



Yongpan Sheng¹, Zenglin Xu¹, Yafang Wang², Xiangyu Zhang¹, Jia Jia², Zhonghui You¹ and Gerard de Melo³

¹ University of Electronic Science and Technology of China, China; ² Shandong University, China; ³ Rutgers University, USA

S_1

Fact Extraction

- Document Ranking
- Coreference Resolution
- Open-Domain Knowledge Extraction

S_2

Fact Filtering

- Fact filtering algorithm

This is achieved by optimizing for a high degree of coherence between facts with high confidence

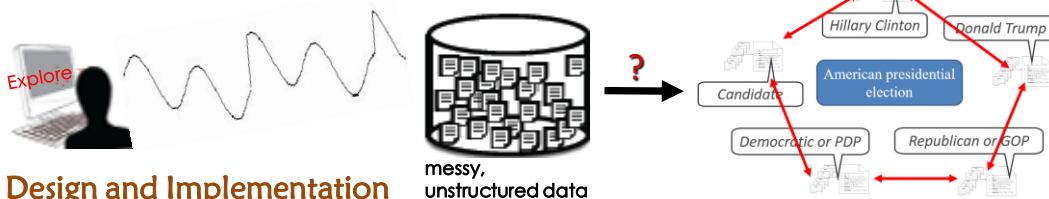
S_3

Conceptual Graph Construction

- Merge potential entities and concepts, whose labels present equivalent meanings
- A heuristic algorithm to find a full graph that is connected and satisfies the size limit of 25 concepts

Motivation

This is challenging to mine and organize meaningful concepts and their semantic connections from a set of related documents under the same topic in the information extraction.



Design and Implementation

A large collection of news articles

Doc1: 82 percent of leaned Democrats say Registered voters'd support a clear Clinton...
Doc2: The hypothetical contest compares to a clear Clinton lead among all adults,...
Doc3: Donald Trump, his rivals don't need an apprentice in the White House, his rivals have one right now...
Doc4: Wisconsin Governor Scott Walker told Donald Trump...
...

S_1

Fact candidates as subject-predicate-object triples

0.95 (Billionaire Donald Trump; as a powerful competitor of; Hillary Clinton)
0.90 (Ben Carson; has suspended; 2016 US presidential race)
0.92 (Hillary Clinton campus; may be affecting; Trump)
0.97 (68 percent of Republicans; presently voted; Billionaire Donald Trump)
0.94 (e-mail scandal; might be effecting; 68 percent of Republicans)
...

S_2

Representative facts

0.95 (Billionaire Donald Trump; as a powerful competitor of; Hillary Clinton) ✓
0.90 (Ben Carson; has suspended; 2016 US presidential race)
0.97 (68 percent of Republicans; presently voted; Billionaire Donald Trump) ✓
0.94 (e-mail scandal; might be effecting; 68 percent of Republicans)
...

S_3

Conceptual graph

