

[ubiQus]



# Deploying neural networks in operations

SlatorCon Zurich November 29th 2018

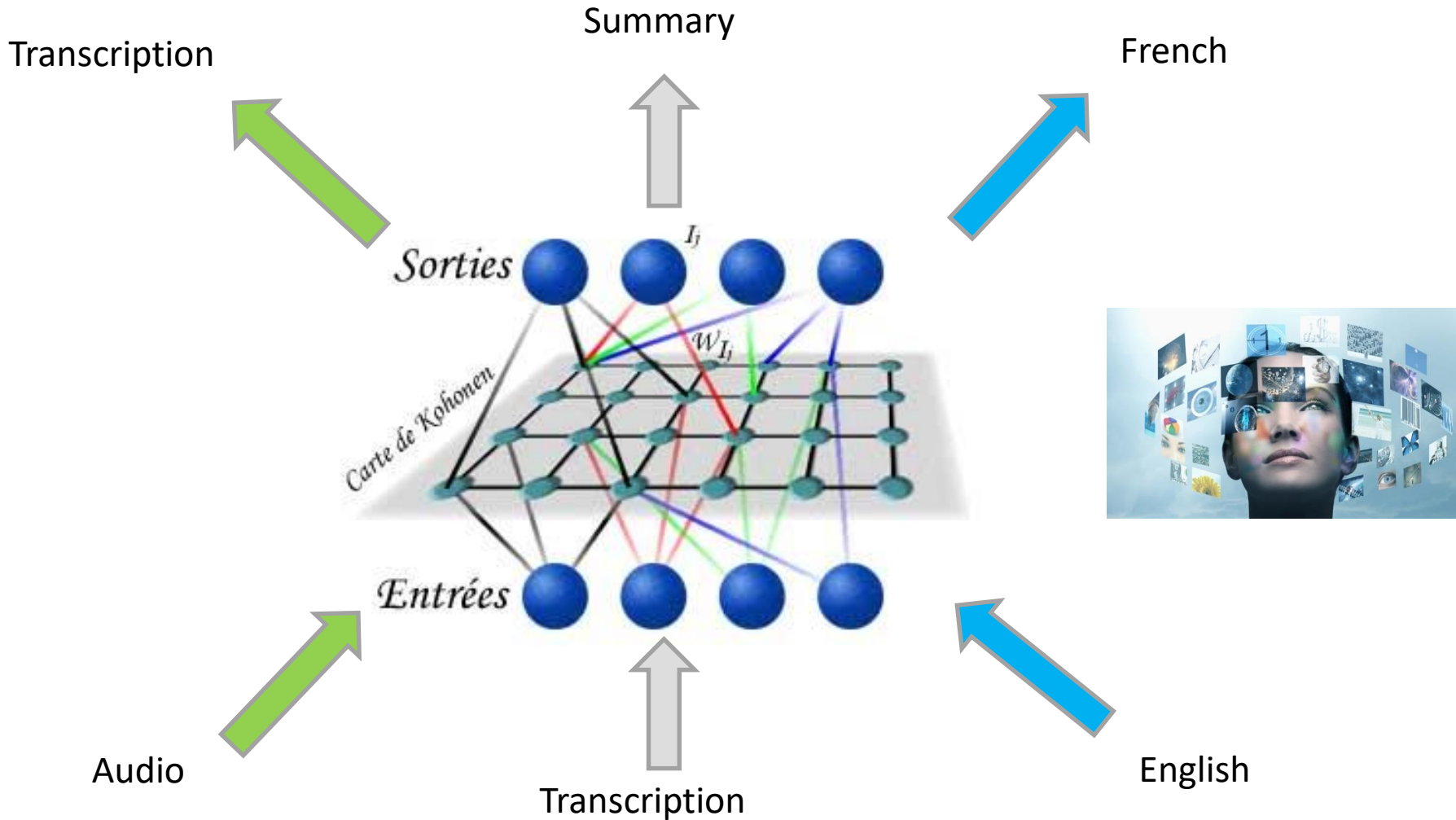


## Who we are

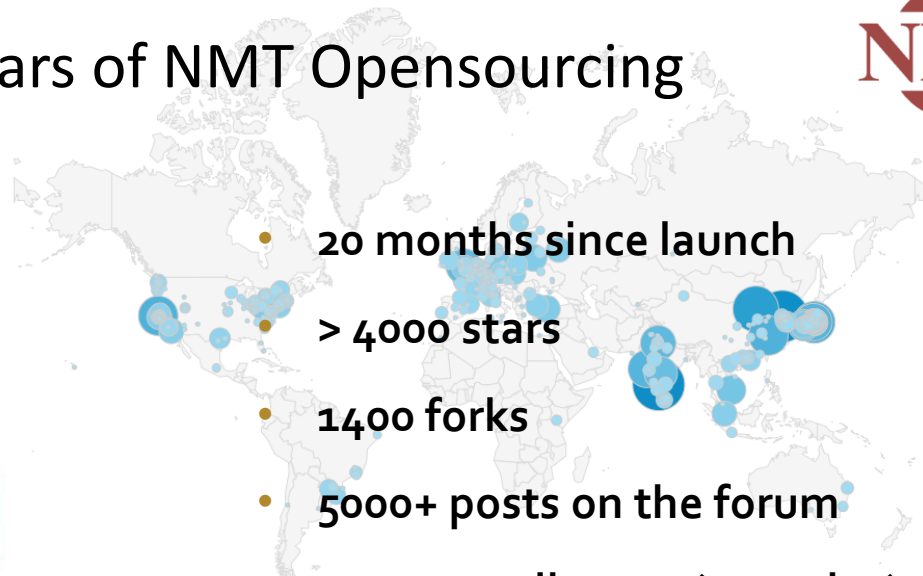
- Founded in 1991 in Paris, now established in France, UK, Ireland, Belgium, Spain, USA, Canada
- Revenue : 75 Million Euros
  - Translation & Interpretation
  - Summarization
  - Transcription
- Very active R&D team in NLP



- Top 20 LSP
- Very active in M&A over the past 15 years

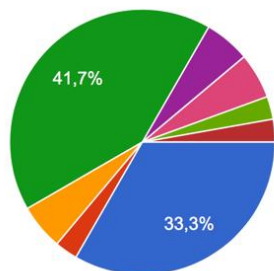


## OpenNMT: ... 2 years of NMT Opensourcing



- 20 months since launch
- > 4000 stars
- 1400 forks
- 5000+ posts on the forum
- 100+ contributors (15 active)
- Many code refactoring

Only ... 5000 lines of code



- Developer
- Linguist, Language specialist
- Translator, or Translation Project Manager
- Researcher / Academics
- Independent Expert
- Executive
- Hobbyist
- First two and last
- Software architect (researcher + de...)

## Key Challenges

- *Build the best systems*
- Convince internally
- Convince translators
- Educate your clients



# [ubiquus] Deployment in Operations

## *Building the best systems*

- Difficult technically speaking but easy when you know what you are doing: clean data, hardware, ...
- Integrate into the existing workflow not to disrupt processes: proprietary developments
- Show reports: educate, explain

## Never forget

- Everyone will point out the errors x10 more than the actual benefits .....





# System Integration

NMT in a workflow

In-house  
Engines

En-Fr

Fr-En

En-Nl

En-Es

opensource

GPU enabled server

proprietary

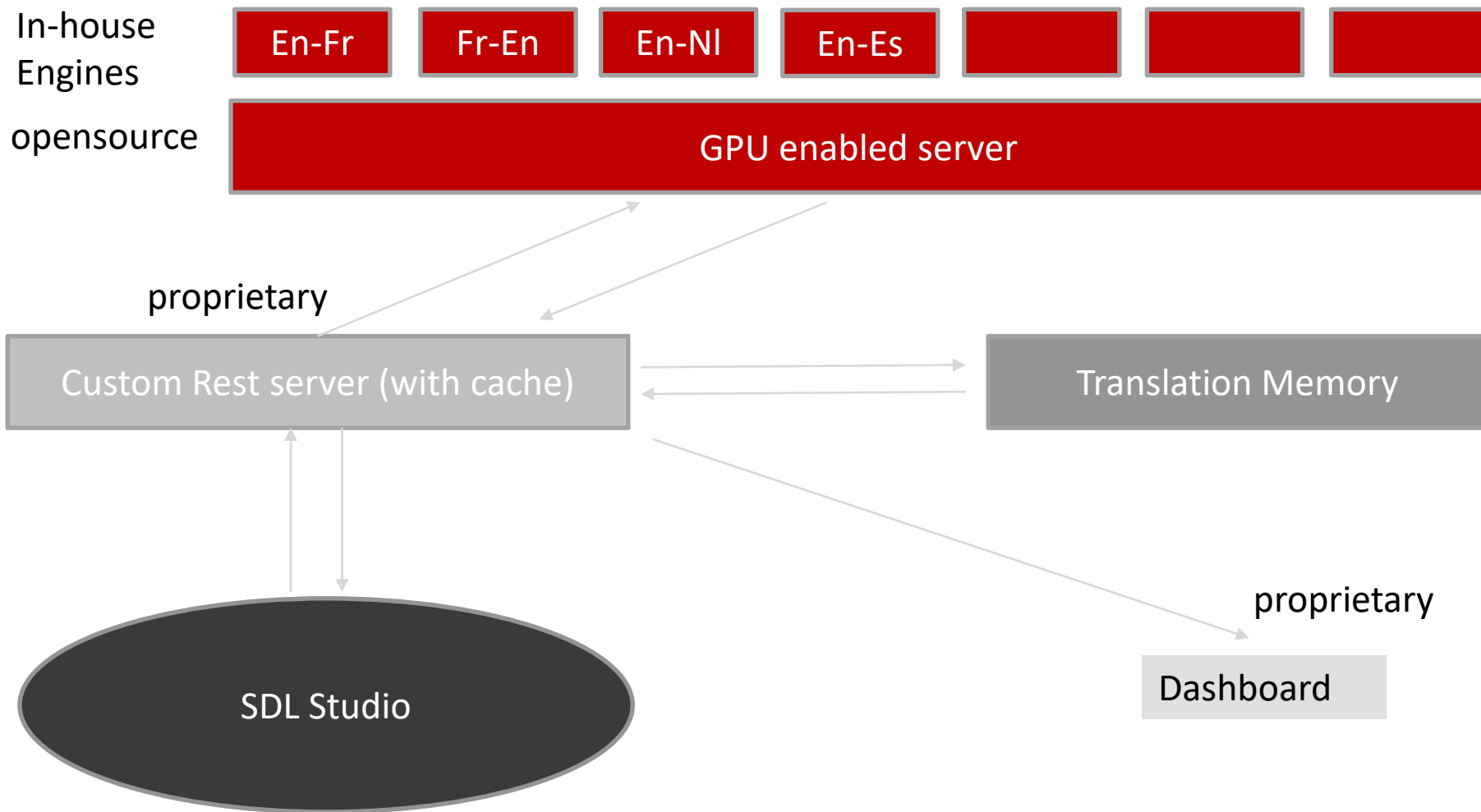
Custom Rest server (with cache)

Translation Memory

SDL Studio

proprietary

Dashboard



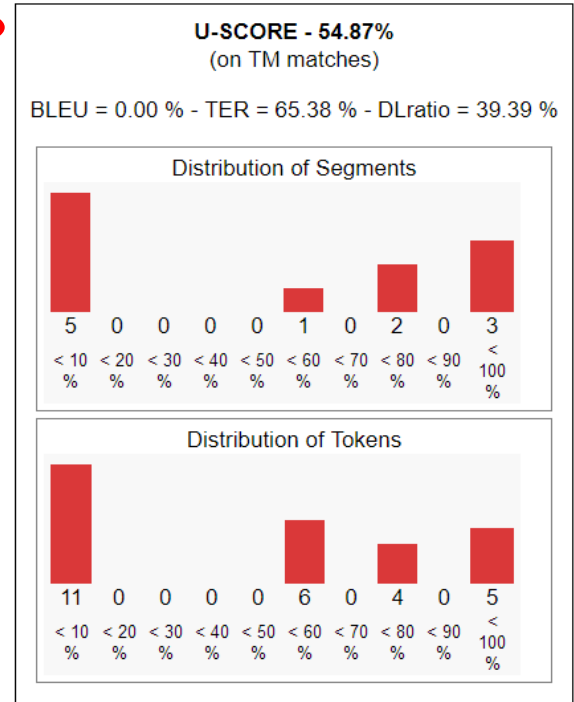
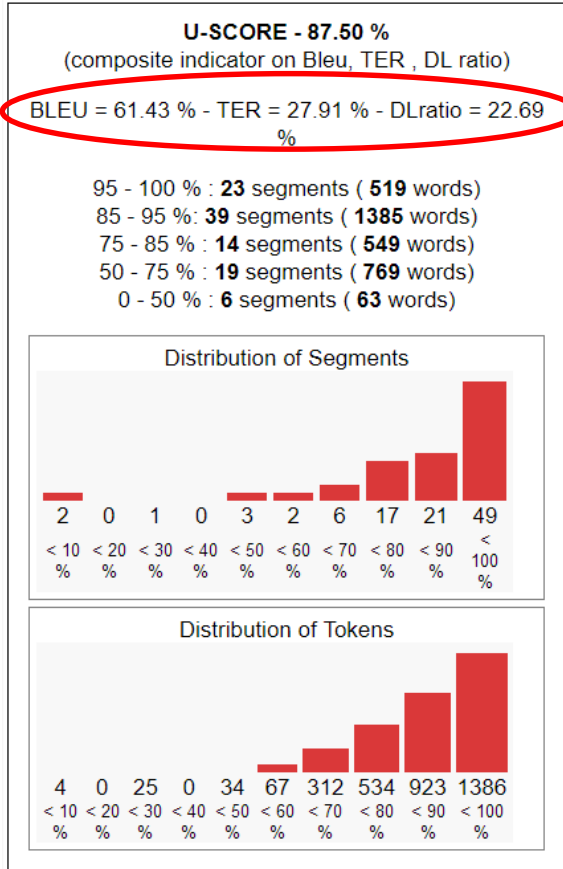


# Deployment in Operations System performance

- Each job sent by the human translator is scored against our NMT engine
- Gives a better sense on how accurate NMT is for each client and each type of job
- Data are then re-used for training

NMT (es --> en) engine was used  
 3043 words / 3398 tokens in source  
 2869 words / 3285 tokens in finished translation  
 2822 words / 3211 tokens in NMT translation

11 segments were TM matches (or AppliedText ref)





*Convincing internal and external linguists....*

- Can be a very long journey: take the time
- NMT output is the best advocate
- Admit when NMT is not good enough, especially with creative content

Period:

Engine (5) ▾

general

- General (de → en)
- General (de → fr)
- General (de → nl)
- General (en → de)
- General (en → es)
- General (en → fr)
- General (en → fr\_ca)
- General (es → en)

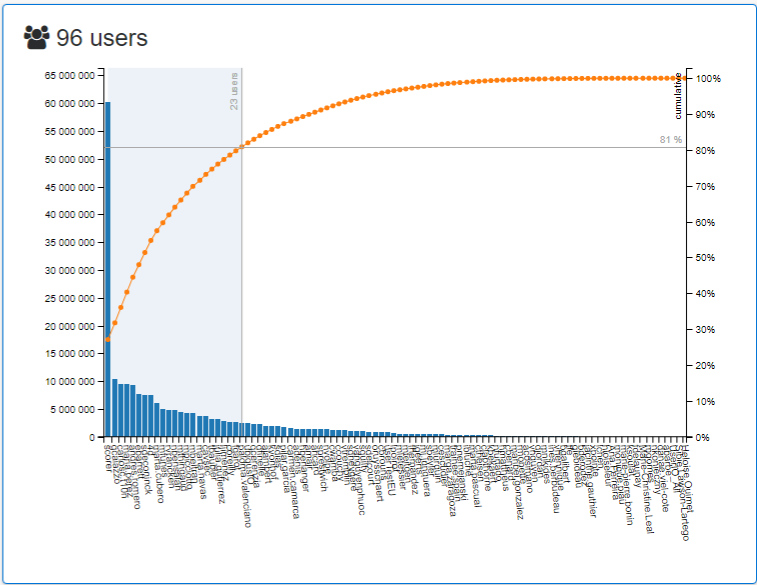
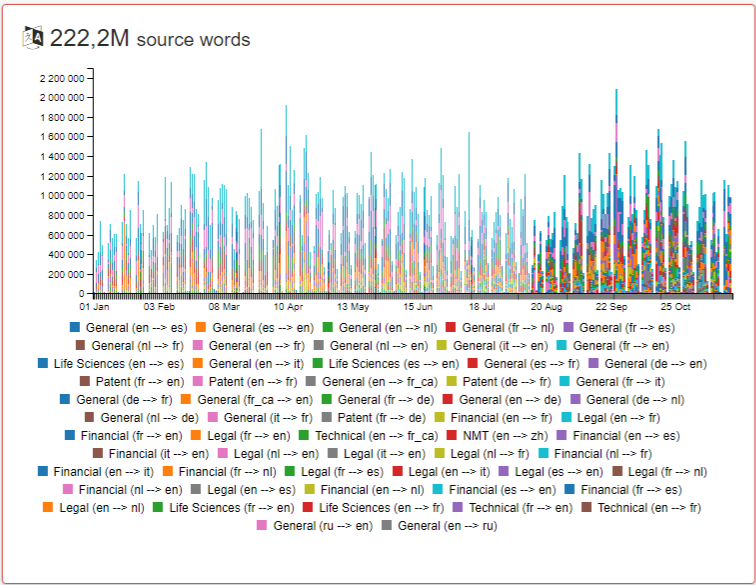
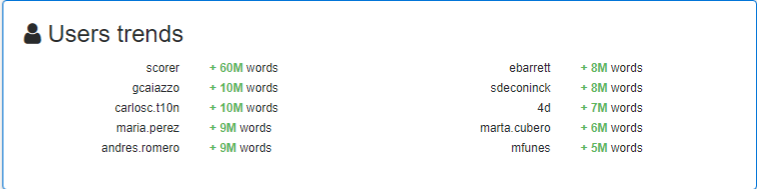
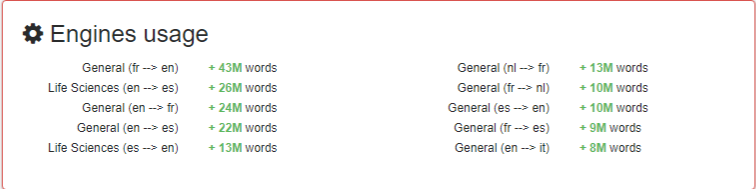
User (105) ▾

admin

- Heloise.Ouimet@Textualis.com
- abarbe@ubiquus.com
- acosimano@ubiquus.com
- claphorne@ubiquus.com
- djordan@ubiquus.com
- fdelaunay@ubiquus.com
- fhermandez@ubiquus.com
- gduvartier@ubiquus.com

Base

- Source word
- Request count
- Response length
- Request length



### *Educate the your clients ...*

- Knowledge of buyers varies a lot:
  - “I don’t want you to use MT”
  - “What would be the cost if you used MT?”
- NMT is just another CAT tool in our workflow.
  - We may sell “FIPO jobs” but not the main goal
  - We do not compete with Systran, Iconic, KantanMT (yet)
- Service service service.

### *Conclusion*

- Deploying NMT / ASR at Ubiquus took less than a year:
  - Half-way in some subsidiaries
  - Fully integrated in some others
- Acceptance highly depends on ownership by management team
- R&D team has to be involved in production workflow, not only in SOTA results.

# [ubiQus]



**Deploying neural networks in operations**

**Thank you !**

**SlatorCon Zurich November 29th 2018**