



# Digital Transformation

**26/1/2018**

# Data analytics: programming

- *Objective: proposing a programme to B2 by taking into account available interpreters and programming constraints*
- *1st test: programme was generated in 30s*
- *Next steps: taking into account all constraints and ACIs*
- *IT Project: integrating DORIS and Pearl*

# Data analytics: terminology extraction

- *Objective: using DORIS to extract terms from documentation and generate glossaries*
- *Use cases: building glossaries from new major proposals (MFF,...); filling gap in « smaller » languages (eg Presidencies)*
- *Next steps: use case to be refined and tested with DORIS*



European  
Commission

Glossaries, docs, search (web)

Glossaries, docs

Lithos

MEDATA

Interpreters

Client

C3/  
Documentation

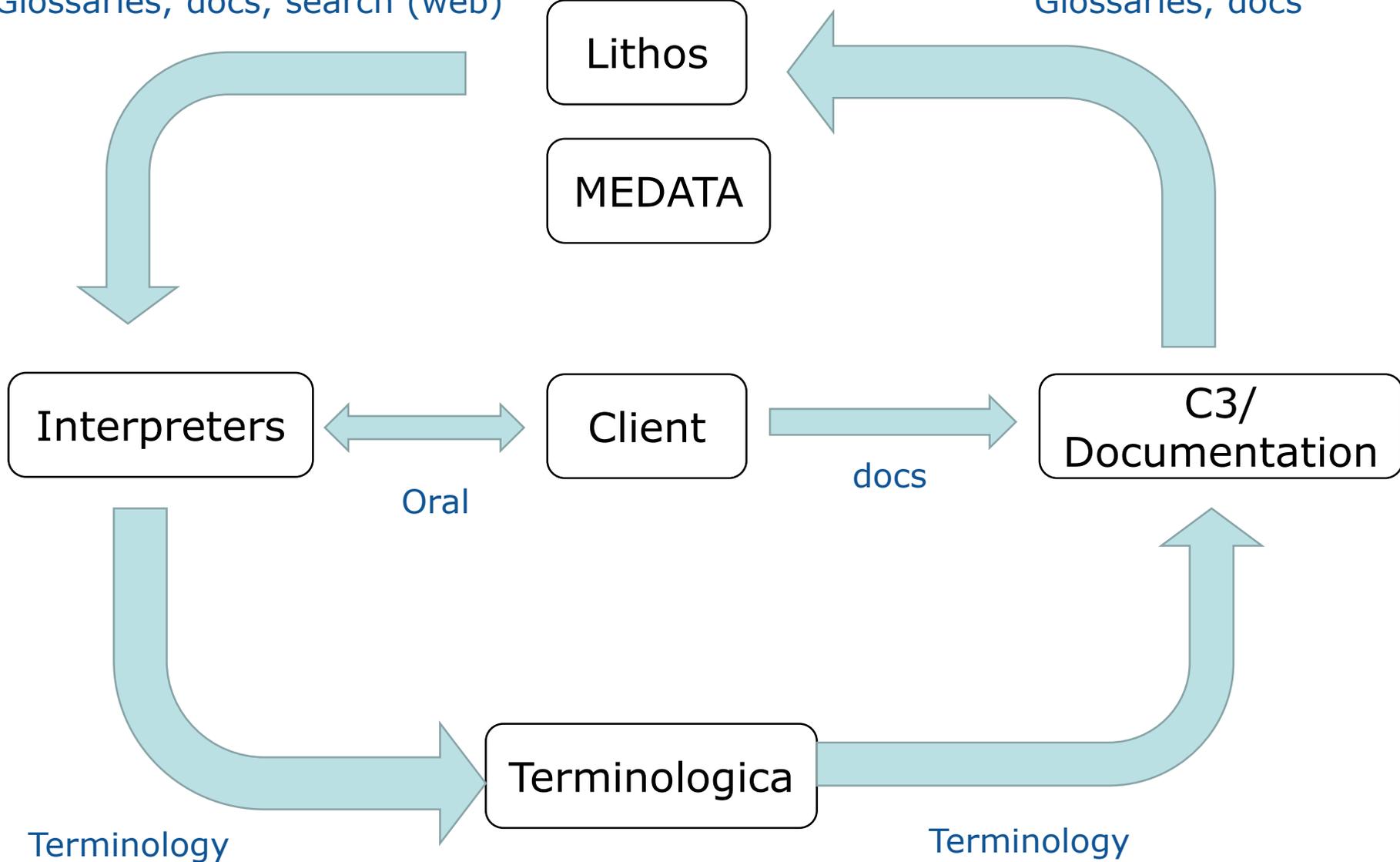
Oral

docs

Terminologica

Terminology

Terminology



# Building the 2<sup>nd</sup> generation of terminology/documentation tools

- *Today's situation: significant terminology work, large number of glossaries, large amount of meeting documents...*
- *...glossaries are managed with excel sheets, Lithos is not integrated in SCIC information system, no systematic feed-back between Meeting Blog and glossaries*
- *...no automatic generation of event-specific glossaries based on interpreter's profile, no way to integrate personal glossaries, no SR/STT*

# Building the 2<sup>nd</sup> generation of terminology/documentation tools

- *What could be a « 2<sup>nd</sup> generation tool »?*
- *Revamped interpreter's interface allowing to build event-specific glossaries and retrieving terms from various sources (IATE, SCIC, personal glossaries) – using InterpretBank?*
- *Lithos is discontinued: glossaries are stored in our database (no excel sheets!)*
- *Support to future speech recognition and speech-to-text solution*

# Research topics

- *Speech recognition*
- *Speech-to-text*
- *Automated terminology extraction based on documents and context (domain)*
- *High-Performance Computing (HPC) & Linguistics*
- *Text-to-speech*

# Research topics

- *Knowledge Center: call for interest in testing speech recognition solutions?*
- *DIGIT Data and AI strategy: securing visibility for our issues/topics*
- *H2020/Connecting Europe Facility in Telecom: securing funding for research topics (alongside translation)*