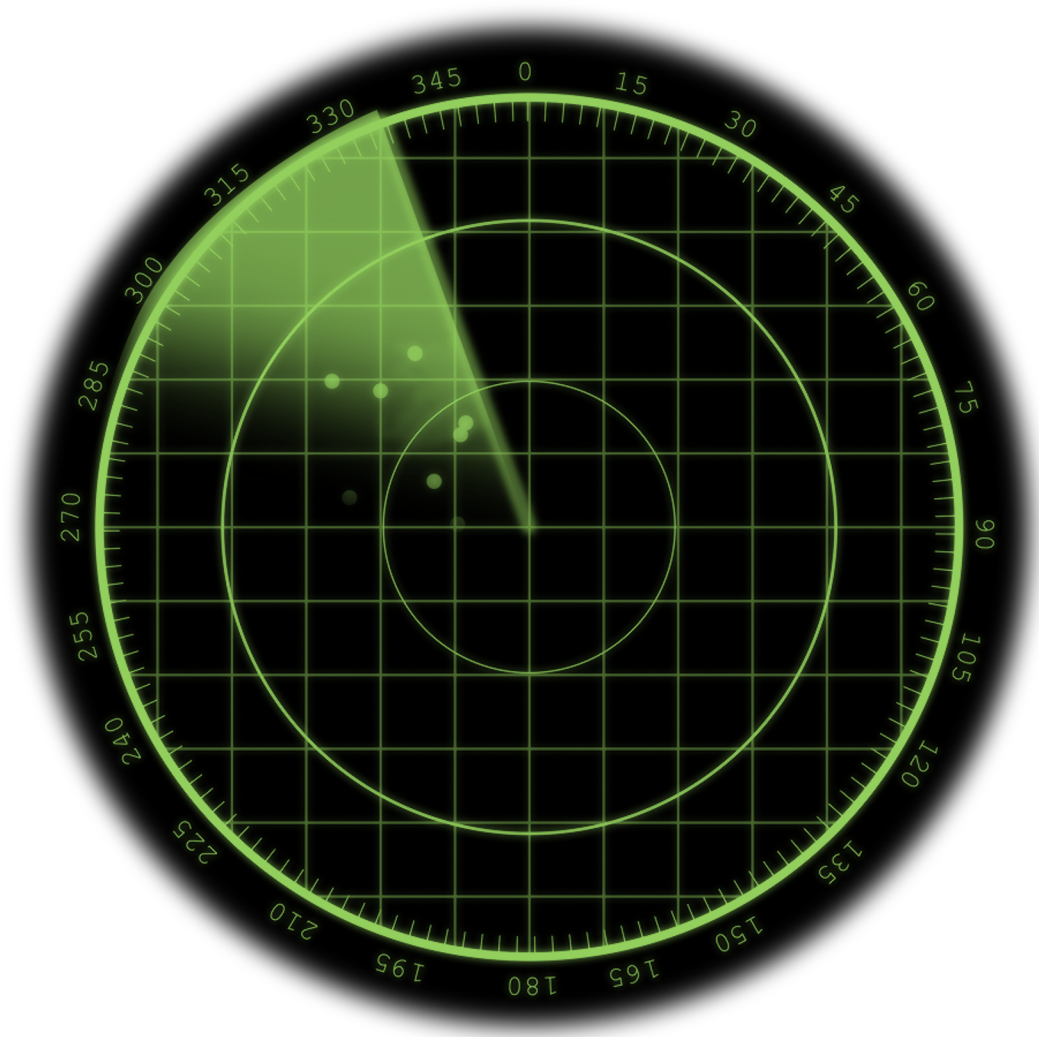




DECEMBER 03<sup>RD</sup>, 2020

# NLP in Analytics & Cognitive

META-FORUM 2020



# Deloitte's Achievements in NLP

- 1 **Text extraction in 55+ languages**
- 2 **Concept extraction in 16 languages**  
(102 planned in pilot)
- 3 **Text pre-processing in 16 languages**  
eventually 100+)
- 4 **Semantic vectors in 16 languages** (currently working with  
multilingual models, with up to 104 supported languages)
- 4 **Knowledge graphs** (Identifying subject-predicate-object  
triplets, who did what to whom, Understanding relationships  
and events)

Fully model based approach which enables to use any language to define the requirements.



# What we boast with



## Fully Automated NLP Techniques

Fully automated solution using state-of-the-art NLP techniques:

- Minimal manual input needed
- Fast implementation and **fast onboarding of a new client**
- **Great scalability** – not limited by amount of monitored companies



## Open-Source Technology

Independent from expensive external data providers, running almost entirely on open-source technology and publicly available information.



## Multiple Languages within one Framework

. Possibility to work with multiple languages within one framework:

- No machine translation until the reporting stage
- Focus even on **non-English sources**



## End-to-end solution in the house

Not only technical tool, but ability to offer subsequent services - consulting or business development.

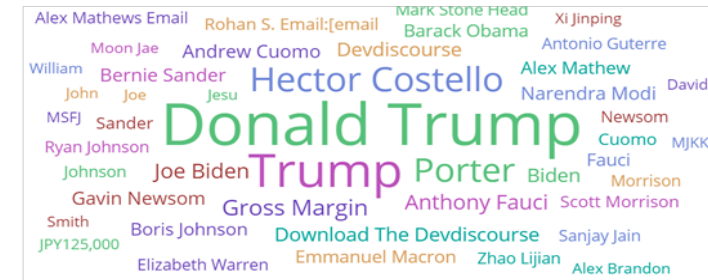
# Potential for ELG

### Tap ELG for:

- **Experienced subject matter experts**
- **Tools ready to test and implement**
- **Data corpus**

## Looking for:

- **New real world cases**
- **Potential partnerships / consortiums to respond to requests**





# Concept Linking

to real world entities

- 1 Organizations, people, places, products, and other “concepts” can be extracted and used to label and filter articles
- 2 Entities in English text can be linked to actual entities in **Wikipedia** (useful for developing a knowledge base)

“Paris is the capital of France”

  
wikipedia.org/wiki/**Paris**  
  
wikipedia.org/wiki/**France**

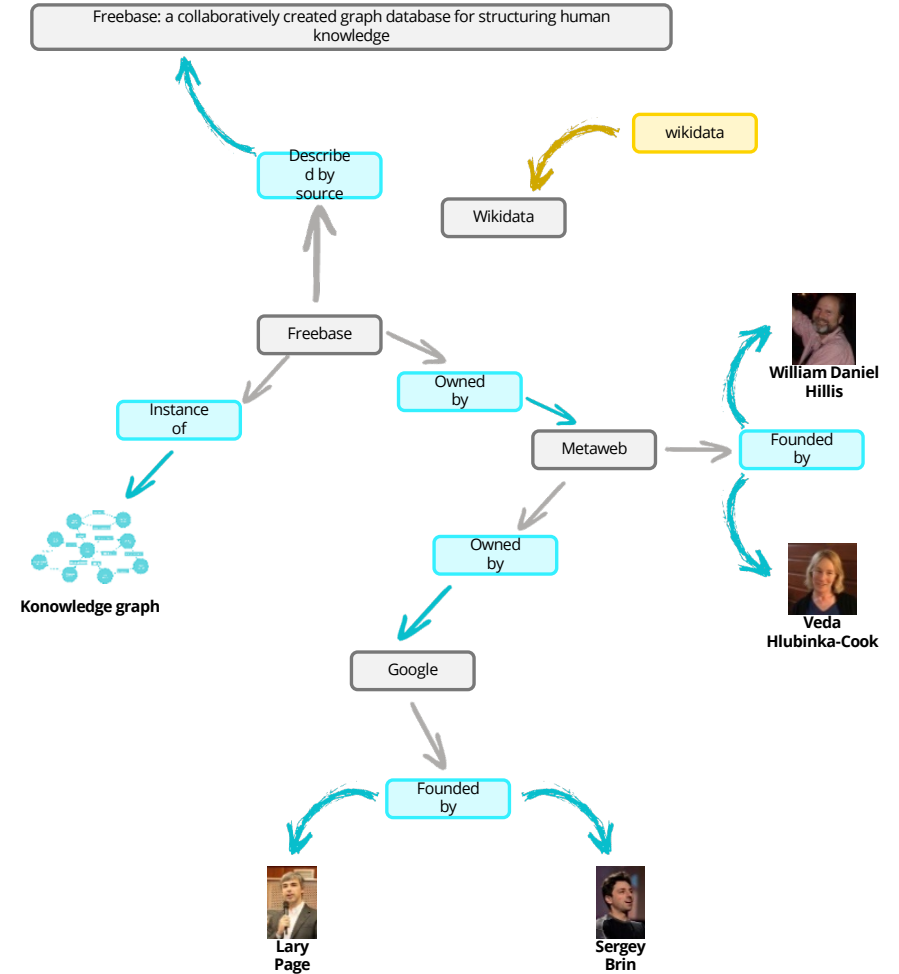


# Knowledge Graph

Extracting key events from text and representing them in a searchable graph:

- 1 **Identifying subject-predicate-object triplets**  
(who did what to whom)
- 2 **Linking concepts to entities in Wikipedia**
- 3 **Understanding relationships and events**

subject	relation	object	subject_type	object_type
launch Data	is in	cloud	NOUN_CHUNK	NOUN_CHUNK
database	hosted on	Microsoft Azure Microsoft Azure Microsoft Azure	NOUN_CHUNK	NOUN_CHUNK
Microsoft customer	know of	problem	NOUN_CHUNK	NOUN_CHUNK
Rotem	to	Discovery of unsecured database owner identified	NOUN_CHUNK	NOUN_CHUNK
database	hosted on	Microsoft Azure Microsoft Azure	NOUN_CHUNK	NOUN_CHUNK
data	hosting	data inside database hosted on Microsoft Azure in Aril	NOUN_CHUNK	NOUN_CHUNK





# Thank You.



**Julius Rusnak**

**CTO Analytics & Cognitive**

**Contact:** [jrusnak@deloitteCE.com](mailto:jrusnak@deloitteCE.com)

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. Please see [www.deloitte.com/cz/about](http://www.deloitte.com/cz/about) to learn more about our global network of member firms.

Deloitte Central Europe is a regional organization of entities organized under the umbrella of Deloitte Central Europe Holdings Limited, the member firm in Central Europe of Deloitte Touche Tohmatsu Limited. Services are provided by the subsidiaries and affiliates of Deloitte Central Europe Holdings Limited, which are separate and independent legal entities. The subsidiaries and affiliates of Deloitte Central Europe Holdings Limited are among the region's leading professional services firms, providing services through more than 6,000 people in 44 offices in 18 countries.