

supervised by Aidan Hogan and Barbara Poblete

[illegible]

Example annotations produced by four EL systems: AIDA (*A*), Babelfy (*B*), DBpedia Spotlight (*D*) and TAGME (*T*).

Joseph Walter “Joe” Jackson, was a steelworker at U.S. Steel. In an interview
D D|T B|D A|B|T D|T

with Martin Bashir for the 2003 documentary Living with Michael Jackson,
A|D|T D B|T A|B|D

Jackson recalled that Joe often sat in a chair with a belt in his hand as he and
A|B A T D|T T

his siblings rehearsed.
D|T

Dataset	Languages
AIDA/CoNLL-Complete	EN
KORE50	EN
IITB	EN
ACE2004	EN
AQUAINT	EN
MSNBC	EN
DBpedia Spotlight	EN
N3-RSS 500	EN
Reuters 128	EN
Wes2015	EN
News-100	DE
Thibaudet	FR
Bergson	FR
SemEval 2015 Task 13	EN,ES,IT
DBpedia Abstracts	DE,EN,ES,FR,IT,JA,NL
MEANTIME	EN,ES,IT,NL
VoxEL _R	DE,EN,ES,FR,IT
VoxEL _S	DE,EN,ES,FR,IT

- Multilingual approaches (without machine translation)
- Best configuration (using machine translation)
- Translating to their best language



Tool	Functionality	Demo	Source Code	Visualization
QRTool	NIF-Dataset Creation	✗	✗	
BENGAL	Automatic NIF Creation	✓	✓	
Eaglet	NIF-Dataset Validation	✗	✓	
GERBIL	Benchmark Framework	✓	✓	✗
Orbis	Benchmark Framework	✗	✗	✓
NiFify	NIF-Dataset Creation / Validation / Benchmarking ✓ Demo ✓ Source Code ✓ Visualization			

The present Ph.D. work proposes to address a variety of open questions regarding the evaluation of EL systems. In this direction, the following research questions are being or will be addressed:

1. What should Entity Linking link?
 - (a) How can we define the goal of the EL task?
 - (b) Is consensus possible on the definition of an “entity”?
 - (c) If not, how can we define benchmark EL datasets and what metrics can we use to reflect the lack of consensus?
2. How well do EL systems perform in multilingual settings?
 - (a) How can we compare EL performance across languages?
 - (b) How do EL systems perform for different languages?
 - (c) Why do results differ across languages?
 - (d) Are multilingual EL approaches necessary with recent improvements in machine translation?
3. How can we adapt and configure EL techniques for different applications?
 - (a) What are the specific applications for EL?
 - (b) What are the different settings that can be considered?
 - (c) Which techniques work best under what assumptions?
 - (d) How can existing EL techniques be best configured to meet the needs of a particular application setting?