

# The elusive encounter of a survey methodologist, a platform engineer and a translation technologist



ESRA – Zagreb, 17 July 2019

**cApStAn.**  
linguistic quality control

Lynn

survey  
methodologist

questionnaire  
author

social science  
researcher



Irmtraud

platform engineer

lead programmer

IT scientist





# Blaise

translation  
technologist

terminologist

active member of  
OmegaT developers'  
community



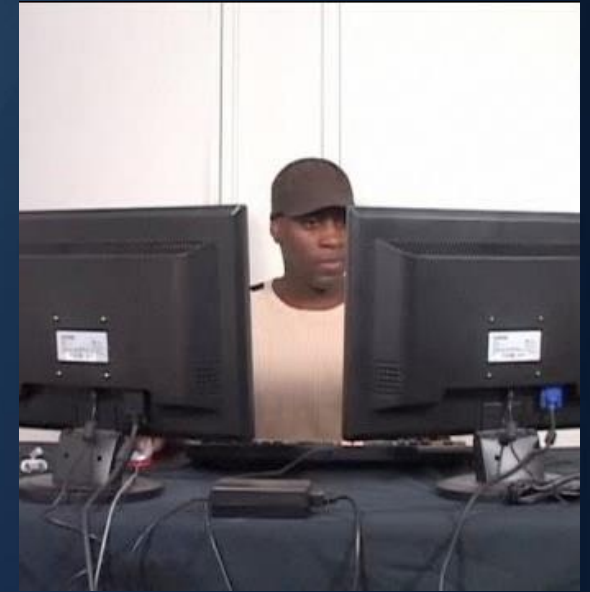
Lynn



Irmtraud



Blaise



# ⊖ wave of 3MC Survey

Lynn



- Mixed-mode survey
- Team translation with team adjudication meetings

Irmtraud



- Everything in one place, from authoring to case mgt
- Built-in translation editor

Blaise



- Standard-compliant XLIFFs
- Alignment of existing translations



# A conversation between two different worlds

Could you tweak the translation editor to accommodate team translation, Irmtraud?

Anything that can be explained in the form of specifications for programmers can be done, Lynn.

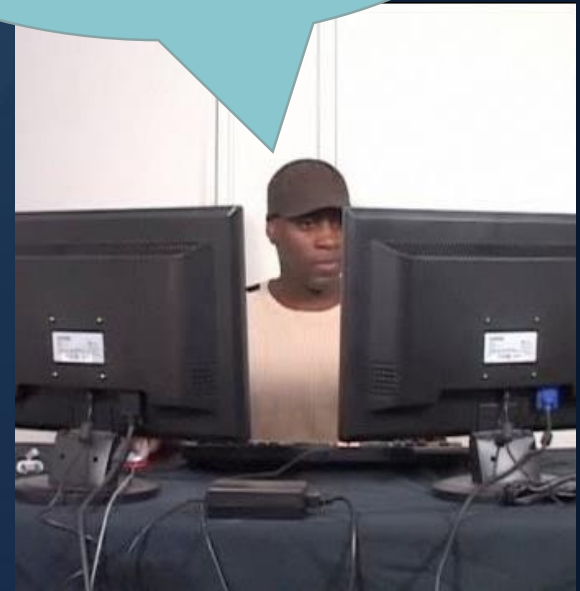
Do we have a budget?



# Leveraging Assets

Could you make sure that translators use existing translations of trend questions, Blaise?

Sure, Lynn. All I need is bilingual files from the previous wave. If you don't have them, give me the master QQ and the target versions, I'll align and produce TMs



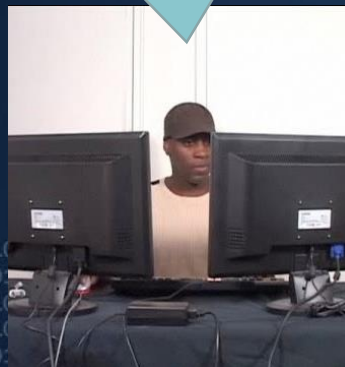


# Exporting and Importing back

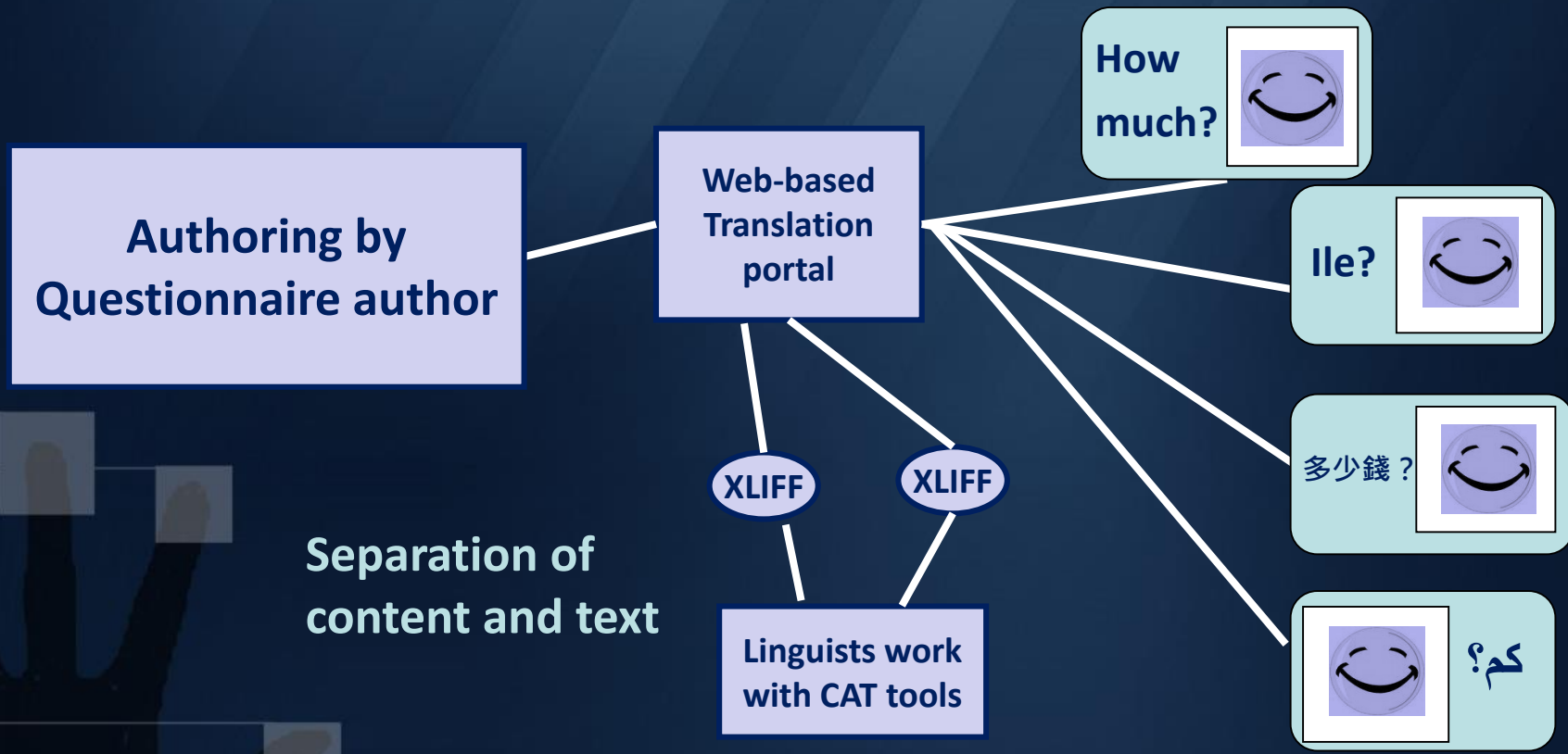
Could advance translations for the cognitive pre-test be produced outside the platform?

Outside the platform?  
Yes, Lynn, but if you want to use these translations later you'll have to copy-paste them into platform.

How about extracting the translatable text from the platform, in XLIFF format?



# translation process outside the platform



Separation of content and text



# Good news for Irmtraud and for all platform engineers

- open-source Java libraries exist
  - Okapi XLIFF Toolkit for XLIFF 2.0
  - Okapi Rainbow for XLIFF 1.2
- allow to create, read, write and manipulate XLIFF, which should facilitate adoption of the standard for any organization



# Segmentation

In order to usefully leverage TMs, I would need sentence-based segmentation.



Do you have standards or specifications for this?





# Good news for Irmtraud and for all platform engineers

- Segmentation Rules Exchange (SRX)
- SRX file defines segmentation rules and can be shared among different parties
- so that segmentation can be applied consistently in the creation of different translation packages by different people and at different points in time.

# Conclusive remarks (1)

- **Non-professional** translators (and SMEs who are not linguists) tend to favour low-tech solution – e.g. all in one platform
- Reluctance to work with tools can be overcome with adequate support
- **Professional** linguists work (much) better if they can use their tools
- Professionalization improves consistency and overall quality



# Conclusive remarks (2)

- XLIFF and other data exchange standards have the *potential* of streamlining the translation process
- If not used properly, standards can become a prison in which translators are tortured in unpredictable ways
- know-how of a translation technologist is needed to handle the level of detail that the data standard does not cover

# Conclusive remarks (3)

- creating a complex translation workflow for survey instruments benefits from a multidisciplinary approach
- Ideally, survey designers/questionnaire authors, platform engineers, and translation technologists
- should work together from the earliest planning stages

Thank you

[www.capstan.be](http://www.capstan.be)

[steve.dept@capstan.be](mailto:steve.dept@capstan.be)