

Physical textual documents

Transcribed and annotated files
(.txt, XML, spreadsheets, databases etc.)

Network structure

**Textual Sources
Acquisition**

Digitization

Annotation

**Network
Transformation**

Network Analysis

I1: Too many missing documents (data not exhaustive)

I2: Original sources and numerizaion mismatch (Wrong character recognition, transcription error etc)

I3: Missing annotation
I4: Non systematic annotation schema
I5: Errors in Named Entity Recognition and Disambiguation

I6: Bad network model

I7: Wrong representation
I8: Wrong choice of algorithm/measure