# Document-level NMT

Josef Jon, António Lopes, Rachel Bawden, João Godinho
Lucy Park, Christian Hardmeier, Tereza Katreniakova

# Current status

| | STATUS | EN-FR | EN-DE |
|---|---|---|---|
| **Baseline** | Converged | Europarl | Europarl |
| **Concat 2to1** | Converged | Europarl | Europarl |
| **Concat 2to2** | Converged | Europarl | Europarl |
| **Multi-source (Voita)** | Implemented (Debugging) | - | - |
| **Multi-source (Zhang)** | Training | - | Europarl |
| **2-step Document Decoder** | Implementing | - | - |
| **Cache** | Not started | - | - |
| **Hierarchical Attention** | Implementing | - | - |

# First Results

| | Bawden's dataset Lexical choice (EN-FR) | Bawden's dataset Anaphora (EN-FR) | Mueller's dataset EN-DE) |
|---|---|---|---|
| **Baseline** | 50% | Running | Running |
| **Concat2to1** | 49.5% | Running | Running |
| **Concat2to2** | 50% | Running | Running |

# Next steps

- Finish implementing methods
- Run experiments with different datasets
- Evaluate all methods with the automatic metrics
- Possibly evaluate on EN-CS (Josef dataset)