

SEARCH ENGINE ENTITY CARDS(EC's)

What's in the box? And its impact on users.

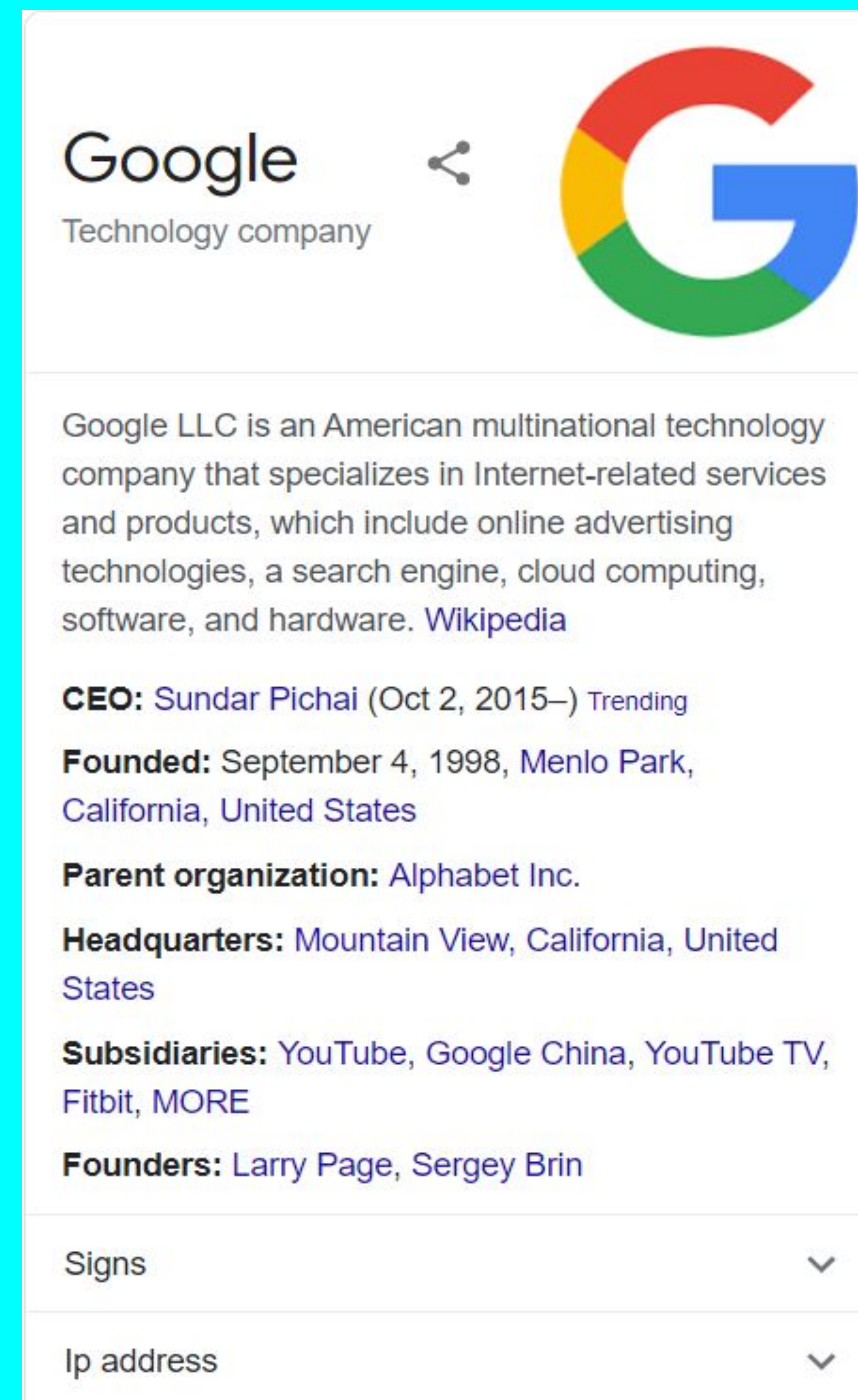
1. BACKGROUND

- 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"



The screenshot shows an entity card for Google. At the top, it displays the word "Google" and "Technology company" next to the Google logo. Below this, there is a paragraph of text: "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". Underneath, there are several key facts: "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", and "Founders: Larry Page, Sergey Brin". At the bottom, there are two expandable sections: "Signs" and "Ip address", each with a downward arrow.

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



3. RESULTS & CONCLUSIONS

- DBPedia
- DBPedia Spotlight as tool
- Bing for search results

4. FUTURE WORK

- Implementation
- Evaluating general web queries on Spotlight
- Evaluation of finished product

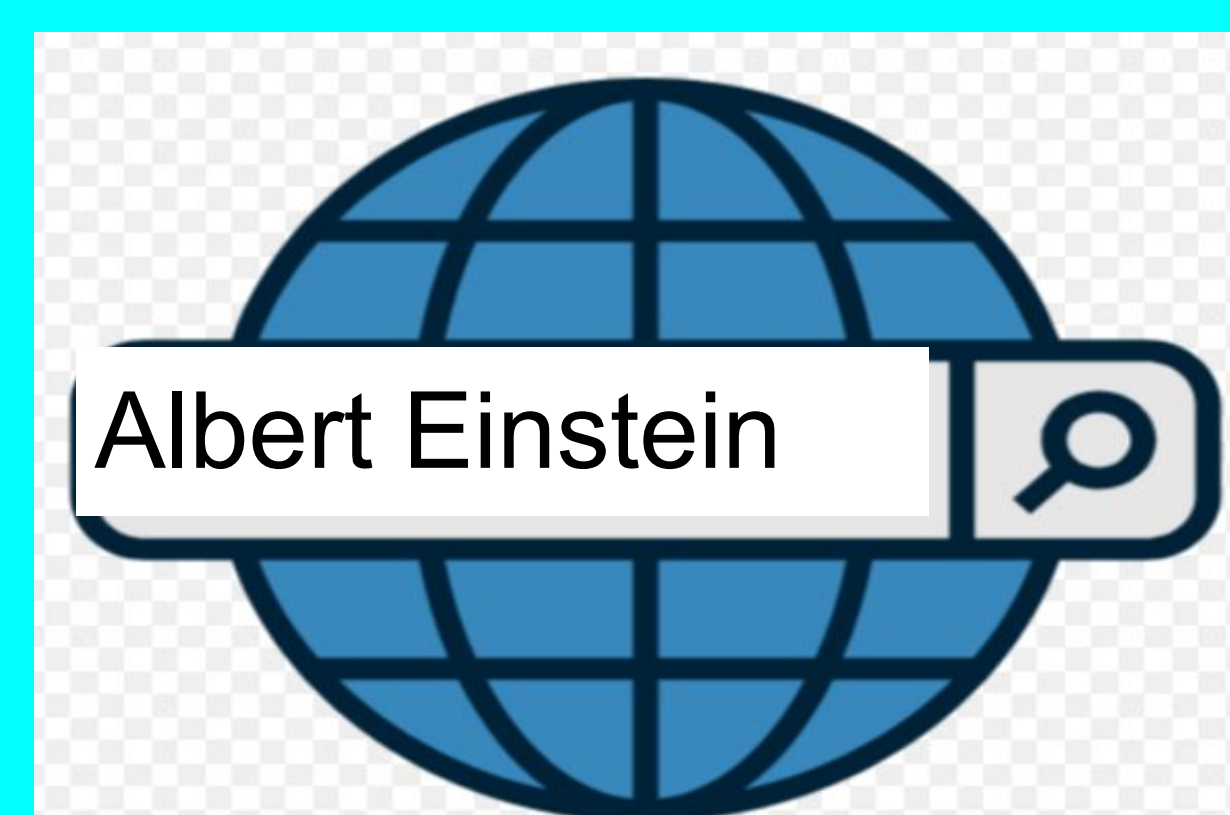
SEARCH ENGINE ENTITY CARDS(EC's)

Content of EC's and its impact on users.

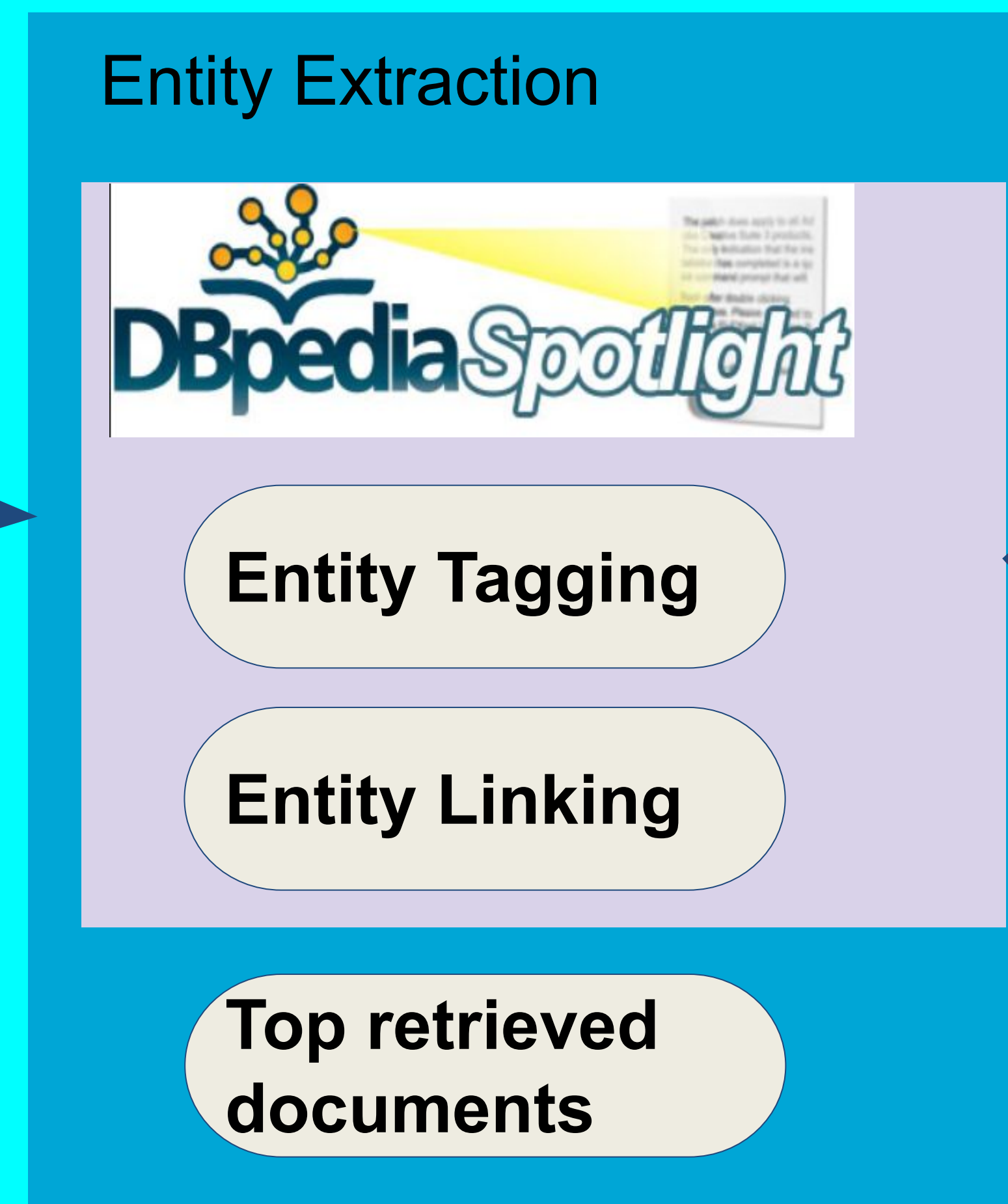
Yash Kalia, Y.Kalia@student.tudelft.nl | Supervisor: Claudia Hauff

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

APPROACH TO IMPLEMENT EC



QUERY

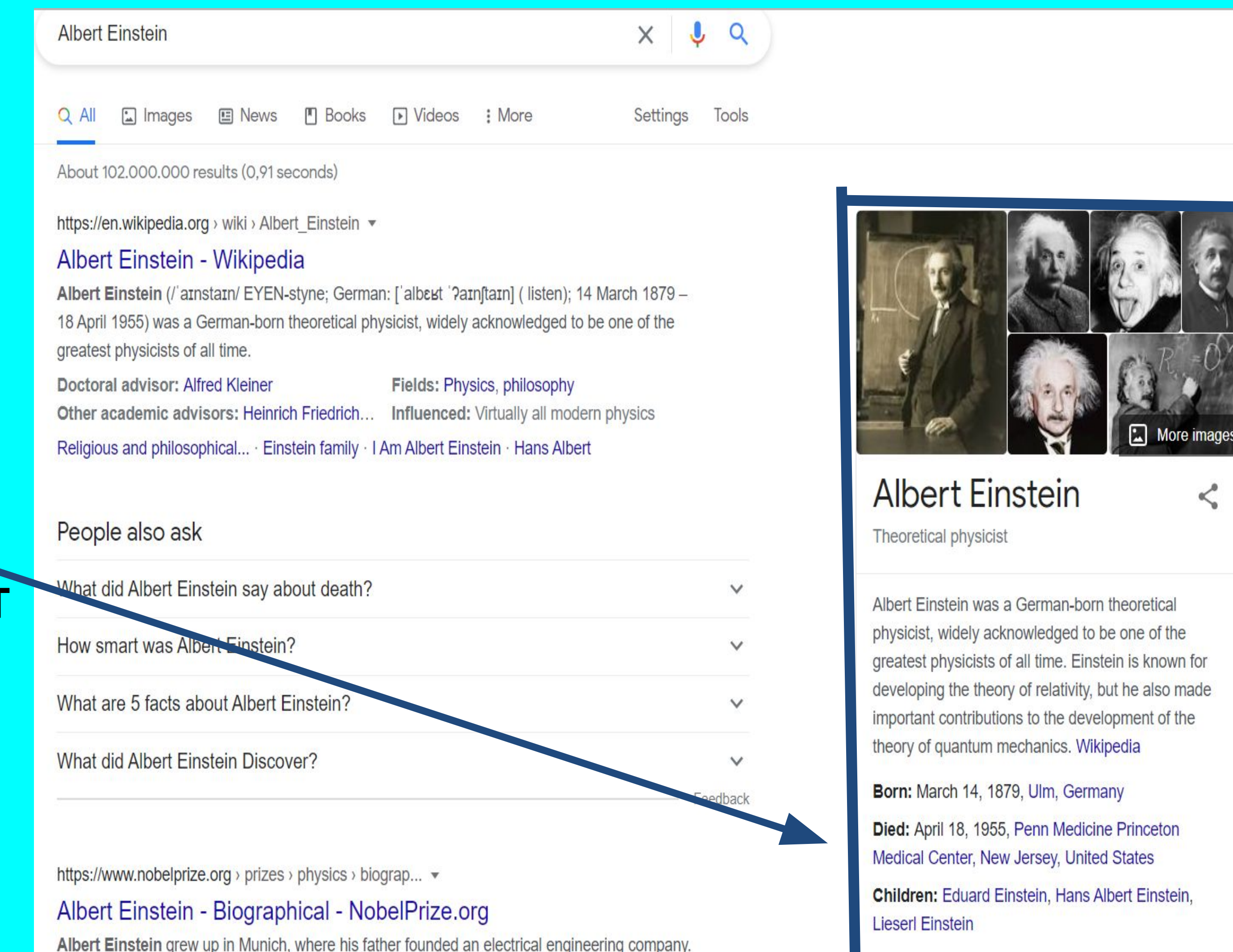


ENTITY

https://dbpedia.org/page/Albert_Einstein



RELEVANT FACTS



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

3. RESULTS & 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 Wikipedia because its a structured KB and offers pinpoint information with tags.

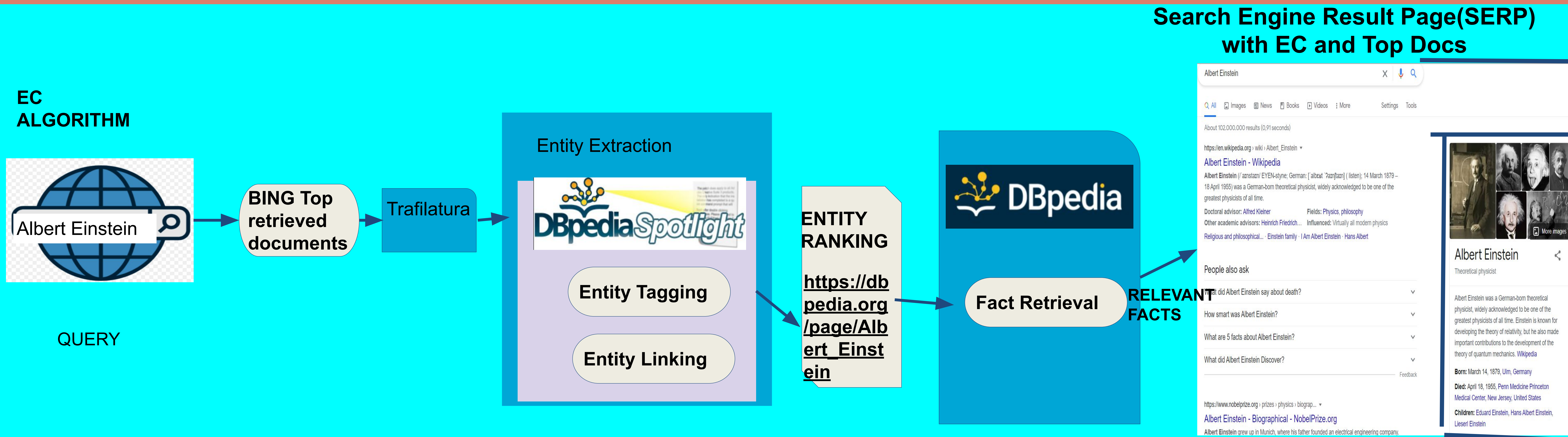
4. FUTURE WORK

- Implementation and further research into fact ranking algorithms and how to display the chosen facts on th screen.
- Offline evaluation of the final product using questionnaires to rate results from different search engines and SearchX EC's

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

3. RESULTS & 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 Wikipedia because its's a structured KB and offers pinpoint information with tags.

4. FUTURE WORK

- Implementation and further research into fact ranking algorithms and how to display the chosen facts on th screen.
- Offline evaluation of the final product using questionnaires to rate results from different search engines and SearchX EC's

SEARCH ENGINE ENTITY CARDS

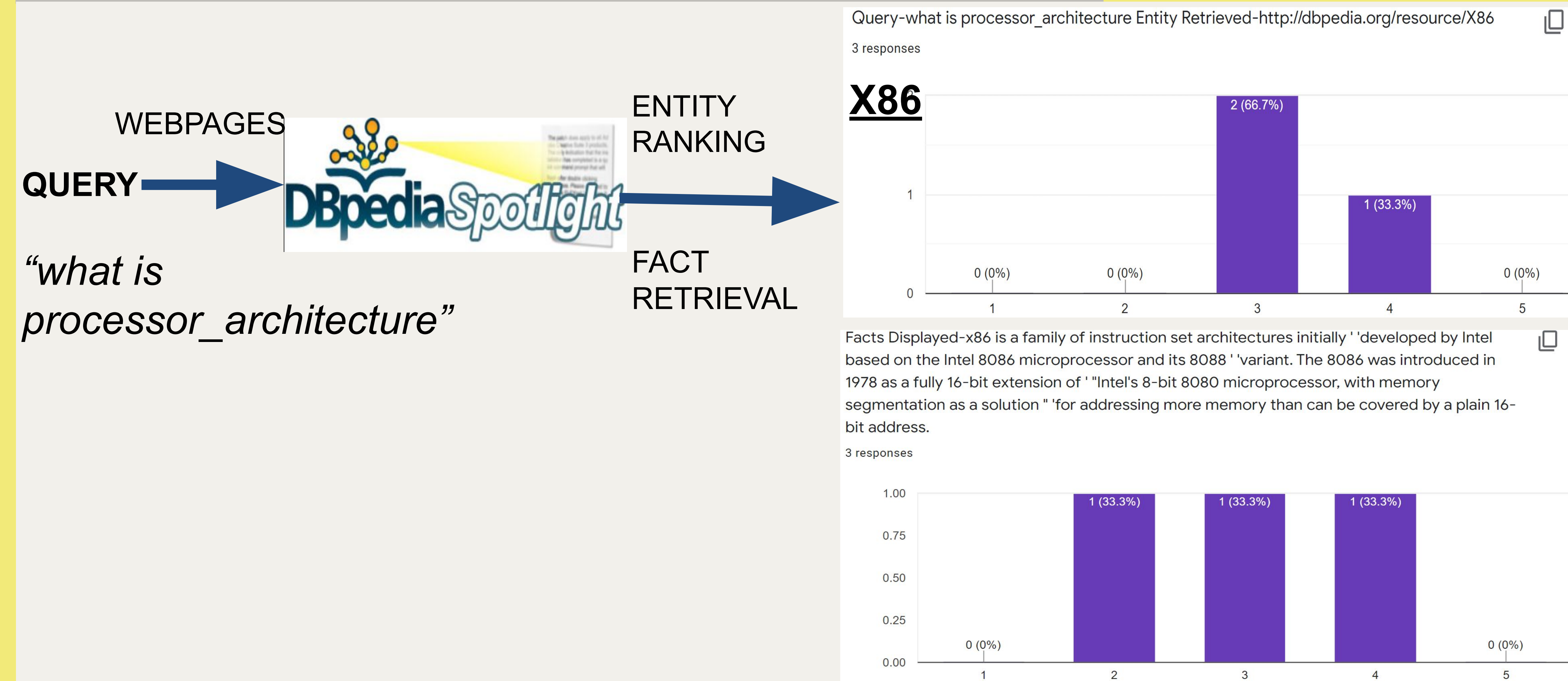
Yash Kalia, Y.Kalia@student.tudelft.nl | Supervisor & Responsible Professor: Claudia Hauff

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-

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

Development of EC Algorithm



EVALUATION & RESULTS

DBpedia Spotlight Retrieval Results

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

- On average **"Slightly relevant"** to **"Moderately relevant"** bracket in Likert Scale. Improvement over using only the query.

Fact Retrieval Results

- "On average **Not relevant**" to **"Slightly relevant"** bracket in Likert Scale.
- The fact retrieval algorithm does not access specific information about the entity from DBpedia.

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 unambiguous 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.