

DBPEDIA-ENTITY V2:

A TEST COLLECTION FOR ENTITY SEARCH

Faegheh Hasibi

Norwegian University of
Science and Technology

Fedor Nikolaev

Wayne State University

Chenyan Xiong

Carnegie Mellon University

Krisztian Balog

University of Stavanger

Svein Erik Bratsberg

Norwegian University of
Science and Technology

Alexander Kotov

Wayne State University

Jamie Callan

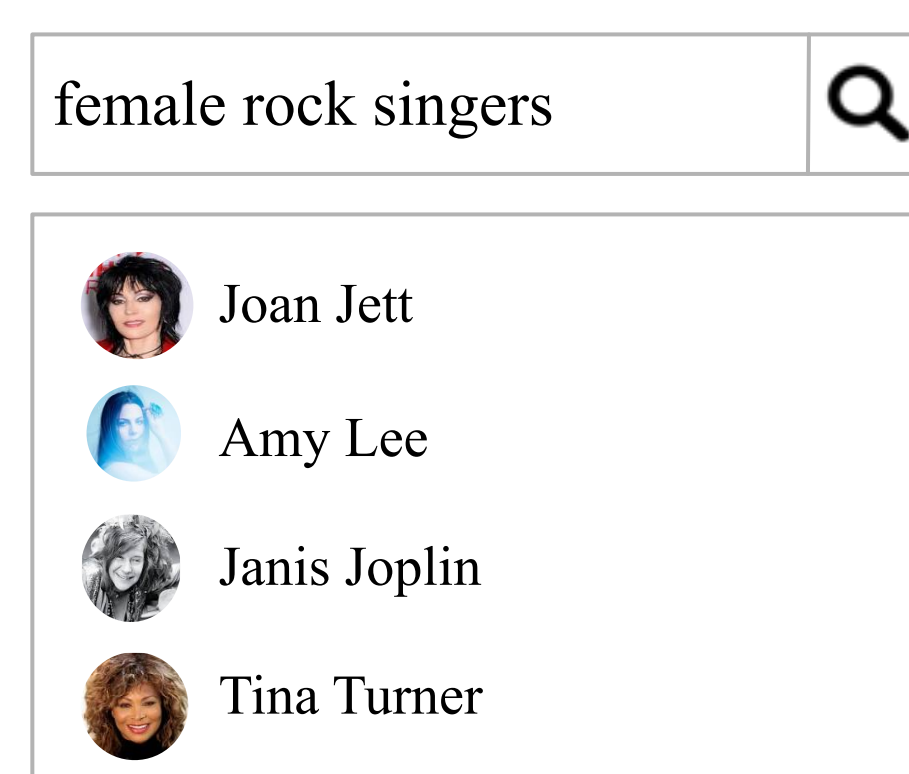
Carnegie Mellon University

DBpedia



+

Entity search



→

**DBpedia-Entity
collection**



SUMMARY

- *DBpedia-entity v1** has been used as a standard test collection for entity search
- *DBpedia-entity v2* is the updated version
 - Based on a recent DBpedia dump (2015-10, with 6.2M entities)
 - Graded relevance judgments and explicit non-relevant judgments for all queries
- Consistent in collecting relevance judgments (vs. heterogeneous judgments in v1)
 - Results pooled from baselines, state-of-the-art retrieval methods, and old qrels
 - Collecting judgments using crowd workers as well as expert annotators
- **What's next?** Show that your model can outperform existing approaches
 - According to our results, recent models can hardly beat traditional baselines

QUERIES AND RESULTS

Query Categories	#queries	Method	v1		v2	
			MAP	P@10	MAP	nDCG@10
SemSearch ES	113	BM25	0.0884	0.0971	0.1893	0.3582
named entity queries; e.g., "brooklyn bridge"		LM	0.1709	0.1837	0.3144	0.4182
INEX-LD	99	SDM	0.1860	0.1880	0.3259	0.4185
keyword queries; e.g., "electronic music genres"		LTR	0.1723	0.1831	0.2446	0.3464
ListSearch	115	BM25-CA	0.2067	0.2056	0.3265	0.4231
list queries; e.g., "professional sports teams in Philadelphia"		FSDM	0.2069	0.2039	0.3279	0.4267
QALD-2	140	BM25F-CA	0.2088	0.2126	0.3361	0.4378
natural language queries; e.g., "Who is the mayor of Berlin?"		FSDM+ELR	0.2210	0.2089	0.3295	0.4335
Total	467					

* K. Balog and R. Neumayer. 2013. "A Test Collection for Entity Search in DBpedia." In SIGIR '13.

