

European Media Wrapper Round Table 2007

Prestospace & the BBC experience

Ant Miller, BBC FM&T

ant.miller@bbc.co.uk



The next 20 minutes

- Who am I
- What is the BBC archive doing?
 - Internal Projects
 - CAI
 - BBC Catalogue online
 - Open Archive Trial
 - iPlayer
 - PCF
 - External Projects
 - Prestospace
 - Semedia
 - Video Active
 - Some other things we've seen!

Internal BBC Projects 1

the core catalogue and taxonomy

- CAI
 - Computer assisted indexing
 - Working within the existing taxonomy
 - Enhancements to current workflow
- Catalogue Online
 - Infax web interface
 - <http://open.bbc.co.uk/catalogue/infax/>

Internal BBC Projects 2

revised metadata environments

- Open Archive Trial
 - Closed trial leading to “Public Value Test”
 - Fully selected and ‘curated’ content
- iPlayer
 - Broadcast schedule oriented download and streaming services
 - EPG type metadata
- BBC Program Pages
 - One web page for every transmitted program
 - Genre and Format categories
- PCF
 - Structured metadata for interactive content

External Projects: Prestospace

- Large project:
 - Preservation
 - Restoration
 - Metadata Access & Delivery - MAD
 - Storage and Archive Management – SAM
- www.prestospace.org
- 36 Partners
 - RAI, INA, ORF, Joanneum, University Tor Vergata, University of Sheffield, IT Innovations, SSL Ltd, Eurix, BBC, Stream UK, Snell&Wilcox, tudio Hamburg, ACM, Hi Stor, CRCDG-CNRS, OntoText Lab. Sirma, Österreichische Mediathek etc



Prestospace - MAD

- Delivered platforms
 - Modular automated metadata annotation platform
 - Search and delivery platform
- Ontological basis for information extraction (IE), annotation and search
- KIM Platform
 - KIM is an extensible platform for semantics-based knowledge management which offers IE-based facilities for metadata creation, storage, and conceptual search



Side Bar- Ontologies

- Structured knowledge base
- Meaningful relationships between components
- Advantages:
 - Power Information extraction and recall
 - Language Independence
- Disadvantage: Limited Domains?



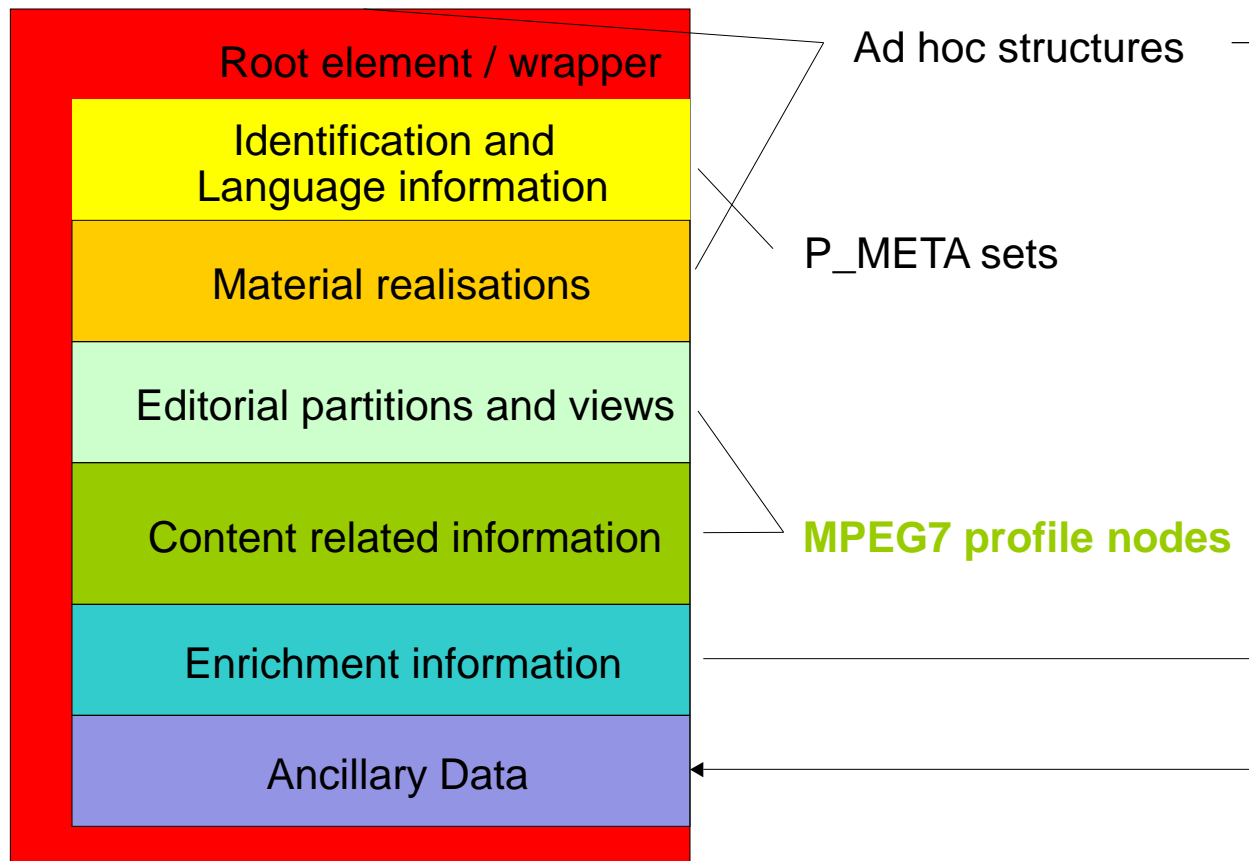
The MAD Data model

(From: From video segmentation to semantic indexing:the Prestospace approach)

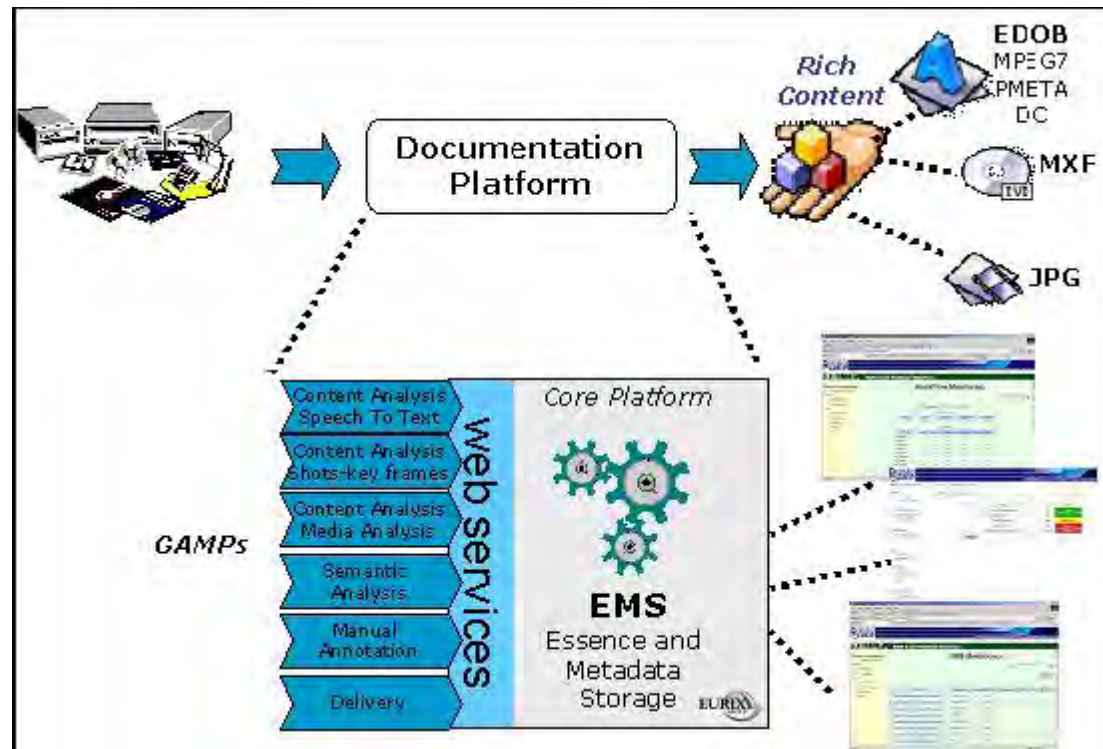
- Supports interoperability with existing standards
 - EBU P/META, ISO MPEG7, Dublin Core, SMPTE MXF and DMS-1
- Results from the analysis of existing users data models
- Imposes a clean separation between editorial and material related information
- Definition of the basic entities and relations
- Supports incremental refinements

Schema of PrestoSpace document format (XML)

(From: From video segmentation to semantic indexing:the Prestospace approach)



Prestospace – MAD Continued



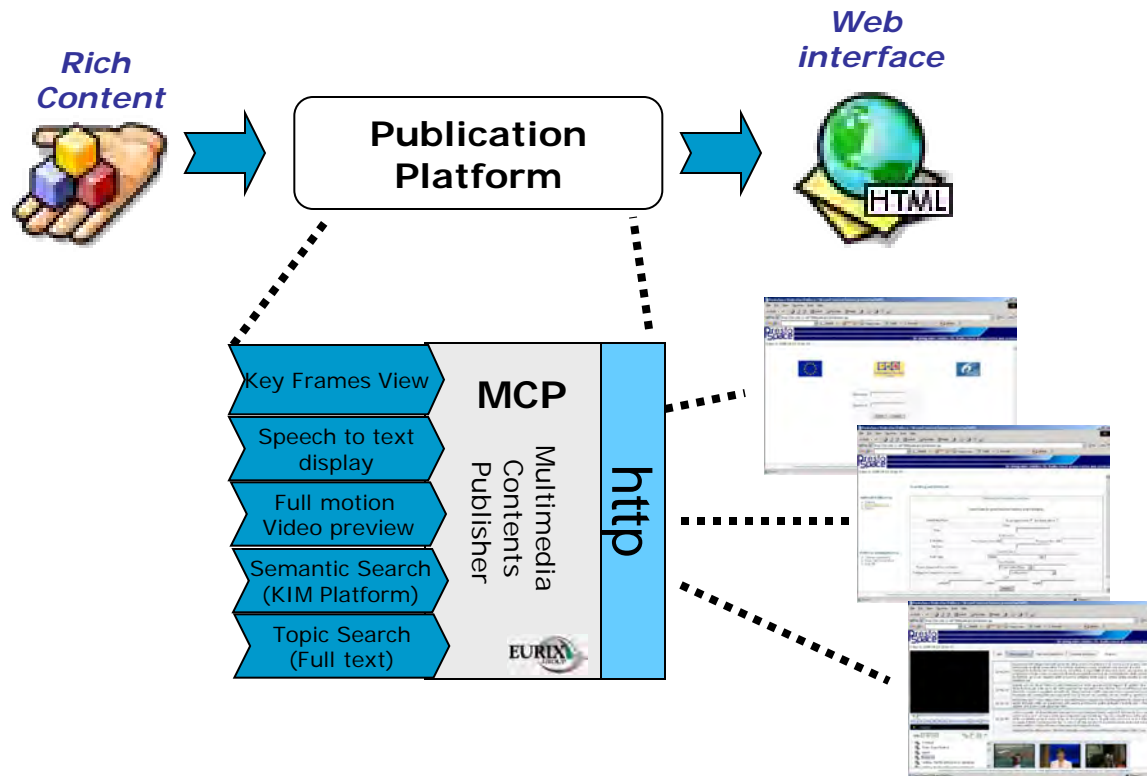
• Process

- Extract text for content
- Cross match to contemporaneous textual sources
- Build structured Ontology

Prestospace – MAD continued

- Result
 - • *Full text search* as usually applied by mostly popular search engines
 - • *Natural Languages Questions*
 - • *Semantic browsing* as navigation through concepts, relations and instances of the ontology

Prestospace MAD- Publication Platform



Prestospace MAD- Publication Platform Continued

The screenshot displays the Prestospace Publication Platform interface within a Microsoft Internet Explorer browser window. The browser's address bar shows the URL: `http://localhost:8080/PublicationPlatform/protected/programmes/tg1_2005-10-2312000/index.html`. The interface features a top navigation bar with the Prestospace logo and the tagline "An integrated solution for Audio-visual preservation and access". Below this, a search bar and a "FIND" button are present. The main content area is divided into several sections: a video player on the left showing a news anchor, a large video player in the center with a timeline below it, and a grid of video thumbnails on the right. The timeline shows various time points, and the thumbnails are arranged in a grid, each with a play button and a timestamp. The left sidebar contains a list of classifications for the video content, including "Other Classification", "Ambiente, Natura e Territorio", "Politica, Partiti, Istituzioni e Sindacati", "Usi e costumi", "Sanita' e Salute", "Giustizia, Criminalita' e Sicurezza", "Musica e Spettacolo", "Sport", and "Sport". The browser's taskbar at the bottom shows various open applications and the system clock.



Key Prestospace Documents:

All available at Prestospace.org

- FROM VIDEO SEGMENTATION TO SEMANTIC INDEXING:THE PRESTOSPACE APPROACH
- Deliverable D15.1 Analysis of current audiovisual documentation models, Mapping of current standards
- Deliverable D15.3 State of the Art of Content Analysis Tools for Video, Audio and Speech
- RitroveRAI: A Web Application for Semantic Indexing and Hyperlinking of Multimedia News*

