Artificial Insanity

Recent Trends in Artificial Intelligence and Machine Learning

Eddy Touma, Co-founder & CEO KLANGOO

Machine Learning vs. (Knowledge-based) Natural Language Understanding

NOMINAL COMPOUNDS

John works in a car factory ⇒ John works in a factory [**that produces**] *cars*
John works in a car dealership ⇒ John works in a dealership [**that sells**] *cars*

HIDDEN EVENTS IN RELATIONAL NOMINALS

John enjoyed the book ⇒ John enjoyed [**reading**] *the book*
John enjoyed the movie ⇒ John enjoyed [**watching**] the movie
John enjoyed the cigarette ⇒ John enjoyed [**smoking**] *the cigarette*

QUANTIFIER SCOPE AMBIGUITIES

John visited a house on every street in his neighborhood
⇒ John visited a [**different**] *house on every street in his neighborhood*

PREPOSITIONAL PHRASE ATTACHMENTS

I read a story about evolution in the last ten minutes
⇒ I read a story about evolution [**and finished reading it**] *in the last ten minutes*
I read a story about evolution in the last ten million years
⇒ I read a story about evolution [**that happened**] *in the last ten million years*

4 Technical Reasons why Data-Driven and Machine Learning NLU is a Myth, Walid S. Saba, April 1, 2019

Machine Learning vs. (Rule-based + ML + …)

LEXICAL SEMANTICS

John likes to play bridge ⇒ John likes to play [**the game**] *bridge*

METONYMY

The corner table wants another beer
⇒ The [**person sitting at the**] *corner table wants another beer*
The car in front of us is annoying me. Can you pass it, please?
⇒ The [**person driving the**] *car in front of us is annoying me. Can you pass it, please?*

METAPHOR

I wouldn't worry about him, John is a rock
⇒ I wouldn't worry about him, John is [**as solid as**] *a rock*

CO-PREDICATION AND METONYMY

Barcelona was calm after it voted for independence
⇒ [**The city of**] *Barcelona was calm after it* [**'s citizens**] *voted for independence*

REFERENCE RESOLUTION

The trophy did not fit in the brown suitcase because it was too small
⇒ The trophy did not fit in the brown suitcase because [**the suitcase**] *was too small* [**to hold it**]

4 Technical Reasons why Data-Driven and Machine Learning NLU is a Myth, Walid S. Saba, April 1, 2019

Machine Learning vs. (Rule-based + ML + …)

FUNCTION WORDS

a. *I know someone who writes **for** the White House*
b. *I know someone who writes **about** the White House*

a. *John visited **a** house on **every** street in his neighborhood*
b. *On **every** street in his neighborhood, John visited **a** house*

STATISTICAL INSIGNIFICANCE

*The ball did not fit in the suitcase because **it** is too*
a. ***small***
b. ***big***

Professor McDonalds told John that he should soon be done with
a. ***reading** his thesis*
b. ***writing** his thesis*

INTENSION

The president of the United States is elected every 4 years
⇒
✗ *Donald Trump is elected every 4 years*

Maybe in the 1600's Paris was not the capital of France
⇒
✗ *Maybe in the 1600's Paris was not Paris*

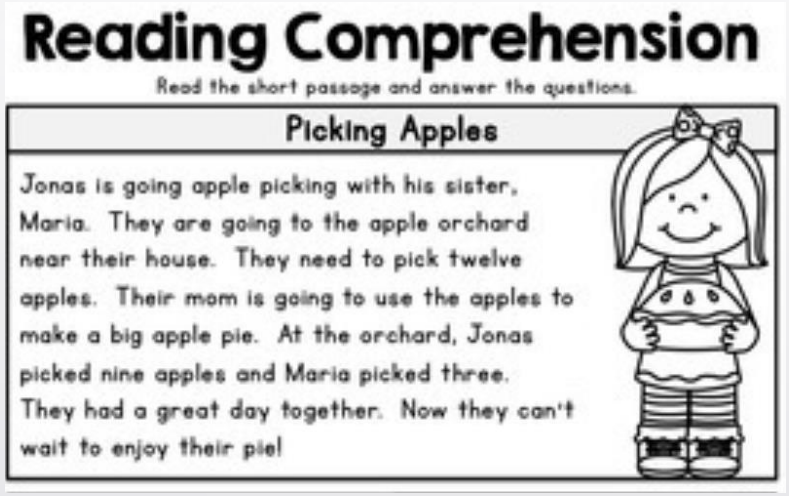
4 Technical Reasons why Data-Driven and Machine Learning NLU is a Myth, Walid S. Saba, April 1, 2019

Bottom line

Machine Learning & Statistical Models

should **ONLY** be used where it **MAKES SENSE**

Sample Correct Application



Reading Comprehension

Read the short passage and answer the questions.

Picking Apples

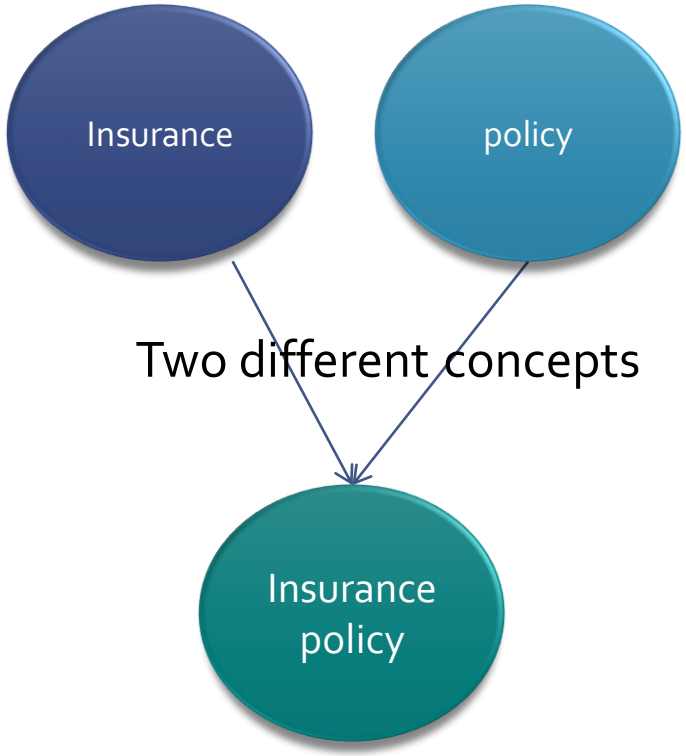
Jonas is going apple picking with his sister, Maria. They are going to the apple orchard near their house. They need to pick twelve apples. Their mom is going to use the apples to make a big apple pie. At the orchard, Jonas picked nine apples and Maria picked three. They had a great day together. Now they can't wait to enjoy their pie!

Finding the "**aboutness**" of a piece of text

I enjoy Jane's **company**.
Eddy's **company** is growing.

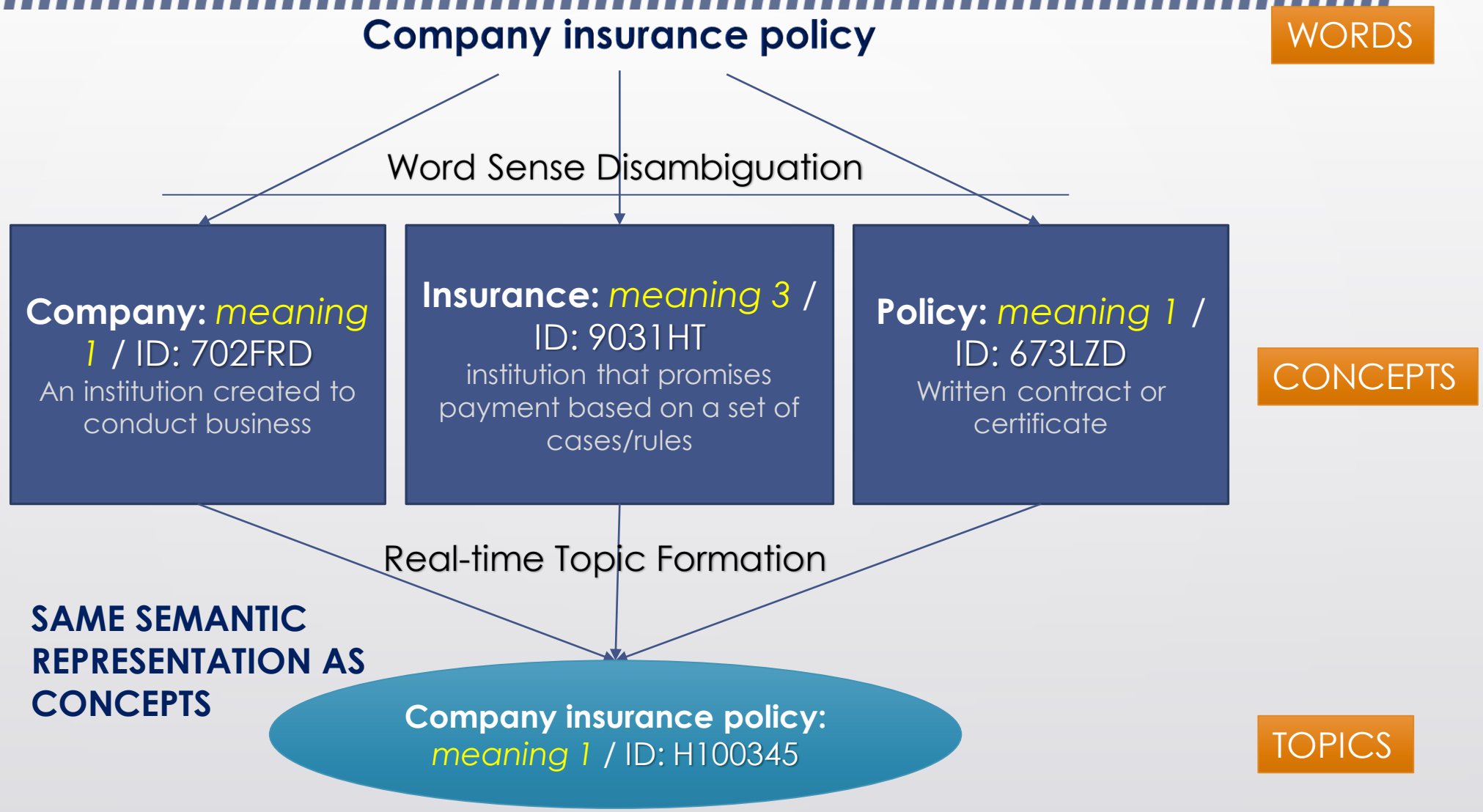
Word Sense Disambiguation
(Semantic Concepts)

Semantic Concepts



Semantic Topics

Topic = the result of 1 or more concepts


Company insurance policy
WORDS
Word Sense Disambiguation
Company: meaning 1 / ID: 702FRD
An institution created to conduct business
Insurance: meaning 3 / ID: 9031HT
institution that promises payment based on a set of cases/rules
Policy: meaning 1 / ID: 673LZD
Written contract or certificate
CONCEPTS
Real-time Topic Formation
SAME SEMANTIC REPRESENTATION AS CONCEPTS
Company insurance policy: meaning 1 / ID: H100345
TOPICS

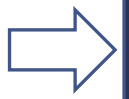
Semantic Relations:

- Parallel universe
- Parallel algorithms
- Multi-threaded computing

Topics Analysis

Semantic Fingerprint

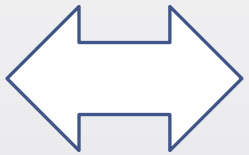
Climate change
Barack Obama
Donald Trump
administration
Washington
Paris agreement
Failed negotiations
Overturn
Pollution
Crisis management
Security clearance



ID 400389 / score 0.72
ID 29366 / score: 0.67
ID HGJ652 / score: 0.53
ID FM3LK9 / score: 0.32
ID HG8UJ0 / score: 0.31
ID JHL982 / score: 0.23
ID AYT372 / score: 0.21
ID 88OOTW / score: 0.19
ID IN73YES / score: 0.16
ID 23PO2WI / score: 0.12

Semantic Matching

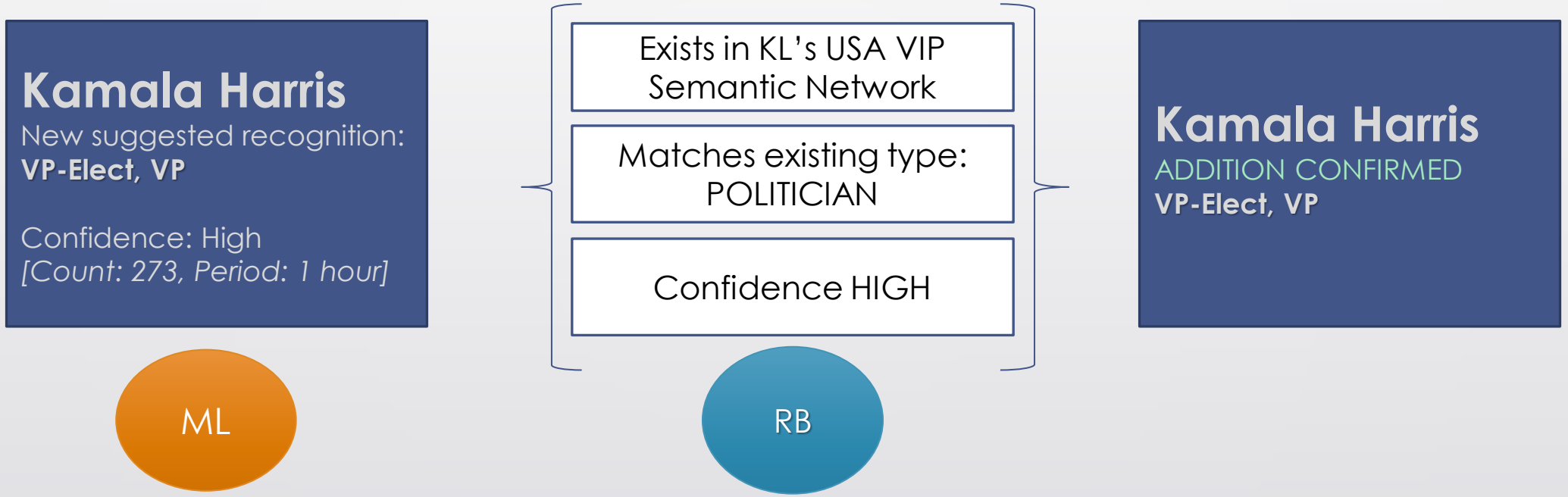
Document A	Document B
ID 400389 / score 0.72	ID 0DO878 / score 0.87
ID 29366 / score: 0.67	ID Y64232 / score: 0.81
ID HGJ652 / score: 0.53	ID KLAHYE / score: 0.67
ID FM3LK9 / score: 0.32	ID 93NLK8 / score: 0.66
ID HG8UJ0 / score: 0.31	ID USH622 / score: 0.59
ID JHL982 / score: 0.23	ID RE293H / score: 0.52
ID AYT372 / score: 0.21	ID BC7229 / score: 0.42
ID 88OOTW / score: 0.19	ID HGJ652 / score: 0.38
ID IN73YES / score: 0.16	ID 400389 / score: 0.21
ID 23PO2WI / score: 0.12	ID JHL982 / score: 0.11



Somewhat Related / 32%

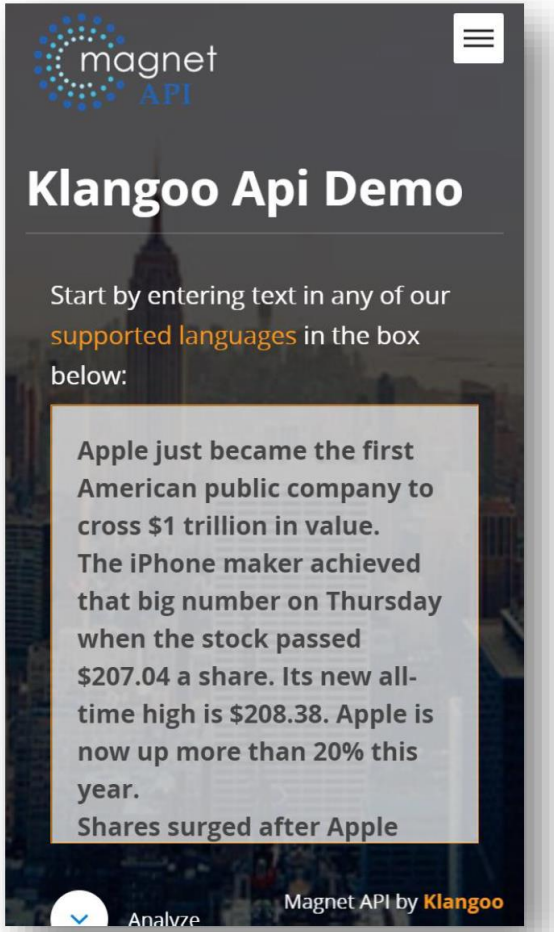
Comparing Texts in different languages

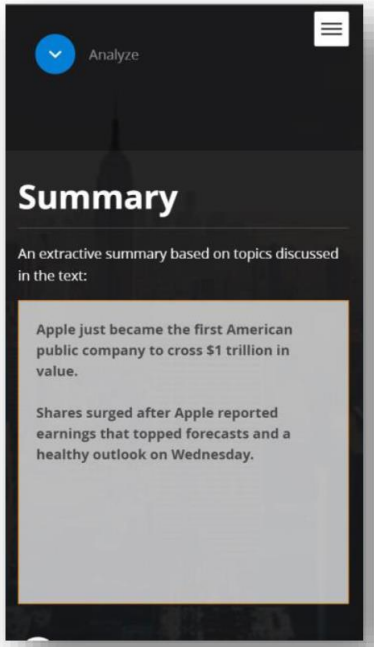
Automated Entity Addition

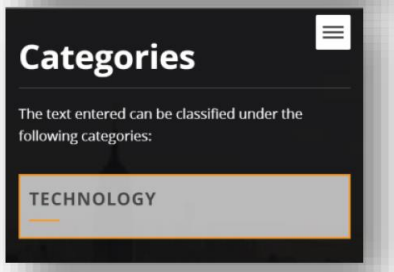


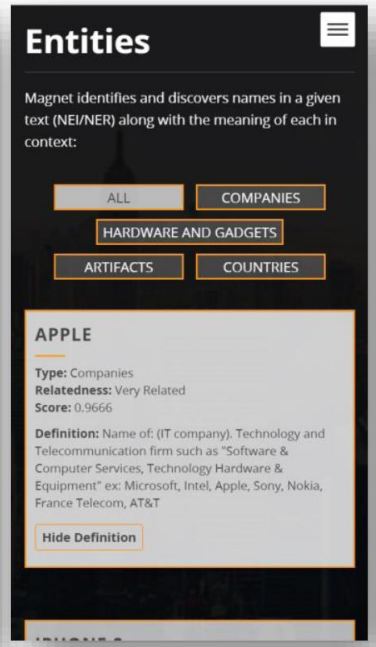
Comparison with the Leading NLP Engines

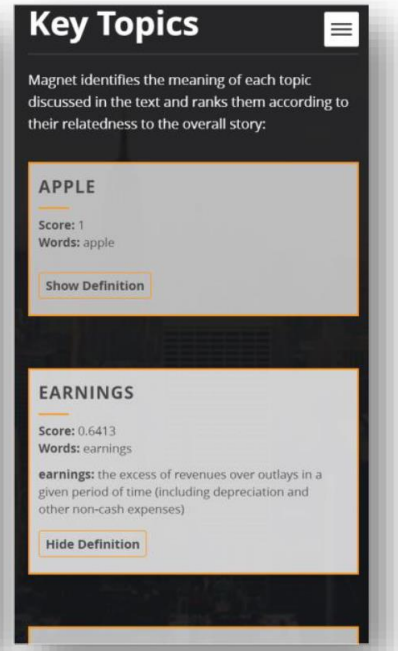
	Key Topics		Entity (NER) Extraction from Text				Categorization	Automated Summary of Text	Semantic Search
	Phrases	WSD*	Types count	WSD*	Languages	Scored			
KLANGOO. NLP	**Yes**	**Yes**	**45+**	**Yes**	**100+**	**Yes**	**100+ Languages**	**Yes**	**Yes**
Google NLP	Yes	No	6	No	10	Yes	EN only	No	No
IBM Watson (bluemix)	Yes	No	26	Yes	10	Yes	9 Languages	No	No
Microsoft Azure (LUIS)	Yes	No	NA	No	EN only	No	No	No	No
amazon AWS Comprehend	Yes	No	8	No	EN & ES	Yes	No	No	No











1. Extraction of named entities
2. Profiling and identifying behavioral patterns
3. Semantically/topically related texts and videos
4. Semantic alerts ("follow" a story or a topic)
5. Semantic automated tagging & categorization
6. Semantic Analytics (growing stories & topics)
7. Story highlights (automatic summary)

On the same subject (*by our very own Dr. Walid Saba*)**:**

- Memorizing vs. Understanding (read: Data vs. Knowledge)
- How the Tech Giants are Hampering Progress in Artificial Intelligence and Cognitive Computing
- 4 Technical Reasons why Data-Driven and Machine Learning NLU is a Myth
- Deep Misconceptions and the Myth of Data-Driven Language Understanding
- Why Ambiguity is Necessary, and why Natural Language is not Learnable
- A Knowledge Graph?
- Oh, Doesn't Google Wish!
- The New AI
- Personal Portals: Semantic, Rule-Based Personalization
- Automated, Ontology-Based Semantic Tagging

Thank you

Eddy Touma
CEO & Co-founder
eddy@klangoo.com