Seminar in:

IBM Watson

Watson translator

Presented by:

Ahmed Albawabiji

What is Watson Translator?

- Source / target language translator.

- Texts translator.

- Offers multiple domain-specific models based on your terminology and language.(conversational, news and patent)

- Takes specific terms and phrases into account.

- Can be teached (forced glossary) e.g. (Burger King).

- Parallel corpus can be provided.

- Can improve itself, when we upload a text in a target language (monolingual courpus).

Language Translation

IBM

Technology

- Uses tokenization to understand a word in a context.
- Allows users with three levels to train a custom translation model:

-Forced glossay: Adds and completely overwrites the orginial tranlation model.

-Parallel phrase customization: Adds a series of matching phrases in the source and target language. Overwrites partly the original translation model.

-Corpus level customization: Allows users to upload a plain text file that contains body of text in the target language that is related to what you are translating. Monolingual corpus helps make your translation more fluent and human.

Training, limitations and state

- It is trainable indeed, but there's constrains:

    - Forced glossary: Limited to a file size of 10 MB. Can upload only one forced glossary file per translation model.

    - Parallel_corpus: Must contain a minimum of 5,000 term and translation pairs.

    - Monolingual_corpus: Must contain a minimum of 1,000 sentences.

- The cumulative file size of all uploaded glossary and corpus files is limited to 250MB.

- Still under testing (alot of laungues are not available yet)

Users & products?

- Could be used to develop smart translation devices.

- Could be itegrated with another APIs like speech to text to get smarter devices.

- Translating communications in real time like emails, online chats and other multilingual communications.

- Could be very useful for companies, which purchase and sell abroad.

- A powerful cell phone App Translator.