SEARCH ENGINE ENTITY CARDS(EC's) What's in the box? And its impact on users.

1.BACKGROUND

- With EC's we get more navigability
- Display dense information
- Links to possible useful pages
- Query specific content.
- Research into how to make EC widget for SearchX

2.RESEARCH METHOD

- Information Retrieval
- Entity Tagging
- Entity Linking
- DBPedia-chosen as Knowledge Base
- DBPedia Spotlight

ENTITY CARD FOR "Google"



Technology company



Google LLC is an American multinational technology company that specializes in Internet-related services and products, which include online advertising technologies, a search engine, cloud computing, software, and hardware. Wikipedia

CEO: Sundar Pichai (Oct 2, 2015-) Trending

Founded: September 4, 1998, Menlo Park, California, United States

Parent organization: Alphabet Inc.

Headquarters: Mountain View, California, United States

Subsidiaries: YouTube, Google China, YouTube TV, Fitbit, MORE

Founders: Larry Page, Sergey Brin

Signs

Ip address

Contact: Yash Kalia Y.Kalia@student.tudelft.nl Supervised by: Claudia Hauff





V

V

3.RESULTS & CONCLUSIONS

)BPedia DBPedia Spotlight as tool • Bing for search results

4.FUTURE WORK

 Implementation Evaluating general web queries on Spotlight Evaluation of finished product

SEARCH ENGINE ENTITY CARDS(EC's) Content of EC's and its impact on users. Yash Kalia, Y.Kalia@student.tudelft.nl | Supervisor: Claudia Hauff



1.BACKGROUND

- EC's are a common tool in modern search engines.
- With EC's the users get easier navigability. Dense information about topics can be displayed in a easy to read manner. EC's also provide images and links to useful related webpages for quick lookup and less searching.
- Research is in how to make EC widget for SearchX which is a search engine for research purposes.

2.RESEARCH INTO

 Information Retrieval
 Entity Tagging and Entity
Linking and tools like
Tagme, Stanford NER and
Spotlight.
 Suitable knowledge
bases(KB) for entity facts and
a linking/tagging tool which
works well the knowledge
base.
 Fact Ranking techniques
based on relevance and
importance

CONCLUSIONS • DBPedia is a suitable and rich enough knowledge base and offers a linking and tagging tool DBPedia Spotlight which is suitable for entity extraction. Preferred over Wlkipedia because its's a structured KB and offers pinpoint information with tags.

Albert Einstein About 102.000.000 results (0.91 seconds) https://en.wikipedia.org > wiki > Albert_Einstein -Albert Einstein - Wikipedia Albert Einstein (/ ainstain/ EYEN-styne; German: [albest '?ainftain] (listen); 14 March 1879 -18 April 1955) was a German-born theoretical physicist, widely acknowledged to be one of the reatest physicists of all time. Doctoral advisor: Alfred Kleiner Fields: Physics, philosophy Other academic advisors: Heinrich Friedrich... Influenced: Virtually all modern physics Religious and philosophical... • Einstein family • I Am Albert Einstein • Hans Albert People also ask Fact Ranking What did Albert Einstein say about death RELEVANT FACTS How smart was A What are 5 facts about Albert Einstei What did Albert Einstein Discover? Summarization https://www.nobelprize.org > prizes > physics > biograp ... • Albert Einstein - Biographical - NobelPrize.org Albert Einstein grew up in Munich, where his father founded an electrical eng

3.RESULTS &

4.FUTURE WORK

- screen.
- SearchX EC's

Search Engine Result Page(SERP) with EC and Top Docs

X 🏮 🔍 Settings Tools Albert Einstein Theoretical physicist

Albert Einstein was a German-born theoretica physicist, widely acknowledged to be one of the st physicists of all time. Einstein is known fo ping the theory of relativity, but he also made moortant contributions to the development of the theory of quantum mechanics. Wikipedia

More images

Born: March 14, 1879, Ulm, Germany Died: April 18, 1955, Penn Medicine Princeton Medical Center, New Jersey, United States Children: Eduard Einstein, Hans Albert Einstein Lieserl Einstein

 Implementation and further research into fact ranking algorithms and how to display the chosen facts on th

 Offline evaluation of the final product using questionnaires to rate results from different search engines and



SEARCH ENGINE ENTITY CARDS(EC's) Content of EC's and its impact on user Yash Kalia, Y.Kalia@student.tudelft.nl | Supervisor: Claudia Hau



1.BACKGROUND

- EC's are a common tool in modern search engines.
- With EC's the users get easier navigability. Dense information about topics can be displayed in a easy to read manner. EC's also provide images and links to useful related webpages for quick lookup and less searching.
- Research is in how to make EC widget for SearchX which is a search engine for research purposes.

2.RESEARCH INTO

- Information Retrieval • Entity Tagging and Entity Linking and tools like Tagme, Stanford NER and Spotlight. • Suitable knowledge bases(KB) for entity facts and a linking/tagging tool which
- works well the knowledge base. • Fact Ranking techniques
- based on relevance and importance

CONCLUSIONS DBPedia is a suitable and rich enough knowledge base and offers a linking and tagging tool DBPedia Spotlight which is suitable for entity extraction. Preferred over Wlkipedia because its's a structured KB and offers pinpoint information with tags.

3.RESULTS &

- screen.
- SearchX EC's

rs	
Jff	

Search Engine Result Page(SERP) with EC and Top Docs

X 🎙 Q		
News 🛯 Books 🕞 Videos 🚦 More Settings	Tools	
(0,91 seconds)		
i→Albert_Einstein 💌		
kipedia		
EYEN-styne; German: [ˈalbɛʁt ˈʔaɪnʃtaɪn] (listen); 14 March 1879 –		
n-born theoretical physicist, widely acknowledged to be one of the		
einer Fielder Dhveice philosophy		
einer Fields: Physics, philosophy Heinrich Friedrich Influenced: Virtually all modern physics		
· Einstein family · I Am Albert Einstein · Hans Albert	More	
n en	Albert Einstein	
	Theoretical physicist	
say about death?	✓ Albert Einstein was a German-born theoretical	
nstein?	physicist, widely acknowledged to be one of the	
	greatest physicists of all time. Einstein is known	
lbert Einstein?	 developing the theory of relativity, but he also ma important contributions to the development of the 	
Diagonal	theory of quantum mechanics. Wikipedia	
Discover?	Born: March 14, 1879, Ulm, Germany	
F		
	Died: April 18, 1955, Penn Medicine Princeton Medical Center, New Jersey, United States	
prizes > physics > biograp 💌	Children: Eduard Einstein, Hans Albert Einstein,	
graphical - NobelDrize org	officient. Lucaru Linstein, Hans Albert Einstein.	

Lieserl Einstein



 Implementation and further research into fact ranking algorithms and how to display the chosen facts on th

 Offline evaluation of the final product using questionnaires to rate results from different search engines and



SEARCH ENGINE ENTITY CARDS Yash Kalia, Y.Kalia@student.tudelft.nl | Supervisor & Responsible Professor: Claudia Hauff

BACKGROUND & RESEARCH

- EC's offer easier navigability and are easy to read manner.
- EC's also display related images and links to useful related webpages.
- The primary challenge with an entity card widget is **entity** retrieval.
- **Research Question**: "What entity and facts to display in an Entity Card given a query?".
- Entity Ranking is used to sort the retrieved entities in order of highest relevance to the query.
- Ranking function-

 $MentionFrequencyIdf = tf_e * log(N/df(e))$

EVALUATION & RESULTS

DBPedia Spotlight Retrieval Results

only entity mentioned.

DBPedia Spotlight Accuracy Entity Retrieval and Ranking(Relevance) Fact Retrieval(Relevance) 7/30 2.29/5 1.91/5

Both evaluations below done with_~100 queries with 3 evaluators: **Entity Retrieval Results**

Improvement over using only the query.

Fact Retrieval Results

- "On average **Not relevant**" to "Slightly relevant" bracket in Likert Scale.
- DBPedia.



"what is

• Relevant entity was retrieved only when entity was mentioned by name and was the

• On average "Slightly relevant " to "Moderately relevant" bracket in Likert Scale.

• The fact retrieval algorithm does not access specific information about the entity from

Development of EC Algorithm

INSIGHTS & FUTURE WORK Insights • Query is insufficient to decide which entity should be displayed. Web documents provide better results. • Best results for queries which pertain to clear unambigous entities. • The fact retrieval works well for which the abstract of the entity has the relevant facts. **Future Work** • Extract different entity summaries using useful keywords from the query. Retrieving not only text but also images and hyperlinks to related entities.