REVEAL THIS

REtrieval of VidEo And Language
for The Home user in an Information Society

http://www.reveal-this.org

FP6-IST-511689

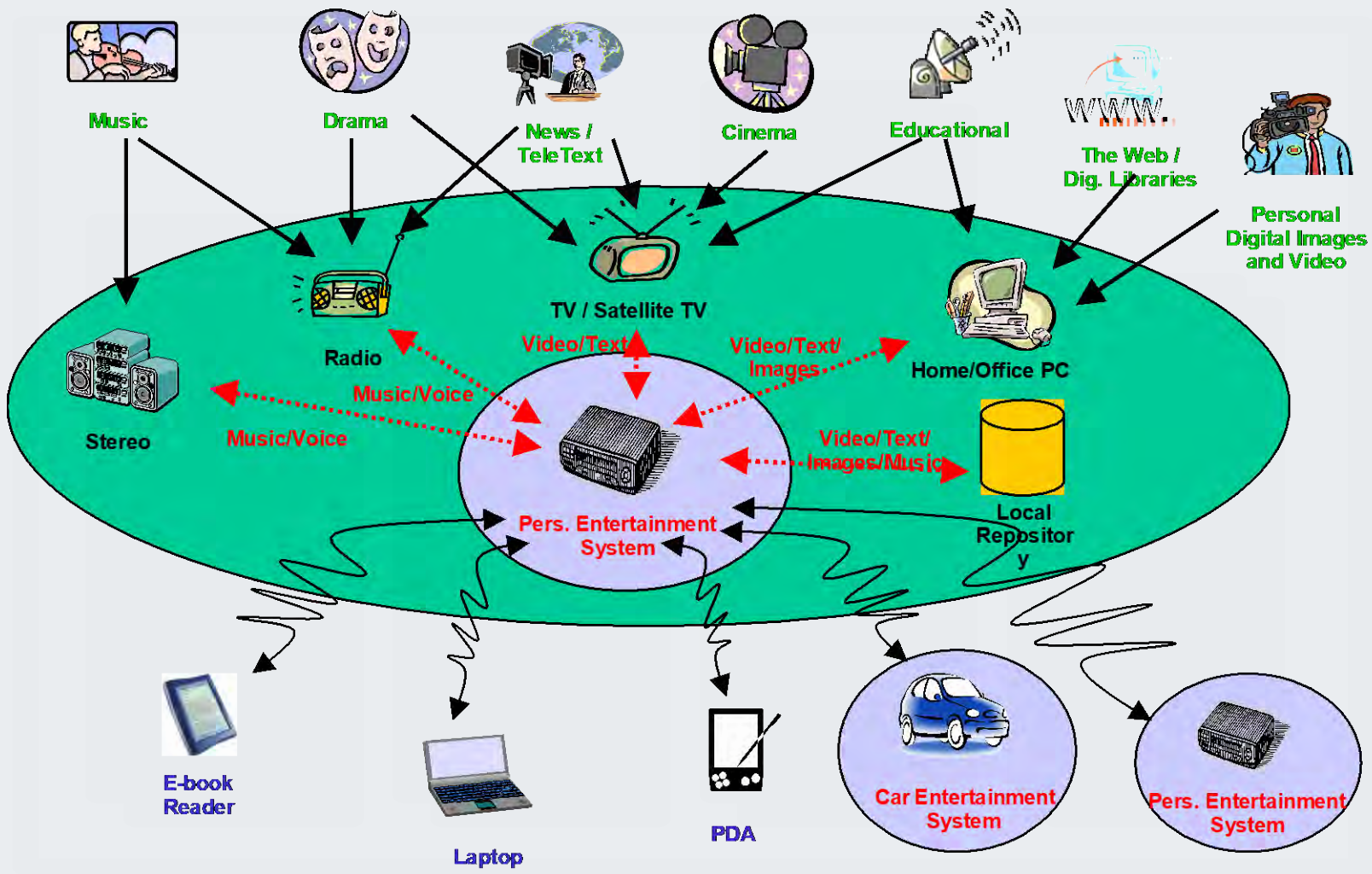*Specific Targeted Research Project* (STReP)
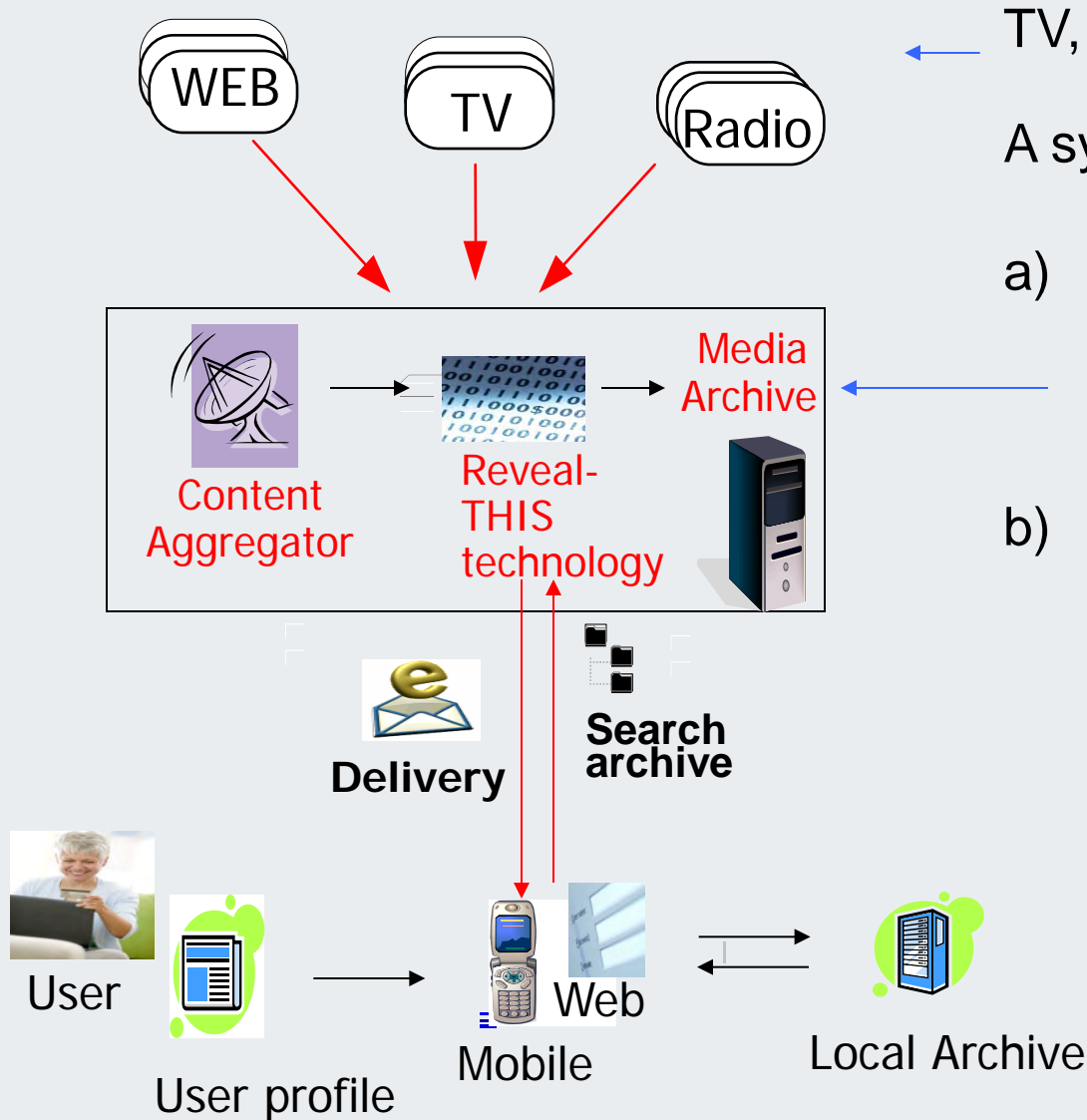
Stelios Piperidis - ILSP

Project objectives

- develop content processing systems that will help European citizens keep up with the explosion of digital content scattered over different platforms (radio, TV, World Wide Web), different media (speech, text, image, video) and different languages

- develop technology able to capture, semantically index, categorise and cross-link multiplatform, multimedia and multilingual digital content, as well as provide the system user with semantic search, retrieval, summarisation and translation functionalities

Vision

Music

Drama

News / TeleText

Cinema

Educational

The Web / Dig. Libraries

Personal Digital Images and Video

TV / Satellite TV

Radio

Video/Text

Video/Text/ Images

Home/Office PC

Music/Voice

Stereo

Music/Voice

Video/Text/ Images/Music

Local Repository

Pers. Entertainment System

E-book Reader

Laptop

PDA

Car Entertainment System

Pers. Entertainment System

Use Scenario

TV, Radio, Web data

A system that offers both types of service :
a) Multimedia and Cross lingual Information Retrieval (pull)
b) Multimedia and Cross lingual information Filtering (push)

Mobile phone and Web interfaces

Potential Users

- **end users** to gather, filter and categorize information collected from a wide variety of sources in accordance with their preferences.
  - o Educated professionals (experienced computer users, with demanding media retrieval needs – pull model)
  - o Laymen (novice technology users with information collection needs – push model)

- **content providers** to add value to their content, restructure and re-purpose it and offer their subscribers, individual or corporate users, personalized content
  - *It is made clear here that REVEAL THIS caters for the retrieval pillar, while leaving aside all possible operations expert editors would wish to perform on retrieved content*

Technical goals

- Augment the content of multimedia documents with semantic information like: characteristic keyframes, identified speakers, faces, entities, topics and fact information

- Develop cross-media and cross-language representations

- Build high-level functionalities including categorization and summarization

- Provide cross-language capabilities (cross-language information retrieval, categorization and machine translation of indicative summaries) based on the latest statistical machine translation technology

- Integrate the above technologies in a unified platform equipped with semantic search, retrieval engine, personalization and delivery mechanisms

Medium specific metadata

➤ **text:** terms/keywords, named entities (e.g. names of persons, places, organizations), topics and facts

➤ **speech:** speech/nonspeech, speakers (e.g. speaker identity), transcriptions and stories

➤ **video and images:** persons, faces, keyframes and categories

Innovative Aspects

➢ semantic enrichment of multilingual multimedia content with topic, terms, entities and event/fact information relevant to user profiles

➢ development of suitable cross-language, cross-media representations

➢ deployment of the above in building categorisation, summarization, search and retrieval capabilities

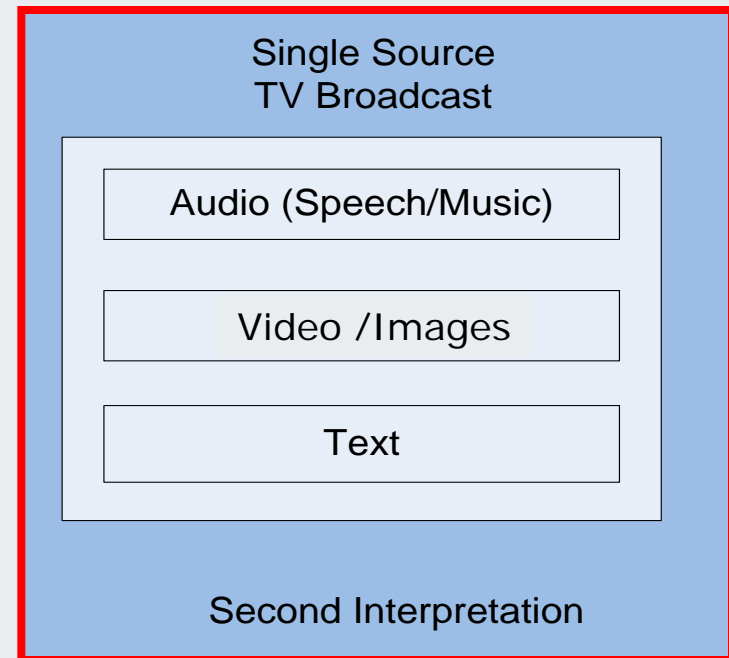# Cross-mediality in REVEAL THIS
the indexing perspective

referring to different sources of information (radio and web text on same topic) → across documents

referring to medium used to convey information within one source (audio, text, image of video segment) → within document

Source A TV Broadcast	Source B Radio Broadcast	Single Source TV Broadcast
News on Elections	News on Elections	Audio (Speech/Music) Video /Images Text
First Interpretation		Second Interpretation

Cross-media processing in REVEAL THIS

referring to medium used to convey information within one source (audio, text, image of video segment)$\rightarrow$ within document

Cross-media indexing
- treat imprecisions & inconsistencies
- process metadata of speech-image-text

Cross-media categorisation
- add to metadata set
- process text and images

Cross-media summarisation
- add to metadata set
- process video and text
- present video+text+audio salient parts using a content/domain specific multimedia discourse grammar

Cross-media processing in REVEAL THIS

referring to different sources of information (radio and web text on same topic) $\rightarrow$ across documents

Semantic retrieval
- retrieval of different multimedia documents for a specific query
- multidocument summarisation
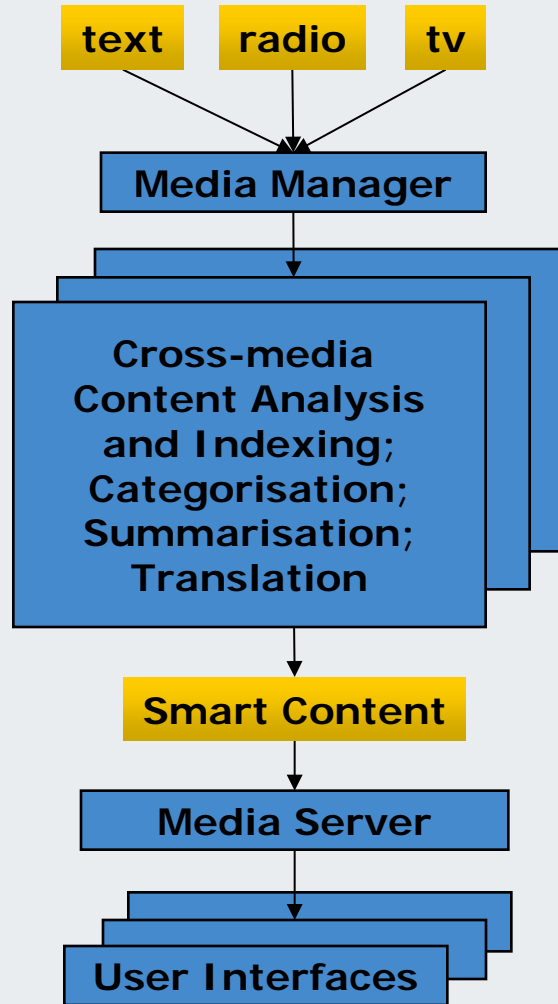
REVEAL THIS in a nutshell

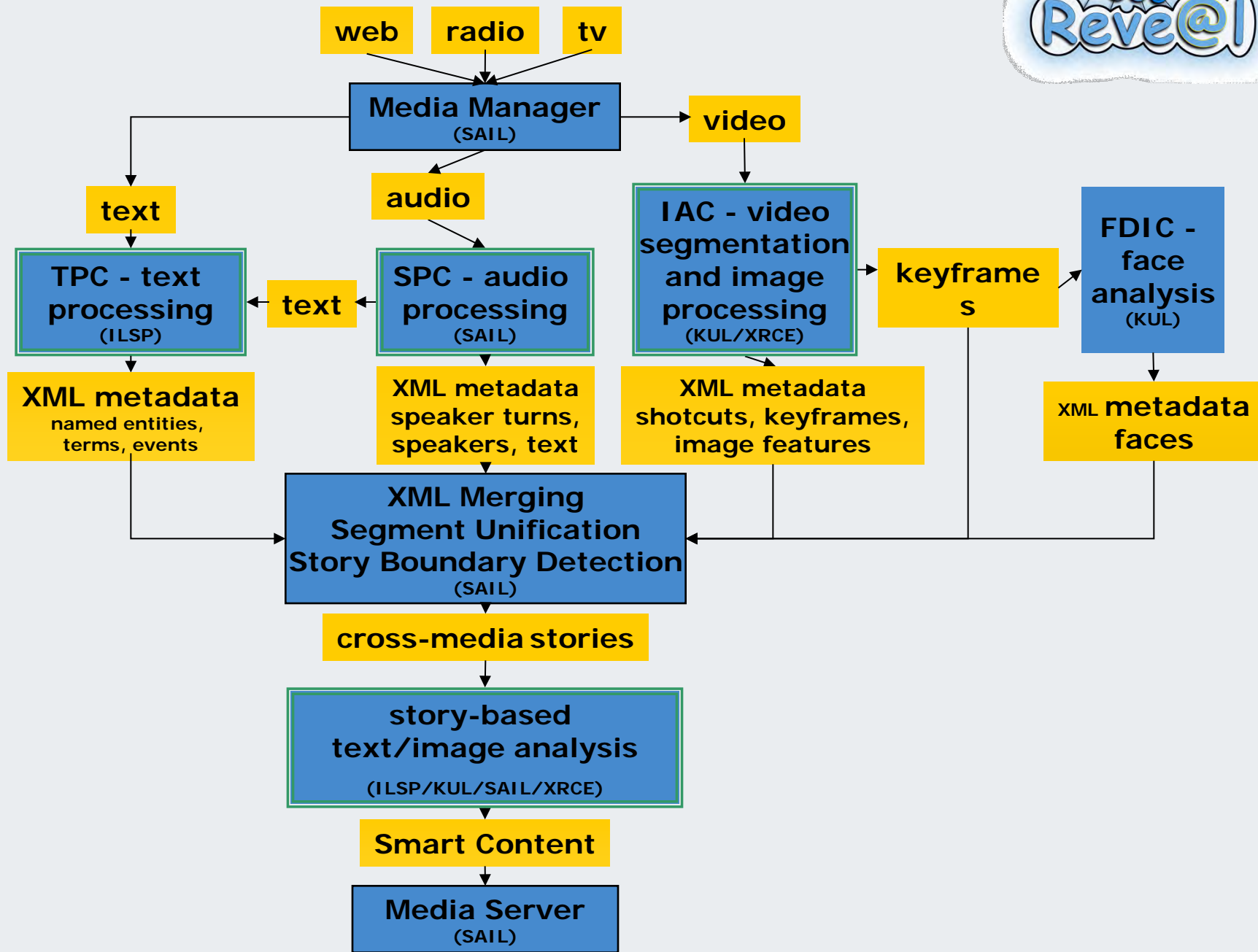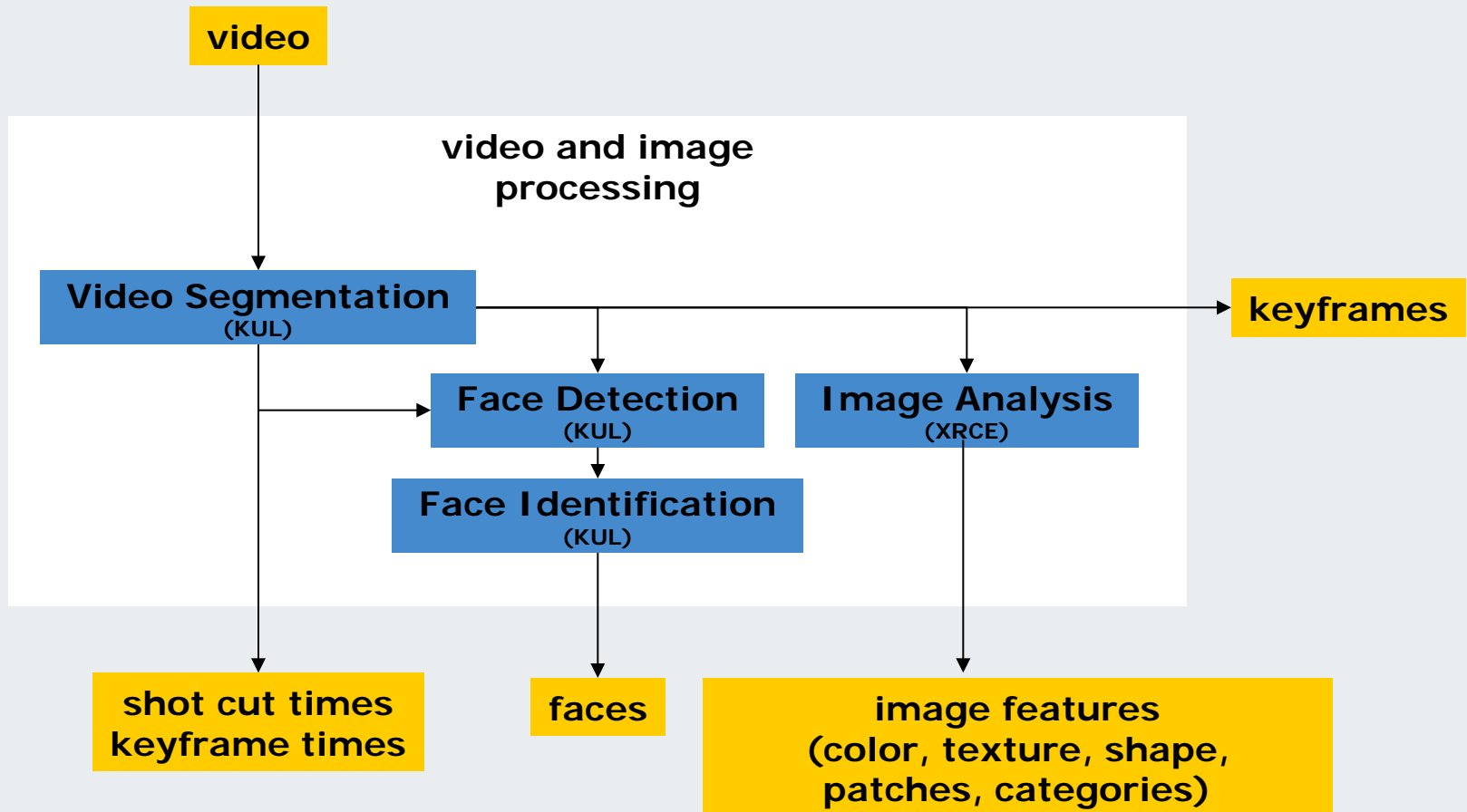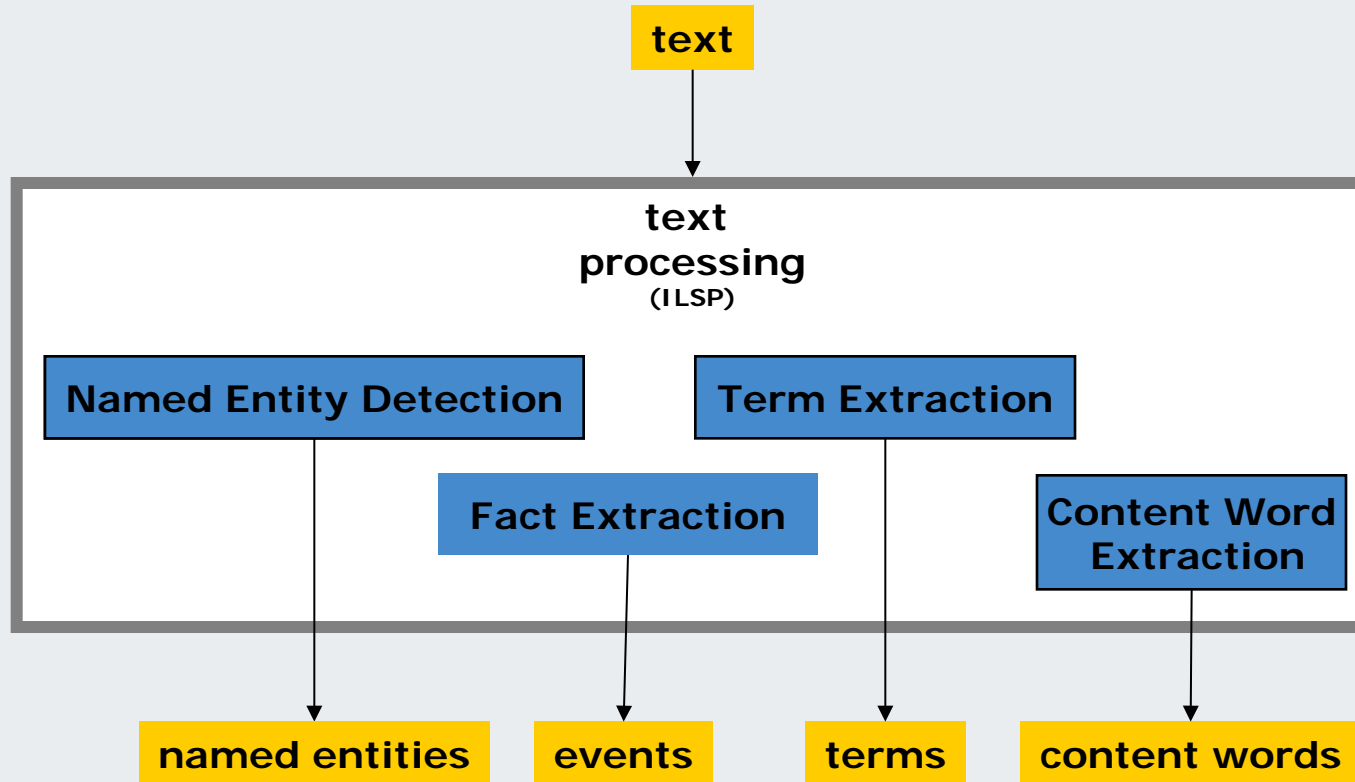|			
Speech recognition in EN	Speech Recognition in EL	Image Categorisation	Face detection & identification
Fact extraction in EN	Fact extraction in EL	Cross-media indexer	Cross-media Categorisation
Text summarisation in EN	Textual summarisation in EL	Scenes and visual summaries	Cross-media summaries
Query Translation	Cross-lingual document translation	Retrieving stories	REVEAL THIS poster

REVEAL THIS

System Architecture

REVEAL THIS Architecture

text → radio → tv

↓

Media Manager

↓

Cross-media Content Analysis and Indexing; Categorisation; Summarisation; Translation

↓

Smart Content

↓

Media Server

↓

User Interfaces

cross-media story
text, named entities, terms, events, speakers, faces, image features

keyframes

story-based
text/image analysis

CMIC – Cross-Media Indexing
(USG)
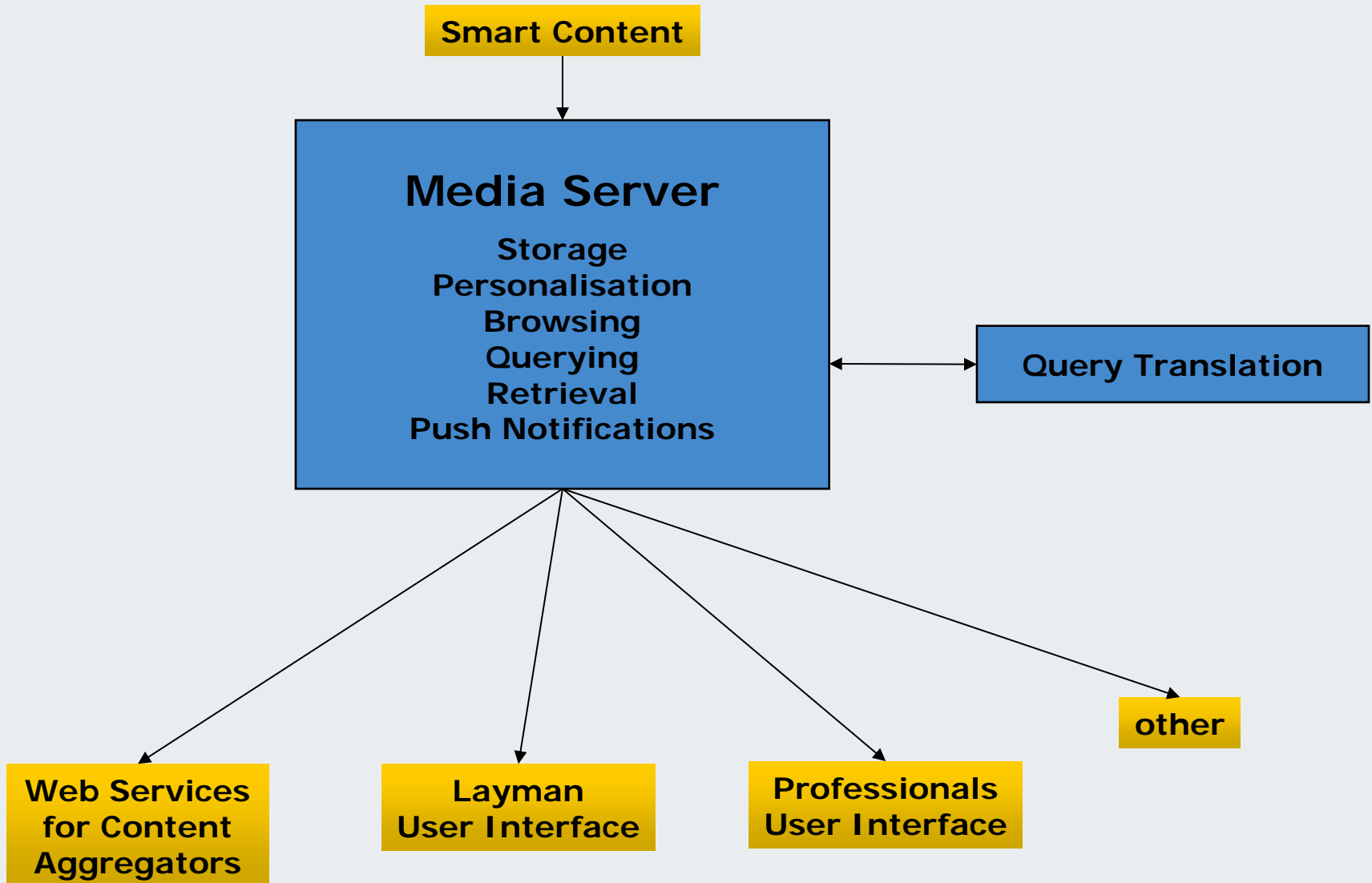
Cross-Media Categorisation
(XRCE)

Cross-Media Summarisation (ILSP)

Visual Summary
(KUL)

Textual Summary
(ILSP)

Translation
(XRCE)

Smart Content

REVEAL THIS Media Server

Smart Content

Media Server

Storage
Personalisation
Browsing
Querying
Retrieval
Push Notifications

Query Translation

Web Services
for Content
Aggregators

Layman
User Interface

Professionals
User Interface

other

Further Information

http://www.reveal-this.org