτων και Δια-σύνδεση γ) την Στοχευμένη Αναγνώριση Μουσικών Ονομάτων Οντοτήτων και Δια-σύνδεση και δ) την Εξαγωγή Γνώσης. Στόχος του ΟΚΕ είναι να αξιολογήσει την απόδοση Συστημάτων Εξαγωγής Γνώσης από το Semantic Web. Τα συστήματα Adel και Fox συμμετείχαν στην εκστρατεία αξιολόγησης (ΟΚΕ) Challenge.

Τέλος, κυρίως θέμα του DEBS 2017 Grand Challenge ήταν η ανάλυση των RDF streaming δεδομένων που παράγονται από ψηφιακούς και αναλογικούς αισθητήρες που βρίσκονται στον κατασκευαστικό εξοπλισμό. Ο στόχος της εκστρατείας αξιολόγησης ήταν να εφαρμοστεί η ανίχνευση ανωμαλιών στη συμπεριφορά του κατασκευαστικού εξοπλισμού. Η εκστρατεία αξιολόγησης συνδιοργανώθηκε με την AGT International για λογαριασμό του έργου HOBBIT.

Τέλος, το HOBBIT εγκαινίασε ένα σύνολο από εκστρατείες αξιολόγησης ανοιχτού τύπου (Open Challenges) που είναι η ακόλουθες: OKE, MOCHA, SQA (Scalable Question Answering) και STREAML (Stream Machine Learning). Η OKE και η MOCHA βασίστηκαν στις εκστρατείες που διοργανώθηκαν στο πλαίσιο του ESWC 2017 και βρίσκονται σε εξέλιξη. Το κύριο task του SQA Open Challenge είναι να αξιολογήσει την δυνατότητα των συστημάτων να επεξεργάζονται αιτήματα χρηστών στηριζόμενα σε τεχνικές επεξεργασίας Semantic Web και τεχνικές συμπερασμού. Η εκστρατεία αξιολόγησης ανοιχτού τύπου Stream ML επικεντρώνεται στο πρόβλημα της αυτόματης ανίχνευσης ανωμαλιών για τον κατασκευαστικό εξοπλισμό.

Εκτός από αυτές τις εκστρατείες αξιολόγησης που στοχεύουν κυρίως στην ώθηση των συστημάτων στα “όριά” τους, το HOBBIT οργανώνει μια ανοιχτή πρόσκληση για πρότυπα αξιολόγησης (benchmarks) που θα ενσωματωθούν στην πλατφόρμα έτσι ώστε πλήθος ατόμων να μπορούν να συγκρίνουν τα συστήματά τους με ένα τυποποιημένο hardware που παράγει συγκρίσιμα αποτελέσματα.

Πληροφορίες για το Project HOBBIT, την πλατφόρμα HOBBIT, τα πρότυπα αξιολόγησης και τις εκστρατείες αξιολόγησης μπορείτε να βρείτε εδώ: <https://project-hobbit.eu>