



OŚRODEK PRZETWARZANIA INFORMACJI  
PAŃSTWOWY INSTYTUT BADAWCZY

# Assessment of historical figures using BabelNet

Szymon Roziewski, Maciej Kowalski

National Information Processing Institute, Natural Language Processing Laboratory, Warsaw

# Motivation

- | Building a simple Sentiment Analysis System.
- | Providing sentiment analysis of famous people based on BabelNet synsets. Actually, they could be fictional.
- | To have some fun :)



## Why assessment of historical figures ?

- | To get a fast sentiment feedback about people playing a role in History.
- | Resulting sentiment might be useful for another purposes: search engines, marketing companies etc.

*Always walk through life as if you have something new to learn and you will.*

*- Vernon Howard*



# Method

- | We used BabelNet for the core fo the system
- | BabelNet provides synset tree for a given entity
- | Each tree leaf carries short gloss of its own definition
- | We collect closest synsets to build a small corpus of text surrounding given term
- | We perform sentiment analysis based on sentiment dictionaries:
  - lexicon dictionary
  - SentiWordNet

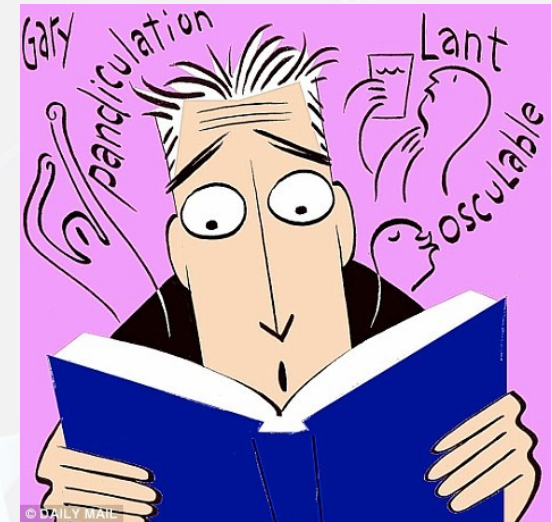
# BabelNet

- | Huge multilingual encyclopedic dictionary and ontology
- | Offers a large network of concepts and named entities with multilingual lexical coverage

# Sentiment Dictionaries

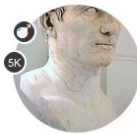
- Lexicon Dictionary – lexicon built at University of Illinois Chicago  
It is divided into positive and negative words
- SentiWordNet (Princeton) – a list of words described by two real numbers:  
the positive and negative impact

Basic forms of words obtained by using morfologik library.



# BabelNet query for Julius Caesar

## Noun



### Julius Caesar, Gaius Julius Caesar, Caesar

Conqueror of Gaul and master of Italy (100-44 BC)

ID: 00014550n | Named Entity



### Julius Caesar (play)

The Tragedy of Julius Caesar is a tragedy by William Shakespeare, believed to have been written in 1599.

ID: 03385292n | Named Entity



### Julius Caesar (1953 film)

Julius Caesar is a 1953 Metro-Goldwyn-Mayer film adaptation of the play by Shakespeare, directed by Joseph L. Mankiewicz, who also wrote the uncredited screenplay, and produced by John Houseman.

ID: 01908470n | Named Entity



### Julius Caesar (TV miniseries)

Julius Caesar is a 2002 mini-series about the life of Julius Caesar.

ID: 02827582n | Named Entity



### Julia Caesaris (sister of Julius Caesar), Julius Caesar (sisters of)

Julia is the name of two daughters of praetor Gaius Julius Caesar and Aurelia Cotta, the parents of dictator Gaius Julius Caesar.

ID: 03177074n | Named Entity

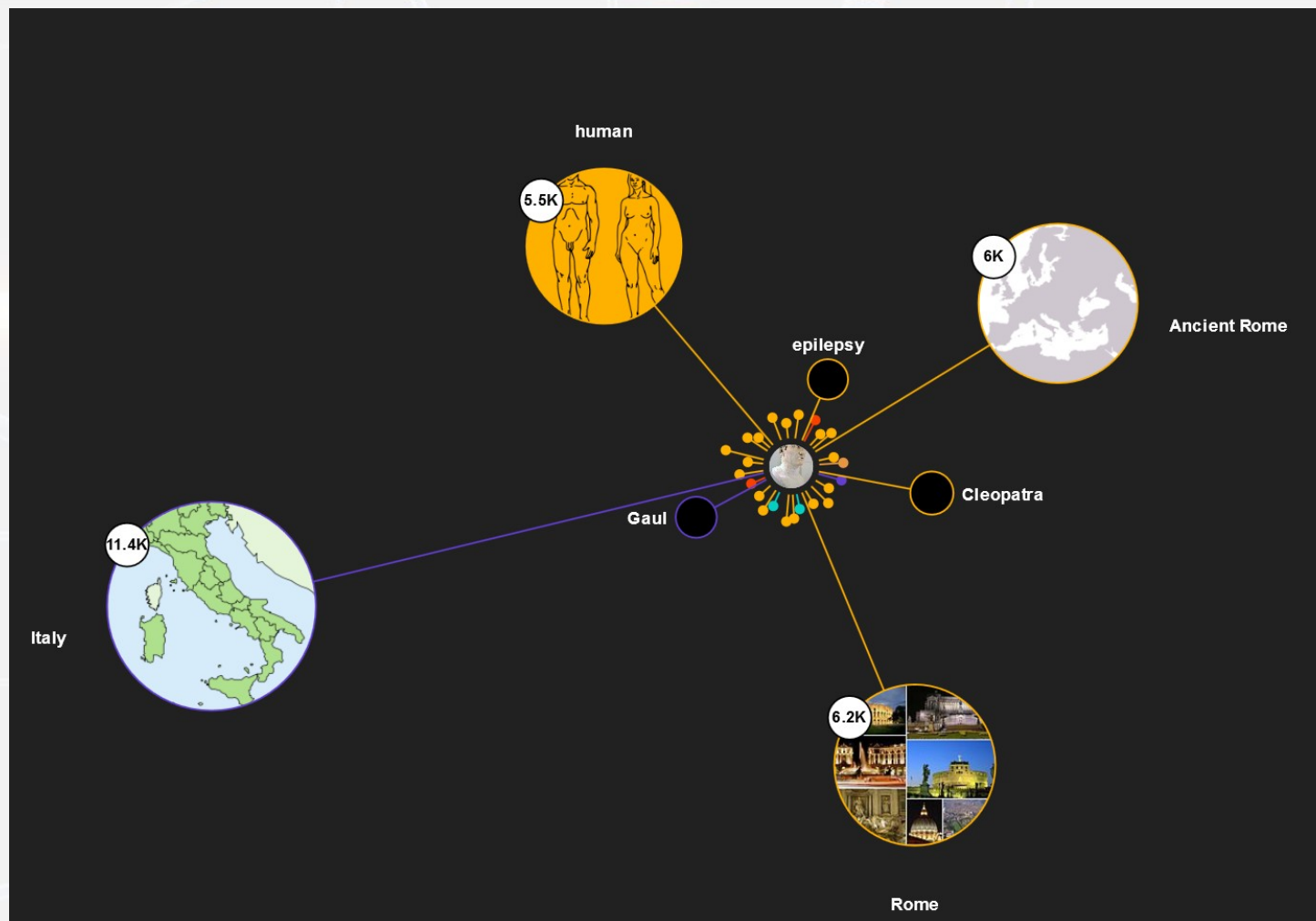


### Julius Caesar (judge)

Sir Julius Caesar was an English judge and politician who sat in the House of Commons at various times between 1589 and 1622.

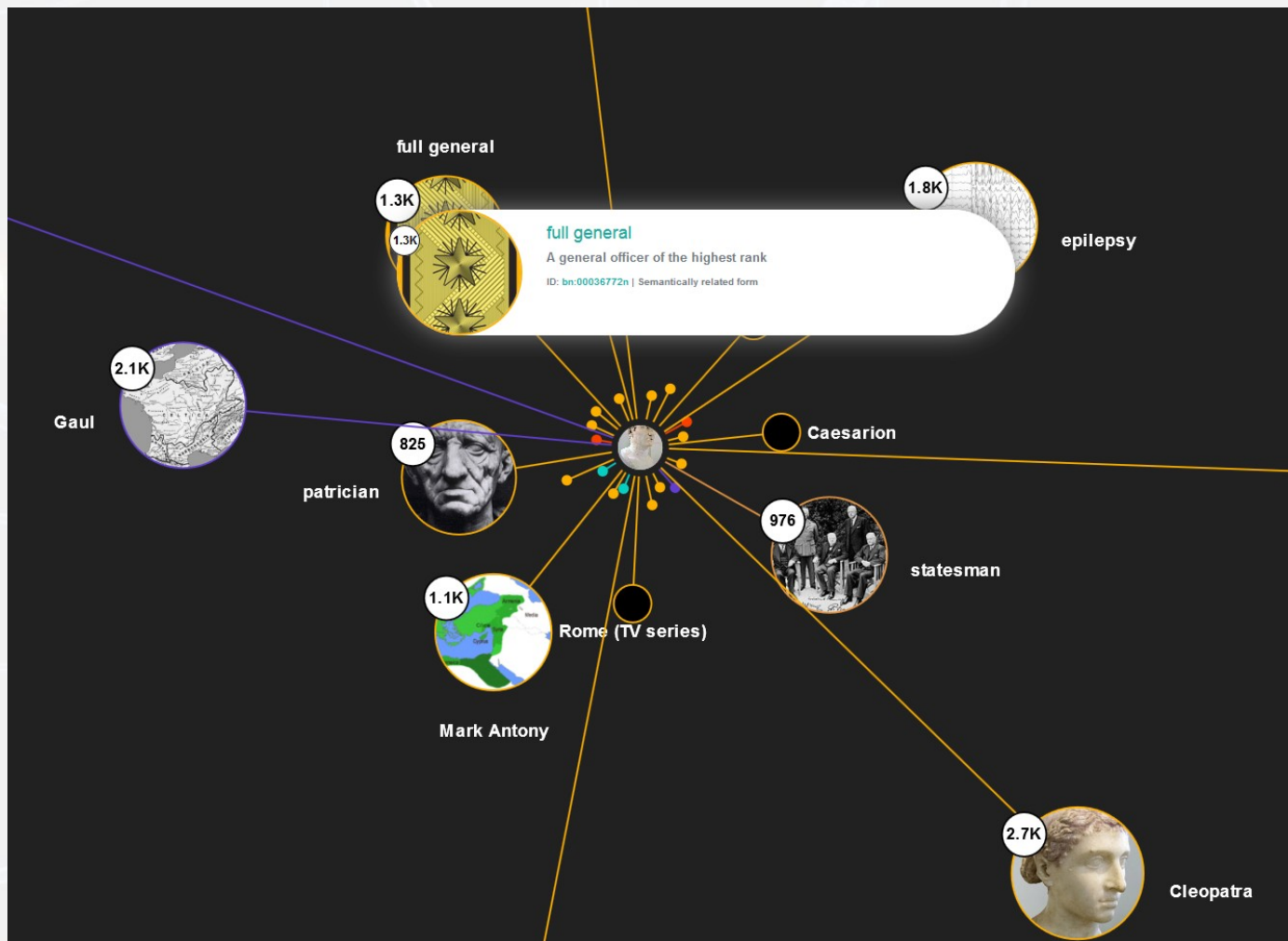
ID: 02545639n | Named Entity

# BabelNet query for Julius Caesar

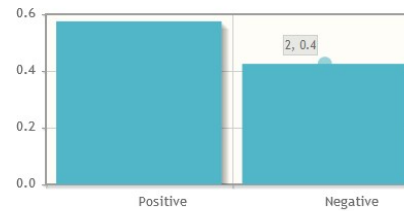
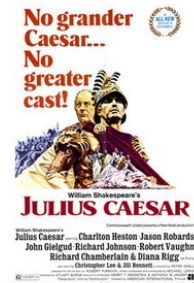
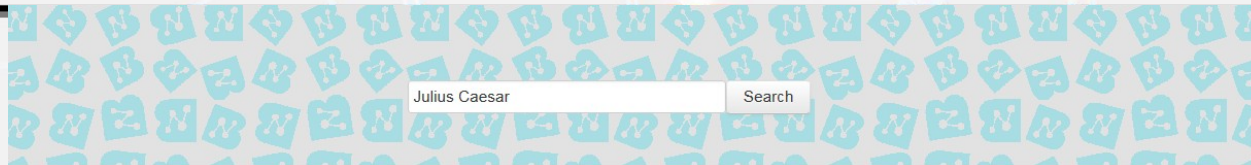




# BabelNet query for Julius Caesar



# Sentiment Analysis of Julius Caesar



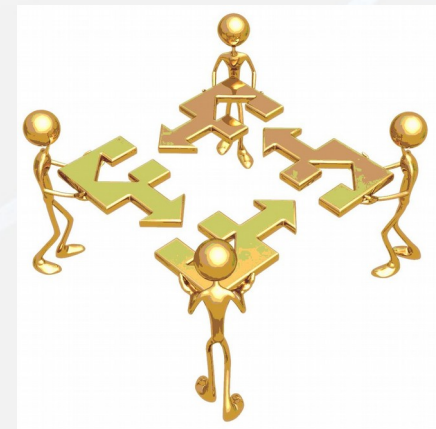
Positive (0.5748116254036604) vs. Negative (0.42518837459634024)

**Positive :)**

Gloss
44 AD, 1944, 2044, etc., the years
A kingdom and sovereign state in Europe comprising the four countries of England, Scotland and Wales in the island of Great Britain, and Northern Ireland in the island of Ireland. Official name: United Kingdom of Great Britain and Northern Ireland. Abbreviation: UK.
Metropolitan Anthony of Surozoh was best known as a writer and broadcaster on prayer and the Christian life.
A large city located on Lake Michigan in northeastern Illinois, USA.
Marek Juniusz Brutus, Marcus Iunius Brutus – polityk rzymski.
New York.
Marcus Junius Brutus Caepio, or simply Brutus, was a Roman senator of the late Roman Republic.
Henry IV, Part 2 and Henry V, in which he dies. He is later brought back in The Merry Wives of Windsor

# Implementation

- | The process has been automated by using Java along with BabelNet API, for quering the semantic network.
- | Parameters used for the synset selection:
  - language = EN
  - BabelPOS = NOUN
  - BabelSenseSource = WIKIDATA



# Sentiment Analysis Results

Historical Figure	Positive impact	Negative impact
Albert Einstein	0.523	0.477
Isaac Newton	0.7	0.3
Al Capone	0.304	0.696
Marylin Monroe	0.602	0.398
Billy the Kid	0.601	0.399
Julius Caesar	0.6	0.4
Cinderella	0.4776	0.5224

# Future Work

- | There were no gold-standard sets to compare with, we assume such method can be developed further.
- | Implementing disambiguating mechanism in order to select the exact synset carrying glosses related to historical figure in question.

# Thank you!



**Ośrodek Przetwarzania Informacji  
Państwowy Instytut Badawczy**

al. Niepodległości 188 b  
00-608 Warszawa

tel.: +48 22 570 14 00  
e-mail: [opi@opi.org.pl](mailto:opi@opi.org.pl)

