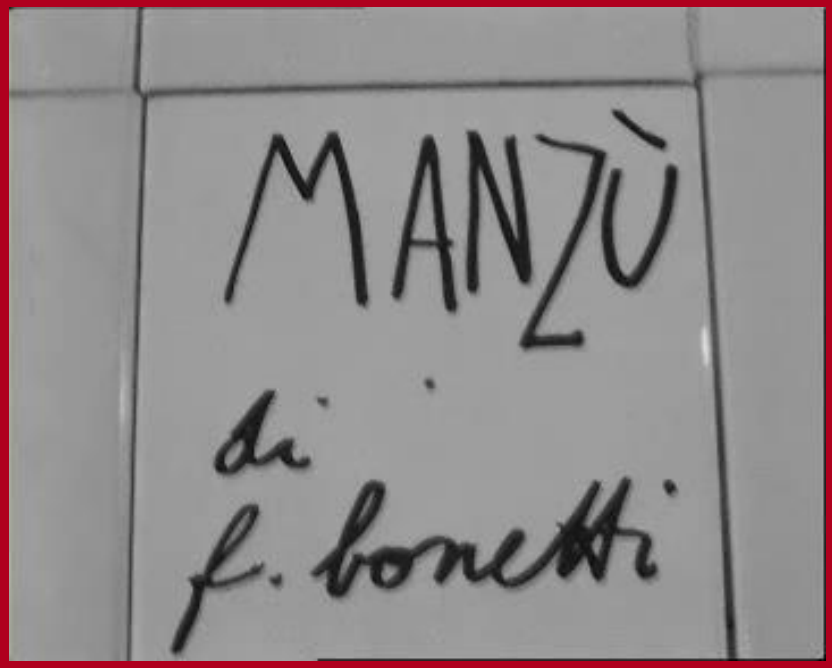
Speech to text at RSI archive

Francesco Veri and Sarah-Haye Aziz

June 11, 2015

Audio Transcription

Come al solito, anche la recente inaugurazione dell'ultima monumentale opera di quell'eccezionale scultore che è Giacomo Manzù, vale a dire la nuova porta del Duomo di Rotterdam ha avuto il sapore di un avvenimento straordinario di risonanza internazionale e per lavori in corso Fabio Bonetti è riuscito ad avvicinare l'insigne maestro bergamasco, a buon diritto ritenuto ormai uno dei più alti interpreti del nostro tempo, artista fra i più grandi del secolo e non solo per la misura del suo talento ma anche per il rigore morale di cui è sempre stato esempio in anni di assorta, tormentata, ispirata attività.



Categorization

Credits: Giacomo Manzù, Fabio Bonetti
Geographic Therms: Rotterdam
Themes: arte, cultura, intrattenimento
Errors

Outline

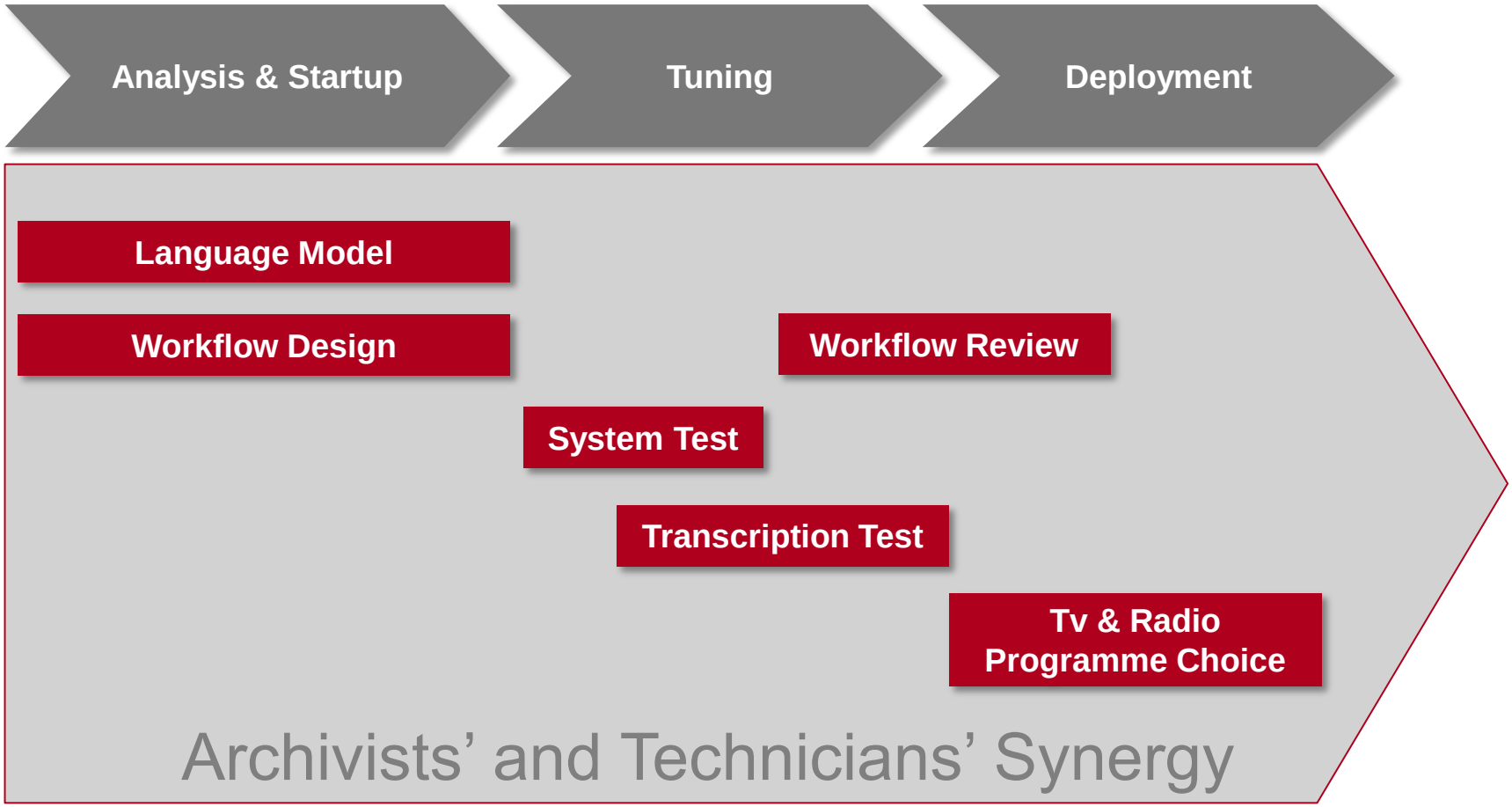
1.	Why an automatic indexing system?
2.	Project timeline
3.	Two paths: system and archivists' workflow overview
4.	Synergy between archivists and IT Staff
5.	Differences between Radio and TV
6.	Solution with Radio
7.	Further implementation

Why an automatic indexing system?

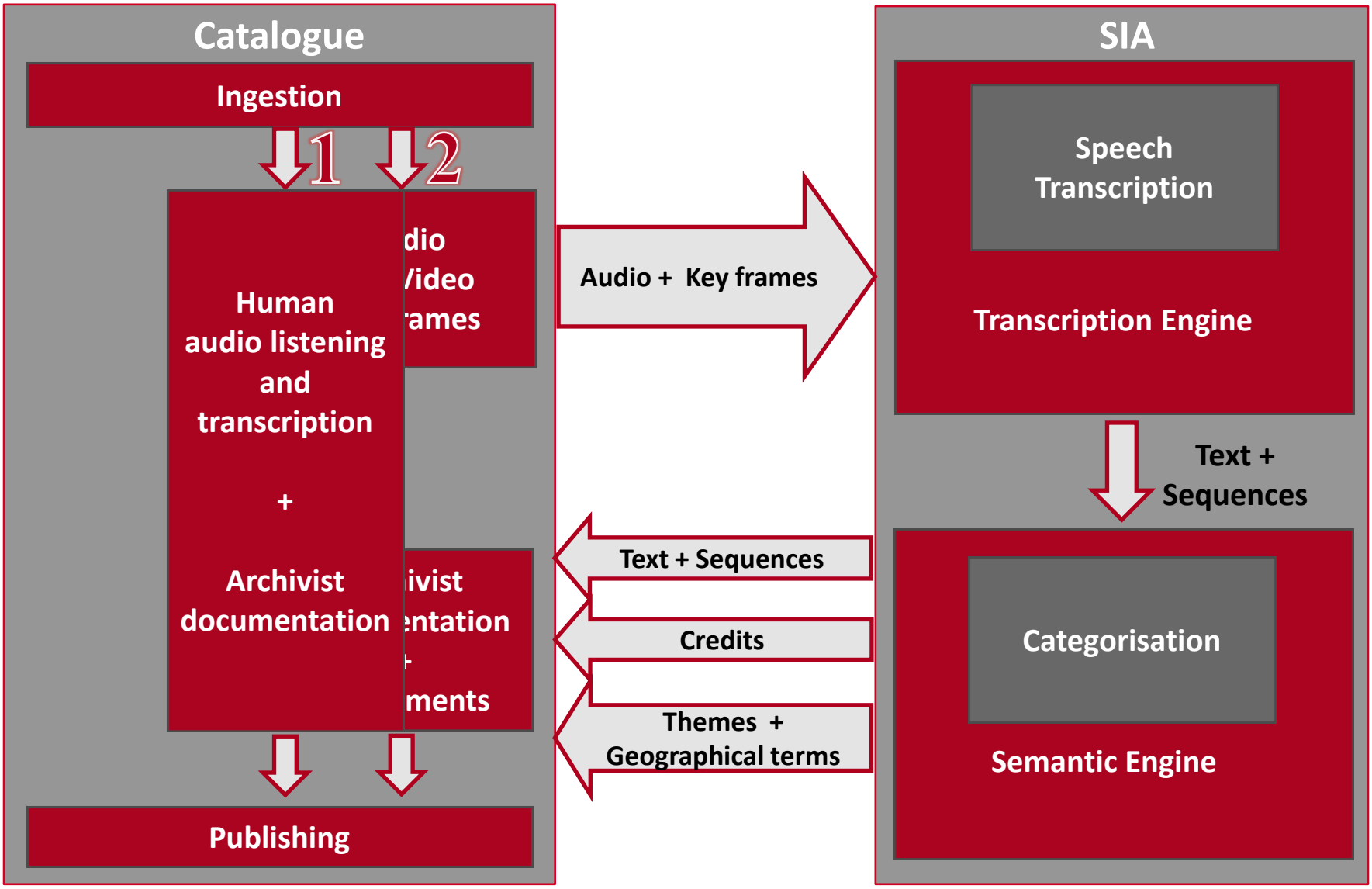
- RSI has a consolidated cataloguing system (CMM) with a well-defined human workflow from 2008
- RSI has plenty of undocumented historical material and no capacity to document it.
- Increase (plus) the documented material by adding an automation but not by substituting (vs) the archivist.

→ **Not *vs* but *plus*!**

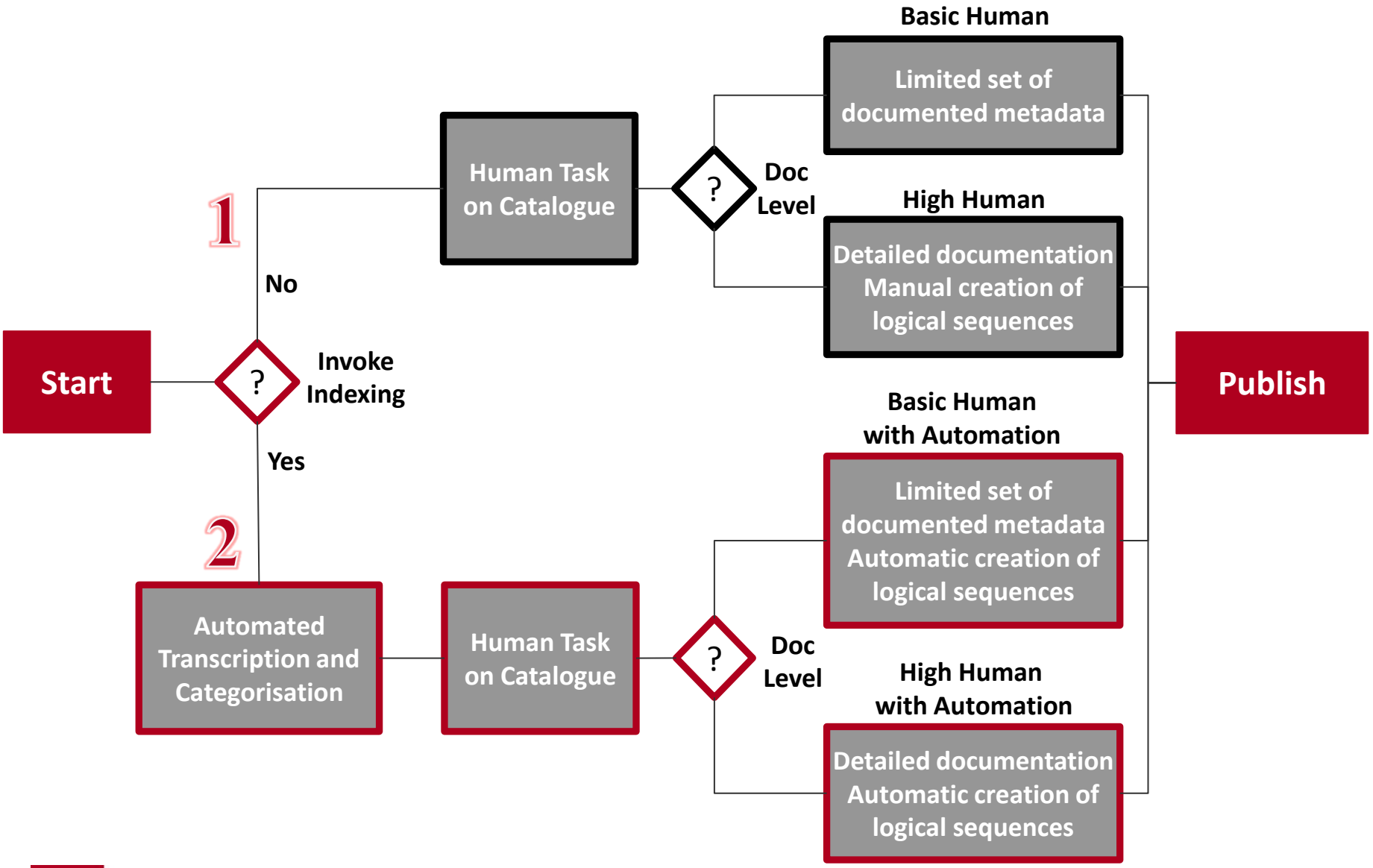
Project timeline



Documenting a material: two paths



The two paths for the archivist



How does it works?

Synergy beween archivists and IT Staff

- Improvement of the automatic indexing system in order to satisfy users' requirements
  - *i.e. using different colours in order to clearly differentiate between human and automatic indexing*
- Archivist becomes familiar with the new system

Does it work? Yes! But…

Differences between Radio and TV

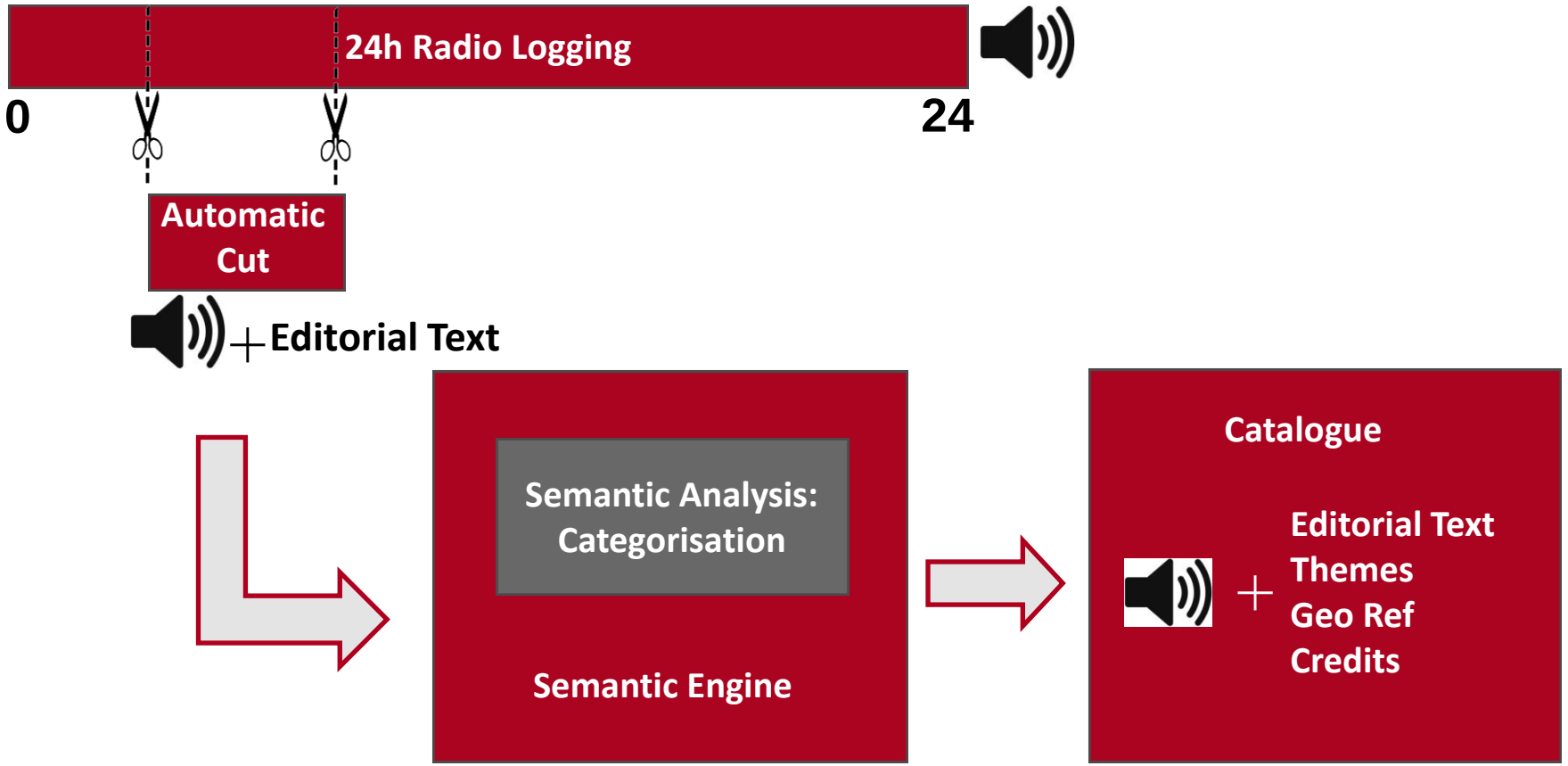
Background Music/Noise → does not help the transcription.

Based only on silences and without key frames, the system creates too many sequences.

Key frames help to locate a change of context.

Speech rhythm and pauses are different between and .

Solution – Capitalize 24h Radio Logging



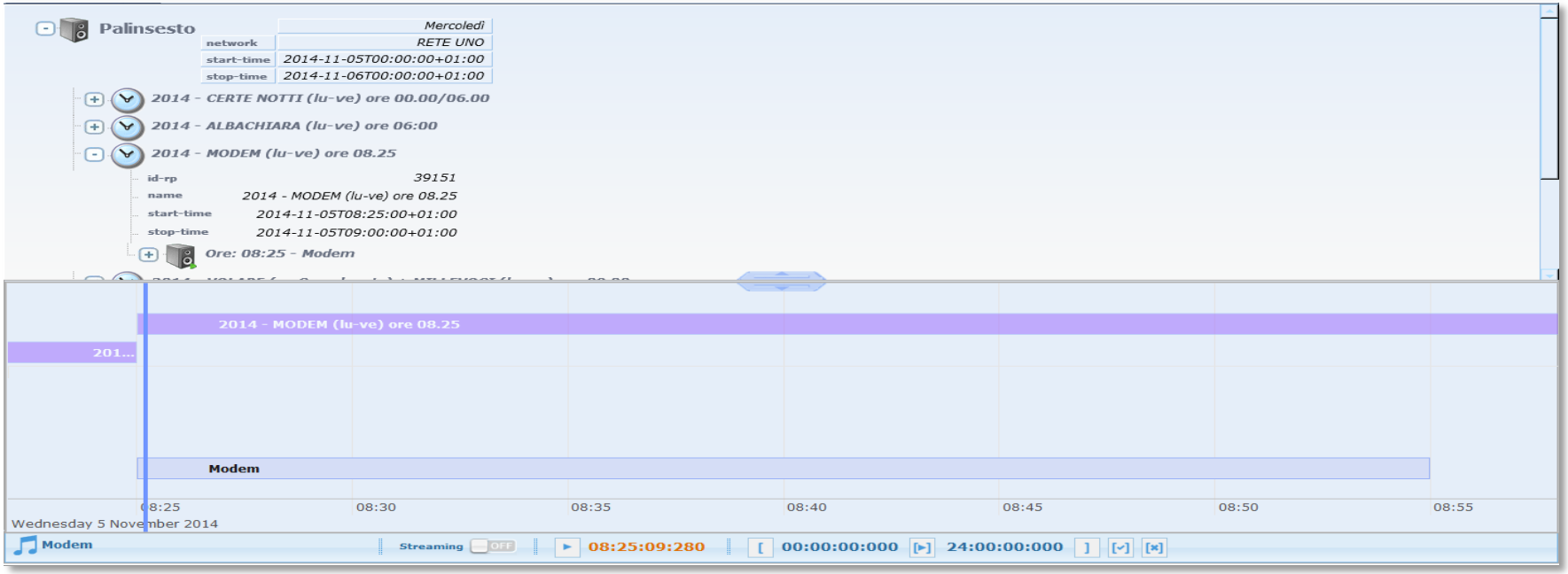
Capitalize editorial texts

- Automatic Semantic Analysis of the editorial text
- Terms not related to our thesaurus become keywords
- Automatic uploading of the editorial texts on CMM

Further implementation

- This solution can be implemented on webpages.Automatic semantic analysis of web content
  - Automatic categorisation of the webpages content

User download from 24h logging



- Audio material retrievable using metadata automatically acquired from the channel schedule
- Download from the 24 h Radio logging