

Do it your own (DIY) Jeopardy Question Answering System

André Freitas and Edward Curry

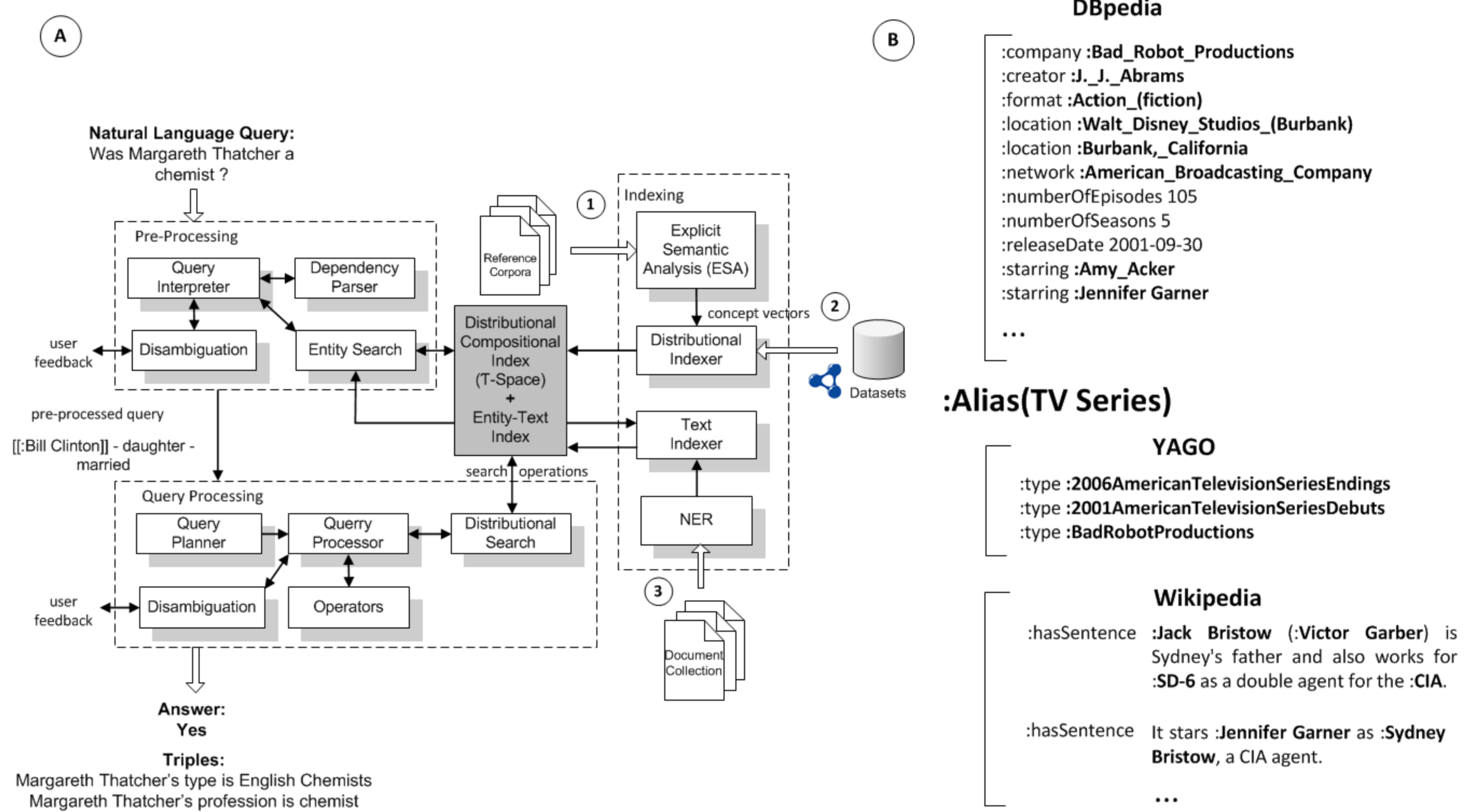
Problem

- Increasing dataset complexity and heterogeneity (Big Data and Web scenarios).
- How to support users querying/searching heterogeneous data without prior knowledge of its representation?

Objective

- Question Answering (QA) over Structured Unstructured Data scenarios.
- Do-it-yourself: customizable for different datasets.
- Support for schema-agnostic queries.

System Components



Treó

Queries over DBpedia



Search: Advanced Search Knowledge Base Admin

Treó Who is the daughter of Bill Clinton married to? Search

"Who is the daughter of Bill Clinton married to?"

Answer

- Chelsea Clinton spouse Marc Mezvinsky
- Bill Clinton child Chelsea Clinton
- Bill Clinton children Chelsea Clinton
- William Jefferson Blythe, Jr. child Bill Clinton
- Virginia Clinton Kelley child Bill Clinton
- Virginia Clinton Kelley children Bill Clinton



Queries over Wikipedia + DBpedia



Treó Sydney's dad, Jack, was a CIA double agent working against SD-6 on this Jennifer Garner show Search

○ Data ○ Vocabulary ○ Text Schema-free SPARQL

Sydney's dad, Jack, was a CIA double agent working against SD-6 on this Jennifer Garner show

Answer

Alias (TV series)

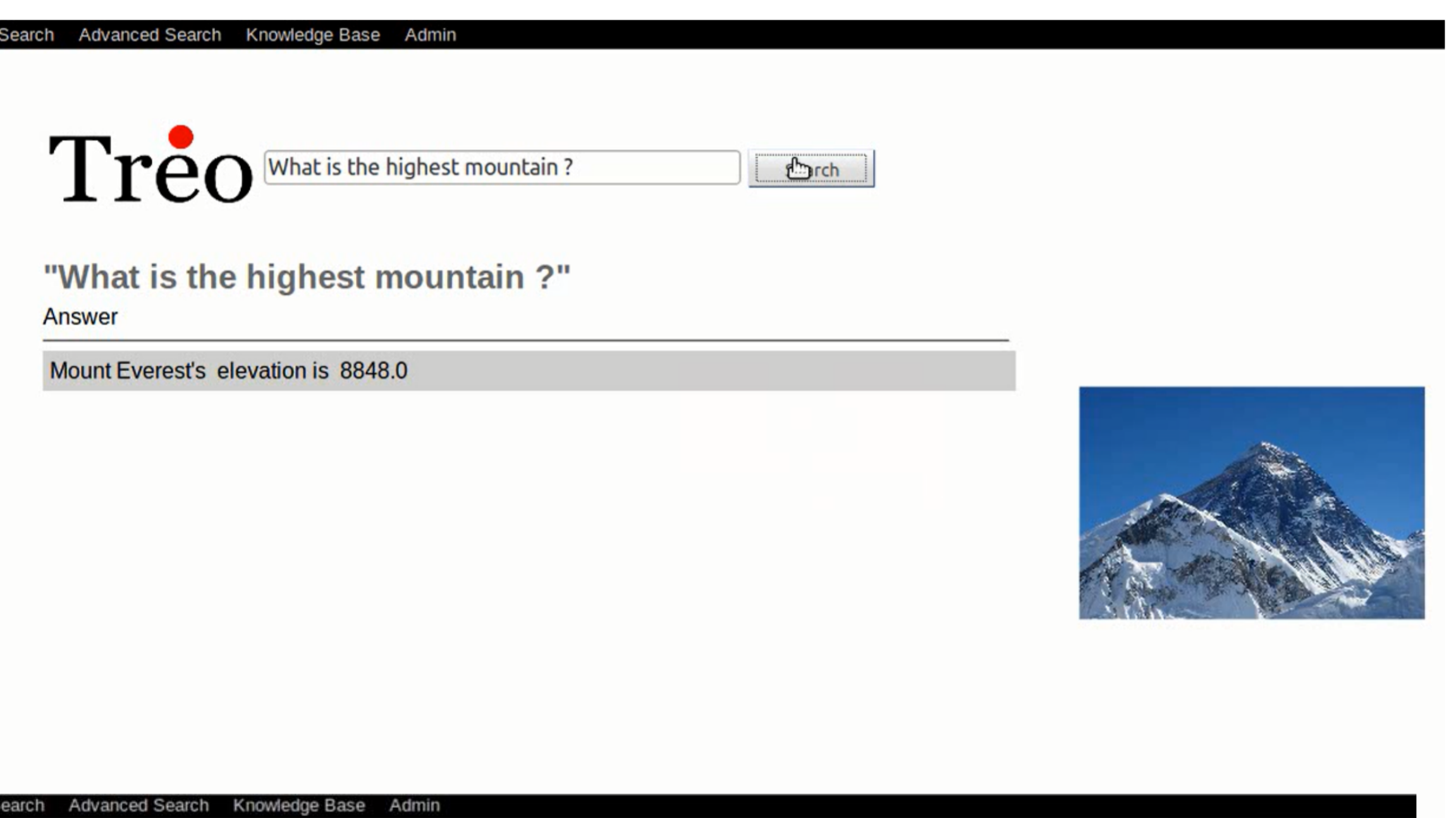
Treó After singing for his supper, he ate "white bread Search

○ Data ○ Vocabulary ○ Text Schema-free SPARQL

After singing for his supper, he ate "white bread and butter"

Answer

Little Tommy Tucker




Search: Advanced Search Knowledge Base Admin

Treó What is the highest mountain? Search

"What is the highest mountain?"

Answer

Mount Everest's elevation is 8848.0




Search: Advanced Search Knowledge Base Admin

Treó How tall is Claudia Schiffer? Search

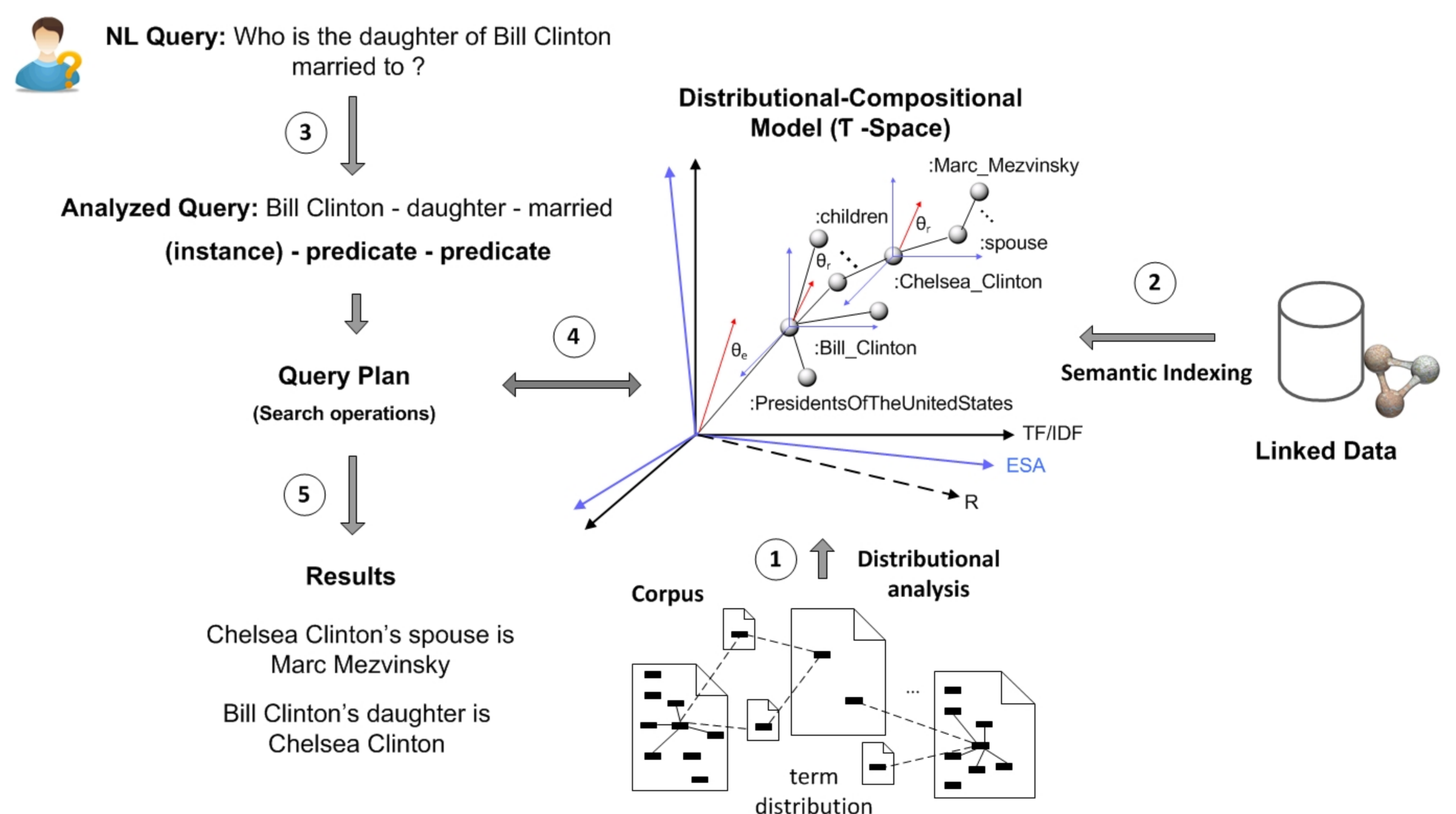
"How tall is Claudia Schiffer?"

Answer

Claudia Schiffer's height is 1.8



Query Processing



References

- A. Freitas, E. Curry, J. G. Oliveira, S. O'Riain, A Distributional Structured Semantic Space for Querying RDF Graph Data. International Journal of Semantic Computing (IJSC), 2012.
- A. Freitas, J. G. Oliveira, S. O'Riain, E. Curry, J. C. Pereira da Silva, Querying Linked Data using Semantic Relatedness: A Vocabulary Independent Approach. In Proceedings of the 16th International Conference on Applications of Natural Language to Information Systems (NLDB), 2011.

Acknowledgements

This work has been funded by Science Foundation Ireland under Grant No. SFI/08/CE/11380 (Lion-2).