

July 21 Expert Group Meeting: Main Outputs

The Challenge in brief

Challenge of the Challenge:

- Ultimate goal: Automatically translated scientific paper published
- In practice: i) Lower barriers for publication in English for scientists and ii) promote access to science for citizens and sometimes for scientists (e.g. science in Chinese)

Strategy:

- Primarily: Tackle scientific challenges of Machine Translation related to scientific documents (incl. tables, images or LaTeX)
- Secondly: Tackle evaluation challenges:
 - ✓ Go from automatic evaluation of isolated sentences translation to document-level translation automatic evaluation
 - ✓ Availability of data

Impact:

- Machine Translation technologies and Natural Language Processing for scientific documents
- Automatic understanding systems for languages other than English
- Document-level metrics

Challenges of the Challenge organisation and questions

Call:

- One track for selecting the R&D projects
- How to generate/select the evaluation organiser consortium?

Tasks:

- Add tasks more complicated with the years (e.g. slides, presentations)
- Is it recommended to include in the challenge spoken language translation?

Disciplines:

- Select discipline with most relevance (connect with citizens, make scientific evidence available to the citizens)
- We should check resources available in a given discipline
- We could consider social sciences and humanities (need expressed in this domain), health, climate, digital sciences

Metrics:

- Start Challenge with existing metrics. Maintain throughout the challenge to provide comparisons
- Integrate in the Challenge objectives: Development of metrics for automatic evaluation of document-level translation

- ✓ Progress is expected from joint work between funded projects and the evaluation organiser consortium. Lead by the evaluation organiser consortium. Need to show that new metrics have correlation with human evaluation
- ✓ The evaluation organiser consortium is expected to evolve into a multilingual framework, so that they can prepare the resources and assess them
- Possibility to evaluate document-level translation via crowd source human judgment or measure of post-editing effort
- François, Rachel and Jan have relevant experience
 - ✓ Form specific working group?

Connect CHIST-ERA Challenge

- Engage with editors and communities that need the technology to check best evaluation structure and check what resources they can make available
- National conferences are good targets: They accept non English papers and some of them translated papers / Rely on evaluation structure of conferences and journals (for instance Spanish Society for Natural Language
 - ✓ Collaboration should be established by the call
 - ✓ Should be done in an open way (in particular, generated data can become training data)
- Computing resources?
- Link with WMT?
- Present Challenge outcomes in other multilingualism initiatives (e.g. <https://european-language-equality.eu/>)

Ranking:

- Based on clusters?

Barriers

- How to attract scientists to play the game (submit papers)?