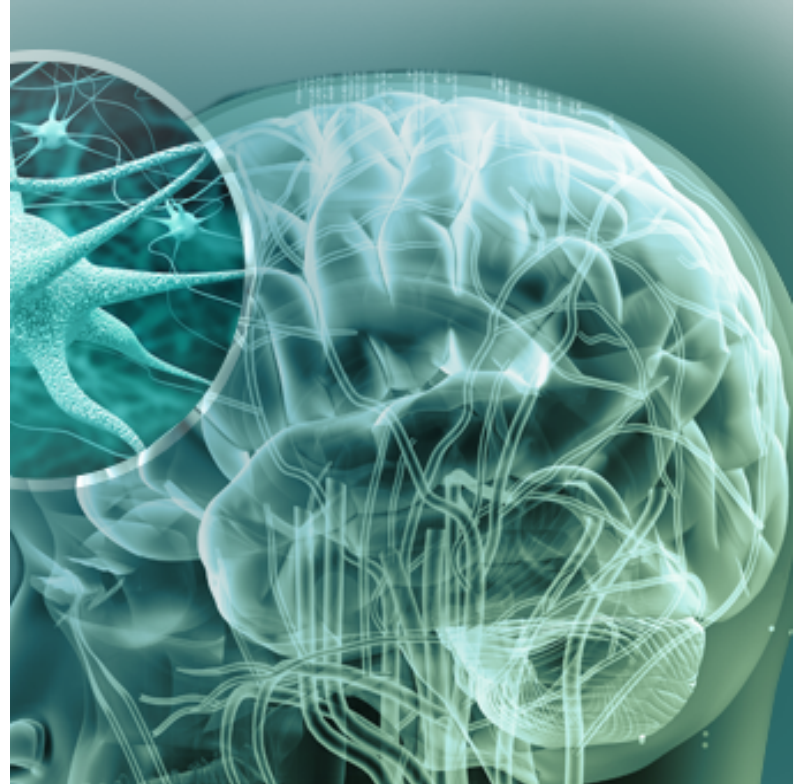


Artificial Intuition for Extraction of Meaning from Texts

- IntuView is a developer and provider of “meaning mining” in unstructured text, including: Idea mining, hermeneutic summarization, Content Categorisation, Sentiment Analysis, Entity Recognition, Analysis, Resolution, and Matching.
- IntuView currently fully supports English, Spanish, French, Arabic, Urdu (Pakistani), Farsi, Hindi, Russian, Hebrew with Turkish and Chinese in the pipeline.



www.intuview.com

- ❑ To normalize streams of information in texts and mine the “meaning” within them in a polyglot global village in which each piece of the puzzle may be in a different language or “sub-language”
- ❑ To extract information from large amounts of “Big Data” in order to discover not only the “known unknown” but the “unknown unknown”.
- ❑ To separate the “wheat from the chaff” and to extract focused information from texts.
- ❑ To identify Implicit information and links between persons, organizations, places, objects and ideas.

- ❑ “Language” is often a political and not a linguistic definition; the different languages - Swedish, Danish and Norwegian – are more inter-intelligible than Moroccan, Libyan, Mauritanian and Egyptian which are all called “Arabic”.
- ❑ Language is replete with cultural memes that are not evident to the non-initiated. These language “registers” can be rooted in social status, education, age, profession and geography.
- ❑ Even within the same language register, words, quotations, idioms or historic references can be “polysemic”; they have different meanings according to the domain and the context of the surrounding text.

“England and America are two countries divided by a common language”

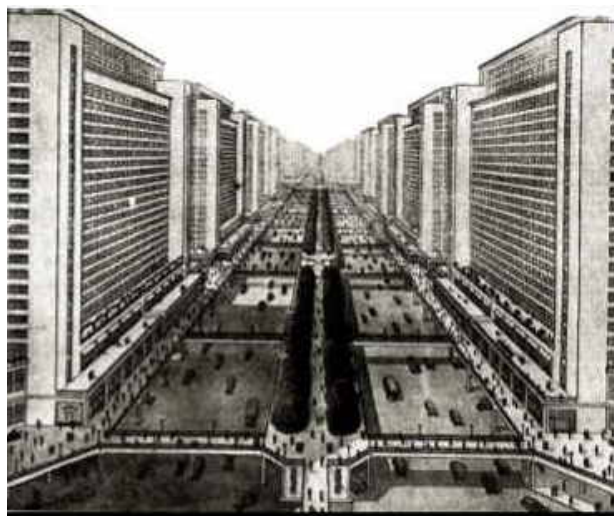
George Bernard Shaw

The Weaver Analogy



"Think, by analogy, of individuals living in a series of tall closed towers, all erected over a common foundation.

When they try to communicate with one another, they shout back and forth, each from his own closed tower. It is difficult to make the sound penetrate even the nearest towers, and communication proceeds very poorly indeed.



But, when an individual goes down his tower, he finds himself in a great open basement, common to all the towers. Here he establishes easy and useful communication with the persons who have also descended from their towers".



What is Artificial Intuition

- "Fast thinking" ("System 1") "is the brain's fast, automatic, intuitive approach... (it) is...more influential...guiding...[and]...steering than "System 2" that is the mind's slower, analytical mode, where reason dominates. (Nobel Laureate Prof. Daniel Kahneman)
- "Fast Thinking" enables us to leverage associative "extra-linguistic" information and social and cultural "memes" in order to:
 - Swiftly understand implicit meanings in previously unread texts.
 - Create impressions, feelings and "intuitive" conclusions, detect anomalies and reduce ambiguity.



The Technology

As a result of these insights, the process of disambiguation of meaning in texts is based on a number of stages:

- Identification of the “register” of the language
- Statistical categorization of a document as belonging to a certain domain, topic, or cultural or religious context in order to reduce ambiguity.
- Using the lexical tokens directly preceding or following the token or tokens in question – in order to provide additional disambiguating information.
- Matching the lexical tokens after disambiguation to unequivocal instances in an ontology, adding the ontological relationships and features to those of the lexical token.

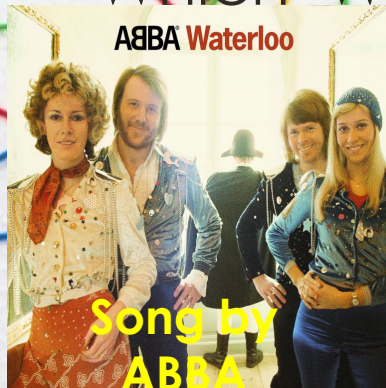


The Problem of Academic Research

- Manual classification of books, articles and other textual materials for academic research does not cover the entire scope of the content: persons, places, organizations and events mentioned; relations between them.
- String searches return polysemes and irrelevant uses of the words and do not return abnormal or antiquated spelling and do not return identical concepts that are worded differently.
- “Keywords” are subjective and are interpreted differently by different people – annotators and users.
- Documents contain a multitude of references to diverse subjects and entities that may not have any “real” relationship between them except for casual co-reference. Searches do not enable complex search for “real” relations between search terms but merely co-reference within the text. Many “hits” may refer to X and Y but without any connection between them.
- Current search engines are not language independent. In a global world and multi-lingual environment, search must be multi-lingual.

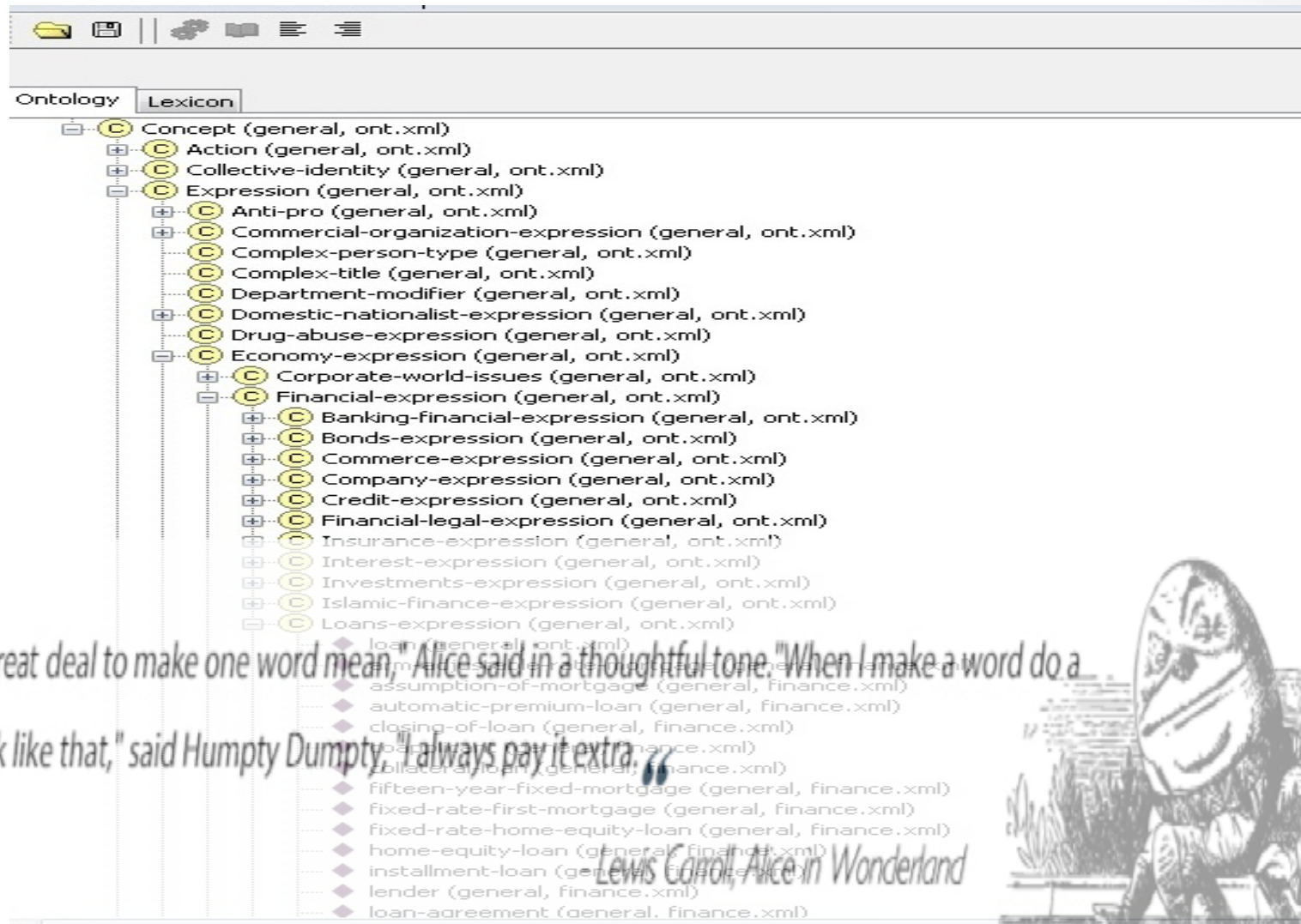


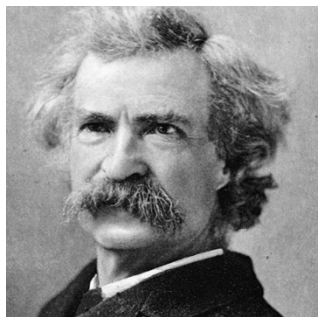
Which WATERLOO do you mean ?



Final Defeat

An ontology is a hierarchal map of concepts and specific instances and their attributes that clearly defines each instance and its relationship with other instances. Many expressions in different languages can all “mean” the same ontological instance, which, in turn, may be linked to different concept parents under different conditions.





“They spell it Vinci and pronounce it Vinchy; foreigners always spell better than they pronounce” - Mark Twain

- ❑ Names reflect cultural naming conventions, are received from different sources and are corrupted by different linguistic backgrounds of the transmitters and the transcribers of the names.
- ❑ A name can identify the person in many ways: Gender; Ethnicity; Religious Sect; Region/City ; Tribal affiliation; Family relations; Social status; Age
- ❑ Al-Zawahiri = Eddhawahri = Eddaouahiri = الظواهري
- ❑ Xi Jin Ping = Hsi Chin P`ing = 习近平

IntuScan identifies the linguistic origin of the name, reverses it to its source orthography, applies cultural naming conventions and statistical models to generate name variants and discovers information such as ethnic origin, gender, religion/sect, status, family/tribal links etc.

Thank You

