

OKE 2017

Open Knowledge Extraction Challenge

Presenter: Michael Röder



OKE @ ESWC 2017
Portoroz, Slovenia



HOBBIT
Holistic Benchmarking
of Big Linked Data



Horizon 2020, GA No 688227

May 30th, 2017

René Speck, University of Leipzig, Germany

Michael Röder, University of Leipzig, Germany

Ricardo Usbeck, University of Leipzig, Germany

Axel-Cyrille Ngonga Ngomo, University of Paderborn, Germany

Horacio Saggion, Universitat Pompeu Fabra, Spain

Luis Espinosa-Anke, Universitat Pompeu Fabra, Spain

Sergio Oramas, Universitat Pompeu Fabra, Spain

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

- Knowledge Extraction Challenge
- 4 Tasks
 - Focused NE Identification and Linking
 - Broader NE Identification and Linking
 - Focused Musical NE Recognition and Linking
 - Knowledge Extraction
- 3rd iteration
- Carried out using HOBBIT benchmarking platform
- Format: NLP Interchange Format (NIF), RDF/OWL-based

- Goal: Identification of NEs and Disambiguation to DBpedia
- 3 DBpedia Ontology classes (**Person, Place, Organisation**)

Florence May Harding studied at a school in **Sydney**, and with **Douglas Robert Dundas**, but in effect had no formal training in either botany or art.

- `dbr:Florence_May_Harding`
- `dbr:Sydney`
- `oke:Douglas_Robert_Dundas`

- Scenario **A**:
 - curated training and evaluation data
 - contains emerging entities
- KPIs
 - Precision
 - Recall
 - F1-measure
- Scenario **B**:
 - synthetically generated data
 - implemented as stress test
- KPI
 - β = F1-score points per second

- Goal: Identification of NEs and Disambiguation to DBpedia
- 12 DBpedia Ontology classes and subclasses
Activity, Agent, Award, Disease, EthnicGroup, Event, Language,
MeanOfTransportation, PersonFunction, Place, Species Animal, Work
- Scenarios and measures as in Task 1

- Goal of Task 3 A: Identification and Typing of Musical NEs
- 3 Music Ontology classes (**MusicArtist**, **MusicalWork** and **SignalGroup**)
- Goal of Task 3 B: Disambiguation to MusicBrainz
- KPIs
 - Precision
 - Recall
 - F1-measure

- Goal: Knowledge Extraction and Formalization in RDF triples
 - 21 owl:ObjectProperty (**almaMater**, **author**, **award**, ...)
- KPI: comparison of the received graph with DBpedias Knowledge Graph

- ADEL (Plu et al.)
- HmaraNER (Asif et al.)
(canceled participation)
- Baseline
 - FOX (Speck et al.)
 - FRED, FREME and Spotlight
Task 1 – 2 A, 3 A + B

- Second half of 2017: **Open** challenge
 - Easier participation
 - Online leader board
 - (Planned) winner announcement in February 2018
- Nearer future:
Poster session
Results at the closing session
- Now: ADEL participant presentation

Thank You!
Questions?