



Roeland Ordelman

Manager Innovation Integration

R&D Netherlands Institute for Sound and Vision

Researcher Multimedia Retrieval

University of Twente

A photograph of a server room with rows of server racks and blue lighting. The image is slightly blurred, showing the perspective of a long aisle between the racks. The racks are filled with various electronic components and cables. The lighting is a cool blue, creating a high-tech atmosphere.

>1,5 Petabyte per year

**BUSINESS ARCHIVE
DUTCH PUBLIC BROADCASTERS**



**+300K hrs AV
173M Euro**

**LARGE DIGITIZATION PROGRAM
IMAGES OF THE FUTURE**



institutional archivists & broadcast professionals

1. MEDIA PROFESSIONAL

First requirements results

Annotation	Search engine	User interface
Archivist Annotations	Fast, stable, reliable	Purchase selection
Technical metadata	'Traditional' program search	Clear result explanation
Recognized people	Filter/search on any field	User guide
Spoken words	Search by example	History function
Shot type (pan, zoom)	Query suggestion	Preview video in results
detailed annotations	Scalable fast and flexible search	Google + advanced option clear and supportive UI



journalists & academic researchers, students, and educators

2. RESEARCH & EDUCATION

First requirements results

Annotation	Search engine	User interface
Notification of new relevant videos	Boolean operators	Provenance of metadata (where does it come from)
Persistent identifiers for fragments	Selection filters & concepts	Virtual cutter
	Links to associated sources (Wikipedia lemma's, etc)	My collections
		Usage statistics
		Share online & offline





3. HOME USER



Professional Archivist

TRADITIONAL APPROACH: MANUAL ANNOTATION

archive is an **authority** responsible for keeping and protecting its content
quality of annotation assures the authority to keep its authoritative position



Professional Archivist

TRADITIONAL APPROACH: MANUAL ANNOTATION

bottlenecks:


huge amounts of data &
fine-grained access requirements
(granularity)

Chapter 10: The Nervous System

10.1	Introduction to the Nervous System	10.1
10.2	The Central Nervous System	10.2
10.3	The Peripheral Nervous System	10.3
10.4	The Nervous System and Behavior	10.4
10.5	The Nervous System and Health	10.5
10.6	The Nervous System and Aging	10.6
10.7	The Nervous System and Disease	10.7
10.8	The Nervous System and Injury	10.8
10.9	The Nervous System and Rehabilitation	10.9
10.10	The Nervous System and Research	10.10

Chapter 11: The Brain

11.1	Introduction to the Brain	11.1
11.2	The Structure of the Brain	11.2
11.3	The Function of the Brain	11.3
11.4	The Brain and Behavior	11.4
11.5	The Brain and Health	11.5
11.6	The Brain and Aging	11.6
11.7	The Brain and Disease	11.7
11.8	The Brain and Injury	11.8
11.9	The Brain and Rehabilitation	11.9
11.10	The Brain and Research	11.10



enhancing quantities and granularity
of metadata by deploying available information
sources and technology

(SEMI-) AUTOMATIC ANNOTATION STRATEGIES

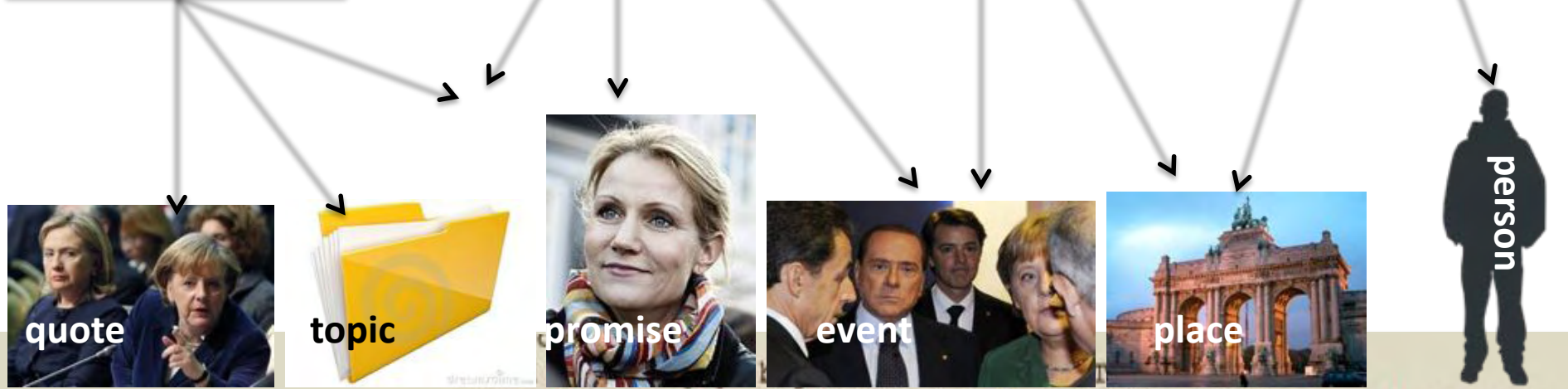
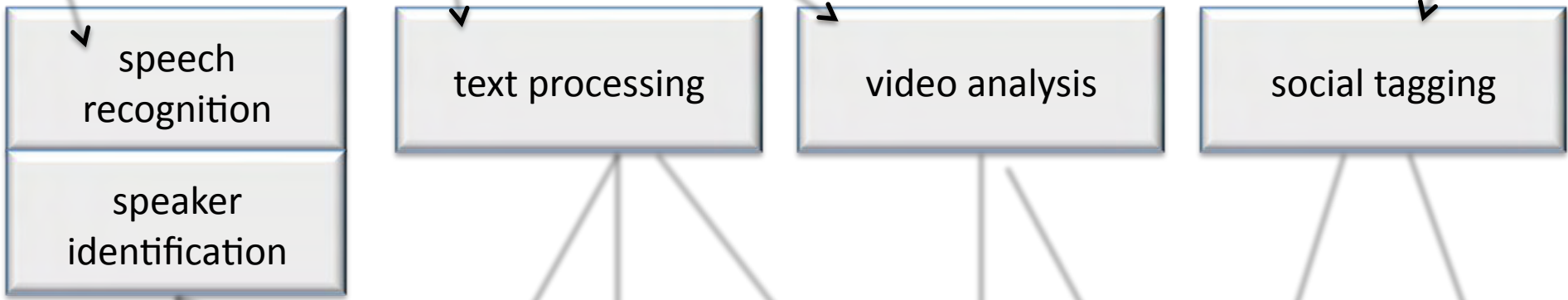
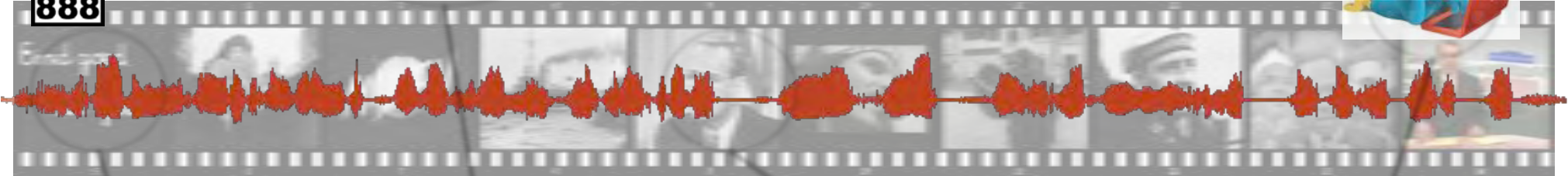
TIME-LABELED ENTITIES

People, Places, Events, Topics, Quotes, Answers, Opinions, Advices, Excuses, Promises, Emotions, ...

Increase annotation levels

- External information sources
 - Subtitles
 - Twitter feeds
 - Information on websites of broadcasters
 - Scripts, auto-cues, etc.
- Technology
 - Audio analysis: speech/speaker recognition
 - Video analysis: visual concept detection, face recognition
- Crowdsourcing
 - Video labeling game

b
TXT 888 The prime minister claimed yesterday that he was not aware ...



METADATA

```
<score AM="-23370.20703" LM="-1678.66980" COMBINED="-25048.87695"/>
</word>
<word wordID="verkiezing" beginTime="7.970" endTime="8.390">
<score AM="-24597.32031" LM="-1744.40735" COMBINED="-26341.72852"/>
```

ANNOTATION TECHNOLOGY



Search on the fragment level



INDEXING AND SEARCH



Time-coded transcript



SPEECH RECOGNITION (SPEECH-TO-TEXT)



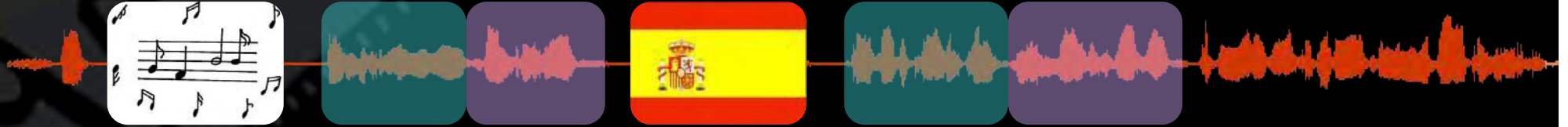
speech/non-speech



language



studio/live



speaker changes

to facilitate easy *browsing & navigation* make use of

STRUCTURE

speaker characteristics



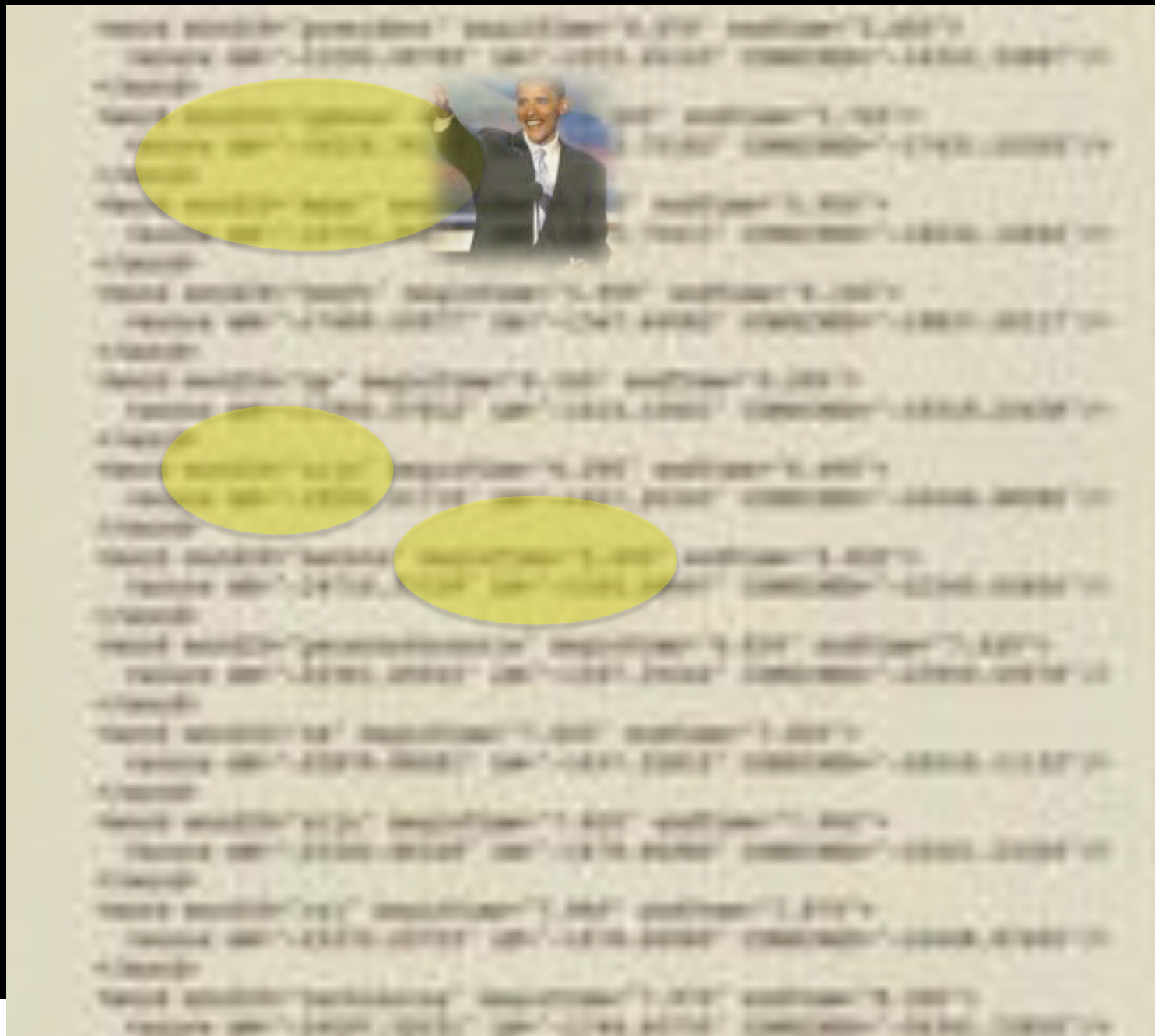
- identity
- gender
- age
- speaking style
- sociolect
- emotional state



Lightly Supervised Speaker Labeling:

- across archive
- deploying metadata priors
- (optionally) using manual supervision

speech recognition





BROWSING COLLECTION ON THE BASIS OF VISUAL CONCEPTS



Visual Concepts
example: singer

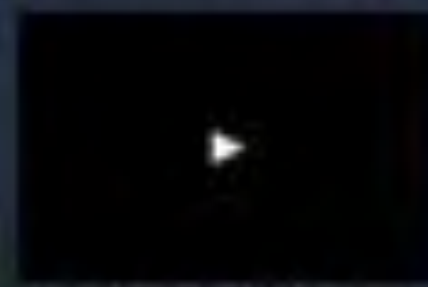
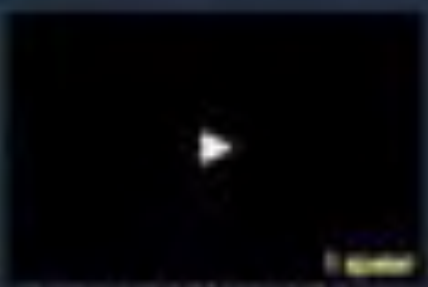
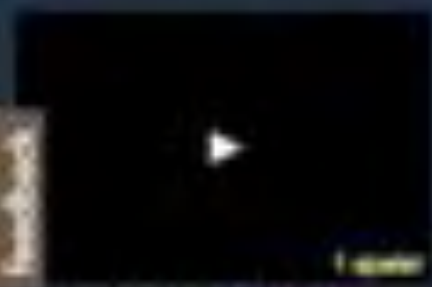
Waisda?

Kies een video.

Vol woorden in die
woordenlijzen wat je niet en hoort.



Je scoort punten als jij en
iemand anders tegelijkertijd
hetzelfde woord invullen.



Waar staat er... [\[link\]](#)

Samenkluis Dory (1996 - 1999)
[\[link\]](#)

De Wendeling [\[link\]](#)

Samenkluis Dory (1996 - 1999)
[\[link\]](#)

Spel mee!

Spel mee!

Spel mee!

Spel mee!

Waisda? hielden

1. Hoort
2. Hoort
3. Hoort
4. Antwoord

Een dagse lieder

1. Hoort
2. Antwoord
3. -
4. -

Waarste lieder

1. Hoort
2. Hoort
3. Hoort
4. Antwoord

Samenkluis bloggers

1. Hoort
2. Antwoord
3. Hoort
4. Hoort

Media Archivist Support System

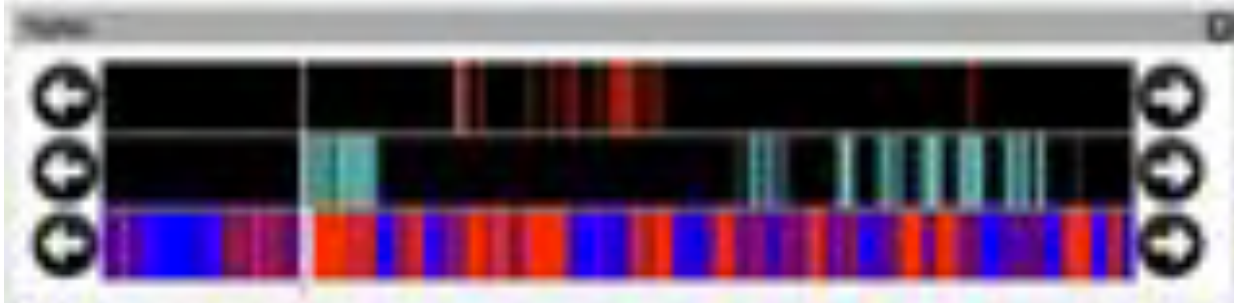
- Manage available annotation sources
 - Authorize and label annotations
 - Add manual annotations if necessary
- Technology aided annotations
 - Quickly skip through video
 - (light?) supervision on automatic annotations:
 - Thesaurus terms suggested using ASR transcripts
 - Speaker/face recognition (bootstrapping)
 - Segmentation (e.g., topic)
 - Feedback loop



- Thumbnail gallery
- Thumbnail 1
 - Thumbnail 2
 - Thumbnail 3
 - Thumbnail 4
 - Thumbnail 5
 - Thumbnail 6
 - Thumbnail 7
 - Thumbnail 8
 - Thumbnail 9
 - Thumbnail 10
 - Thumbnail 11
 - Thumbnail 12
 - Thumbnail 13
 - Thumbnail 14
 - Thumbnail 15
 - Thumbnail 16
 - Thumbnail 17
 - Thumbnail 18
 - Thumbnail 19
 - Thumbnail 20

Text description window containing a list of search results or metadata.

Search for video: [input] [button]



GTAA Trefwoorden

Sort, Filter, Refresh icons

[Geografische namen]

- Kosovo (9)
- Brussel (3)
- Aviano (3)
- Engeland (2)
- Wellington (1)
- Belgrado (1)
- Kathmandu (1)

R&D on video hyperlinking

- navigate between video elements in a source video and related elements in the same video or other multimedia sources
- Linear access: detail-on-demand
- Find relevant anchors in source video, e.g.,
 - based on information need (query)
 - based on topic
- Requires entity labels: FP7-AXES
- MediaEval benchmark evaluation
 - Search & Linking task

R&D on video hyperlinking

- Non-linear access scenario: narratives
- Archive perspective:
 - Create meaningful cross-sections through archive
 - “crowdsourcing”: increase annotation quantities
 - Feedback loop
 - Initial searching & browsing
 - Narrative element recommendation
 - Manual narrative generation
 - Learn/model
 - Recommendation in standard search scenarios



- FP7 Access to Audiovisual Archives
- <http://www.axes-project.eu>



- NL/FES COMMIT – Infiniti
- NL/FES COMMIT – SEALINCMedia