



From Archive Broadcasting to Web-based Archive Services: Examples from INA

Bruno Bachimont

Institut National de l'Audiovisuel

Université de Technologie de Compiègne



Some words about INA





Key facts

The 1974 law concerning freedom of communication, creates seven audiovisual societies from ORTF :



INA is an industrial and commercial public institution, founded on January, the 6th, in 1975

INA position: present



- Preservation and valorization of the French audiovisual memory
 - Radio and TV
 - Soon (!), Media part of the Web
 - Every content that is temporal and that is streamed (hertzian, cable, web, etc.).
- Two main modalities:
 - Audiovisual legal deposit:
 - All broadcasted programs (public or private), with a French production;
 - Archiving the public broadcasting channels:
 - Radio France, France Télévision.

Two types of exploitation



- Legal deposit : a *national* heritage
 - Accessed by researchers to found their scientific inquiry;
 - Tools provided by INA to help researchers consult AV materials.
 - ➔ Contents do *not* belong to INA, nor to State, but to the Nation: cannot be sold and re-exploited.
- Archiving Public channels: a *public* heritage
 - Sold for re-exploitation to producers, journalists, channels, etc.
 - ➔ Contents belong to INA (and many rightholders): they can be sold and re-exploited.

Some numbers:



- Professional Archives :
 - 1,1 million hours of radio and TV programmes;
 - 575 000 hours / Radio ;
 - 535 000 hours / TV
 - + ~ 50 000 hours / year
- Legal Deposit :
 - 930 000 hours (NB: professional archives are partially included in legal deposit).
 - 430 000 hours / TV
 - 500 000 hours / radio
 - + 500 000 hours / year
- 2,5 millions documents covering 113km of shelf space;
 - 8km / year
- 133 years for watching or listening all archives;
- 220 000 hours digitized and accessible on line (B2B).

INA situation: future



- Making accessible the French Audiovisual Heritage to main public and end users:
 - Technologically possible: end users have equipment required to receive and treat contents
 - Culturally necessary:
 - demand for public memory;
 - TV is now an old media: more and more programs have never been watched by contemporaneous people;
 - INA has a monopole on this kind of programs.

From an economical point of view



- Making money with archives:
 - Selling programs or excerpts for re-broadcasting:
 - Small increase in the future because:
 - Broadcasting is already 24h/day and 7d/week;
 - People capacity to watch TV is not extensible (3h30/day !)
 - New incomers are not rich enough to invest in the INA high quality archives
 - Selling programs or excerpts for downloading:
 - From program guides (duration constraints) to catalogs (organisational constraints);
- Conclusions:
 - Slow market increase for broadcasted excerpts.
 - Necessity and opportunity to consider new markets offered by content distribution.

Our perspective



1. *Structuring*

- archived contents into *collections* composed of contents and excerpts selected according to their meaning.

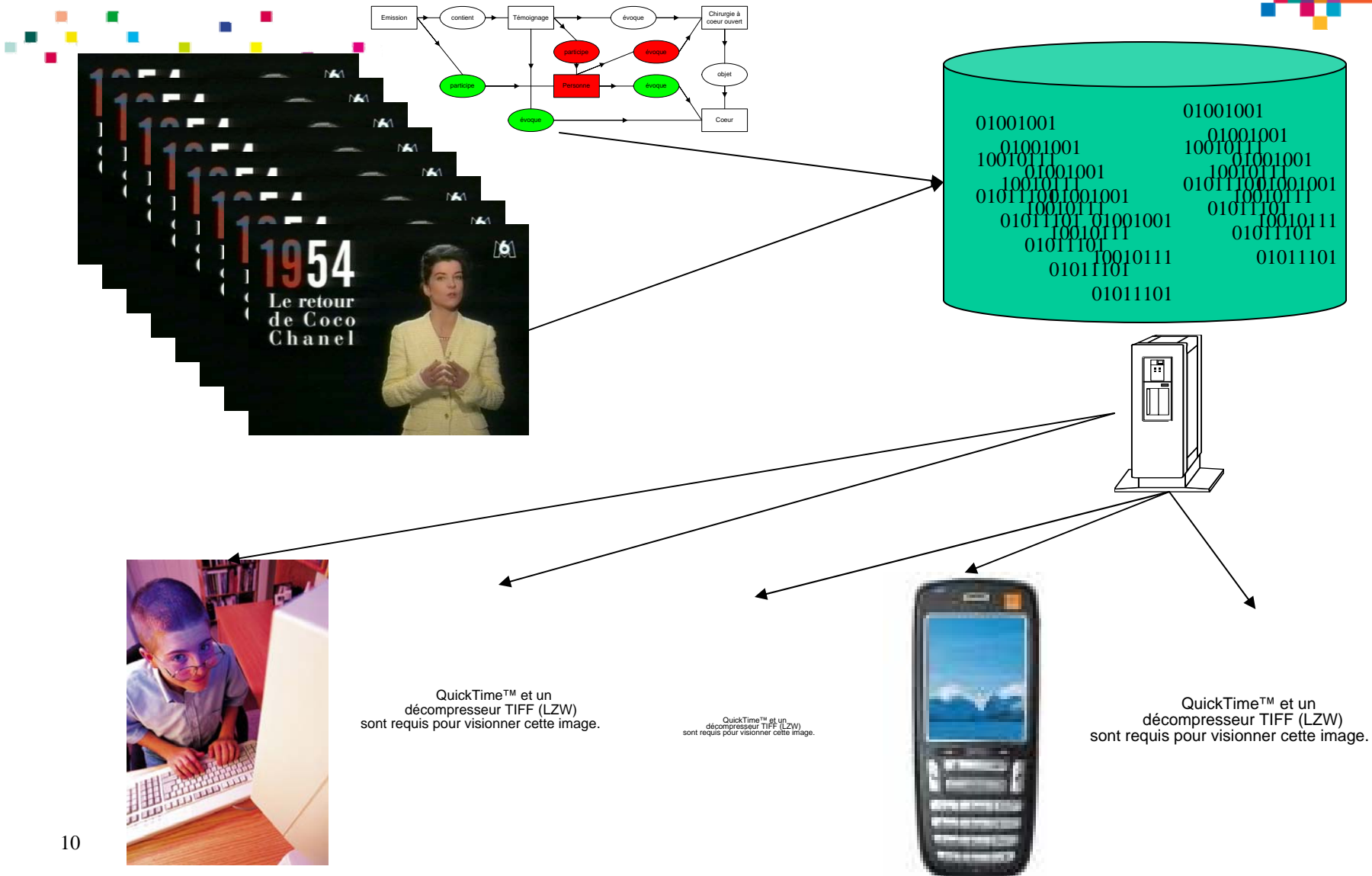
2. *Documenting*

- Collections in order to make intelligible and usable selected contents.

3. *Repurposing*

- Using documentation to edit contents according to
 - the delivering device and format,
 - the targeted audience.

Multiple perspectives





New business, but new
practices



Digital archive distribution



Distributing archives on line is not only a new market, it is a new business and a new professional practice.

- Nature of indexing is changing;
- Status of content is no longer the same;
- The archival stance should be adapted.

From index to metadata



- Traditionally:
 - Index:
 - To point the document relevant to the aimed use;
 - Purpose:
 - To retrieve documents as they are;
- Current trend:
 - Metadata:
 - Information that makes content useful and usable
 - Main property:
 - Can point to any arbitrary part of content;
 - Purpose:
 - To select a segment to reuse it in a new context.

From document to resource



- What is called “document” is from now on a “container” from which one can extract segments.
- Those segments are resources whose meaning is determined by use conditions and exploitation context.
- Initial coherence of the source document is no longer what constitutes content and defines it.

From retrieving to editing



- What motivates indexing is information retrieval:
 - Find out documents that contain and express the desired information.
- What motivates metadata is to select resources in order to create new contents.
 - The resource value is no longer a consequence of its original context (production context) but a consequence of its exploitation context.

From indexing to repurposing



- Objective:
 - Find out resources to create new contents
- Constraints:
 - Segments are decontextualized from their original context;
 - Need for recontextualizing them in their exploitation context.
 - Specific rendering device;
 - Coherence with other segments of the final created content.

Creating new information and new content to integrate selected resources.

Several perspectives



- The Genealogical stance:
 - Resources are selected to be enriched and situated in their original context;
 - The editing process has to publish the archival knowledge..
- The Amnesic stance:
 - Selected resources are reused without any consideration of their original meaning.
 - The editing process is a new creation using archives as stock shots.



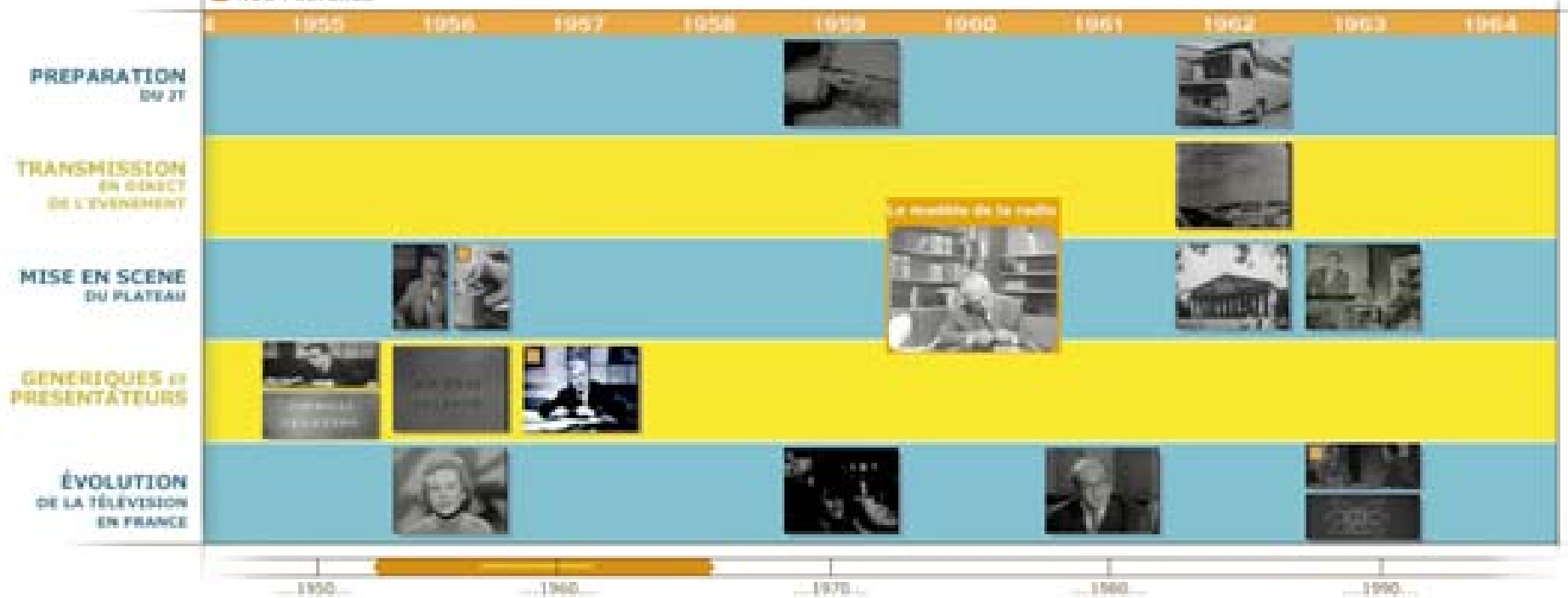
Histoire du journal télévisé

1960 : le modèle de la radio

JT, RTF, 24.11.1960

Claude Dargat semble un journaliste particulièrement décontracté en faisant écouter une correspondance d'Alger sur la tentative d'insurrection menée par Lagailarde et Ortiz. C'est qu'il vit et pense son travail sur le modèle de l'information parlée de la radio : l'image du direct est encore très rare et l'attitude même du journaliste n'est pas pensée pour capter l'attention du téléspectateur.

NOS PRÉFÉRÉS



Main issues



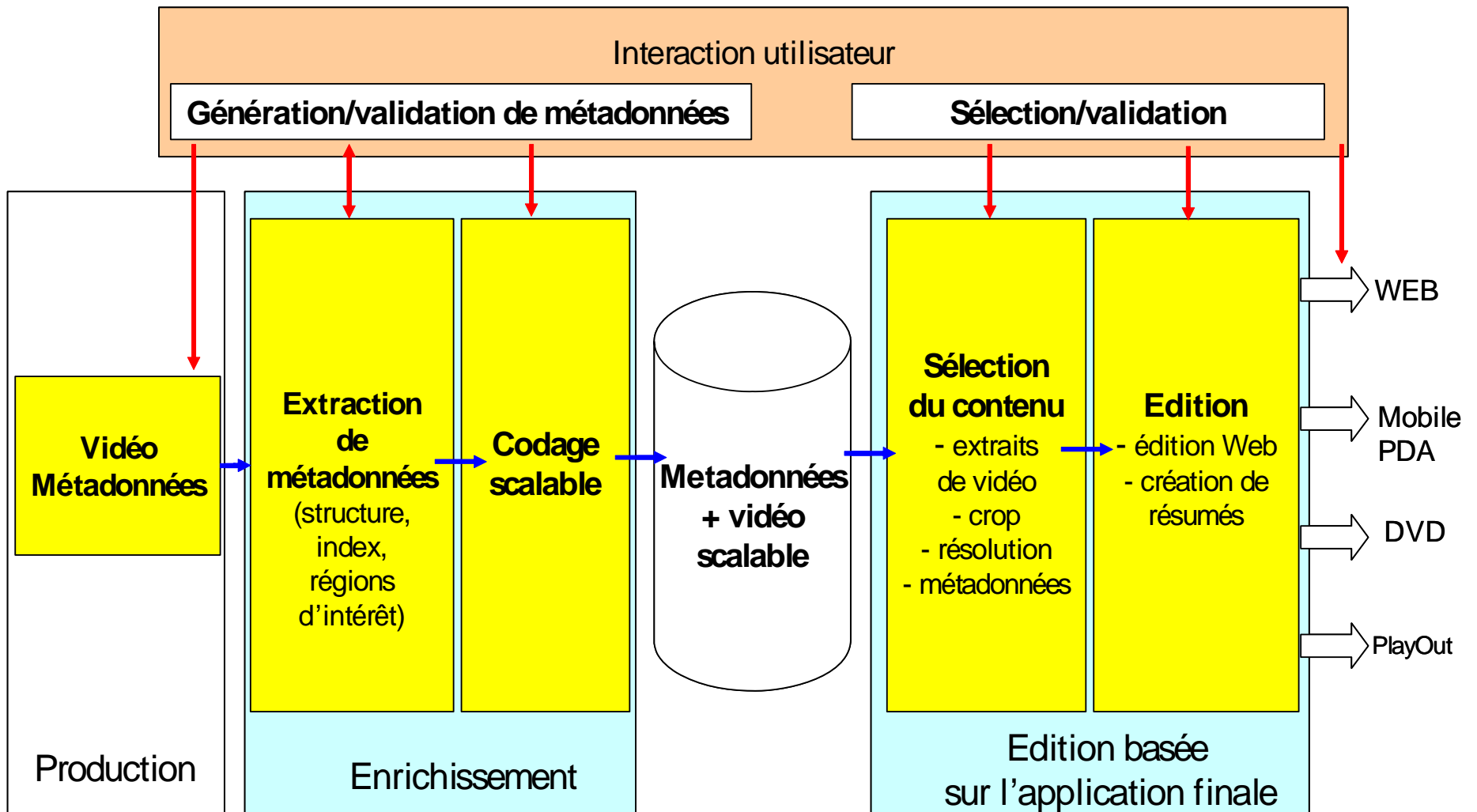
- Technical issue:
 - Dealing with content repurposing on a big scale:
 - Mass editing
 - Need for industrial solutions to manage huge amounts of data as well as complex process
- Cultural issue:
 - Dealing with content repurposing in a “faithful” way:
 - Genealogical editing
 - Need for technical solutions to manage complex knowledge representations and document structures.



Industrial Aspect



Repurposing



Content Share de Thomson GV



The screenshot displays the Thomson GV Content Share interface. At the top, there are tabs for SEARCH, INGEST, PREPARE, EDIT, ARCHIVE, and BROADCAST. Below these are buttons for RECORD, LOG, and APPROVE. The main interface is divided into several sections: a search panel on the left, a central video player, a metadata panel on the right, a video preview strip, and a waveform display at the bottom. The search panel includes fields for NAME, LOCATION, and KEYWORD, along with SEARCH and CANCEL buttons. The video player shows a scene of a person in an orange jacket. The metadata panel displays details such as VELO 4553JUIDS, LOCATION: dvs11, Archive, dvs22, and REEL: 234567. The video preview strip shows a sequence of frames. The waveform display shows audio levels over time. At the bottom, there are buttons for SUBMIT, APPROVE, COPY, ARCHIVE, and TRANSFER.

Include Full text capability

show content according search / task

Allow video segmenting and trimming

Speed-up production with pre-editing

Stream line the operation matching workflow

Handle complex content management

Aggregate Metadata for rich media repurposing

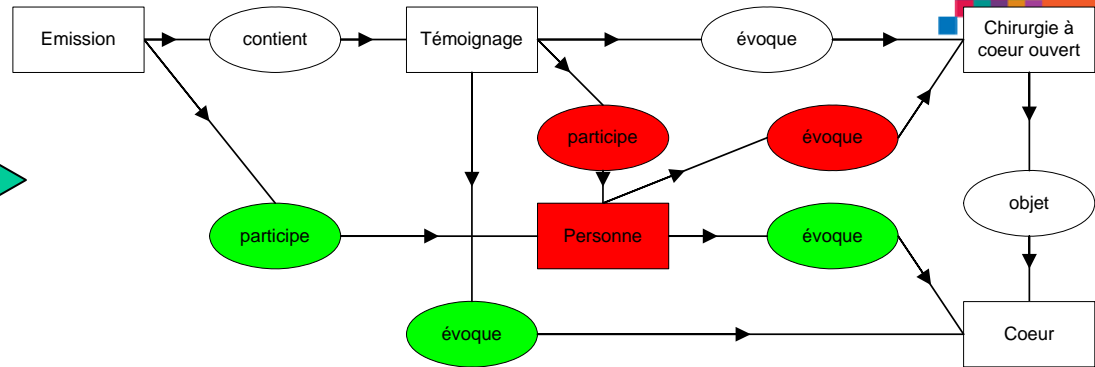
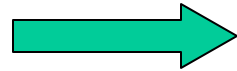
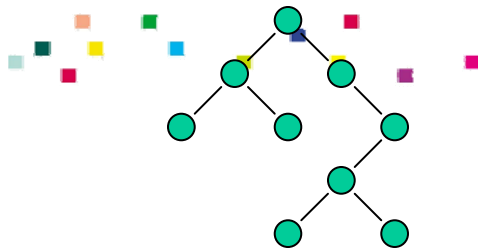
Facilitate operation with task based action



Research aspects



Using what we know



Segmentation | Résultat recherche

PlateauInterview
PlateauImage
Dessin
PlateauAnnonce
Reportage
PlateauImage
Interview
ScouencePlateau
Reportage
Interview
Reportage
Interview
Reportage

Contrôles Média
Durée 0h 54mn 18s 88
Pause
Stop
1

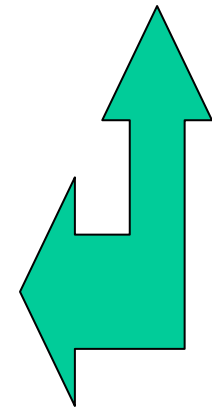
Propriétés Segment
Titre Interview
Durée 00 00 11 16
Début 00 23 39 92
Fin 00 23 51 08
Annotation: Sandy Casar dans le reportage
Associer image

00:23:39.800 00:23:39.920 00:23:51.040 00:23:51.080

00:23:43.408

1 pixel = 4 samples

Créer segment
Supprimer segment



DPALES V2.5-08 (040303) Session OPALES ouverte pour ais sac (pdy auteur -AVCyclisme)

Fichier Espace privé Vidéo Historique Administration Aide

sb : 00:21:35:11, Cur : 00:23:42:16/00:23:41:19, Fin : 00:25:10:23, cache : 202


Outils Recherche Points De Vue Explorer Espace privé Préférences Editeur de diaporama ?

Stade 2 : Descriptif de l'archive V_188 : Stade 2

Outil : Editeur de graphes conceptuels

Reportage Paris-Nice

Reportage Paris-Nice



Journaliste : JeanReneGodart

CoureurCycliste : SandyCasar

EntretienInterview

aCommeAgent

aCommeObjet

estClasseCoureurGeneral

RangEpreuve

epreuveConsideree

position

CourseAEtapes : Paris-Nice

Integer : 2

Consultation de l'ontologie - Ontologie de l'Audiovisuel et du Cyclisme

Navigation

Concepts Relations

Individuel

ObjetBiologique

Personne

PersonnelEpreuve

CommissaireDeCour

DirecteurEpreuve

Journaliste

MedecinEpreuve

Spectateur

PersonnelEquipe

MedecinEquipe

CoureurCycliste

DirecteurSportif

PersonneDeProduction

Participant

ObjetGeographique

ObjetManufacture

Personne

Entités parentes :

PersonnelEquipe

Instances :

SandyCasar

Informations sur la notion sous-jacente :

Notion parents :

PersonnelEquipe

Similarité avec le parent :

Personne employée par une équipe

Similarité avec les frères :

Le rôle de l'employé est précisé

Différence avec les frères :

Il participe aux épreuves

Différence avec le parent :

À un statut particulier dans l'équipe : il participe aux épreuves.

Chercher une entité

Confirmer Annuler

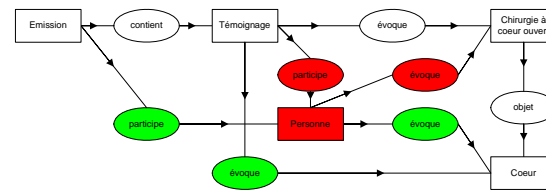
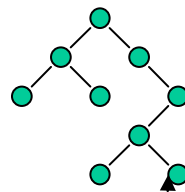
Raccourci vers opadesv.exe

Raccourci vers opades.bat

Démarrer

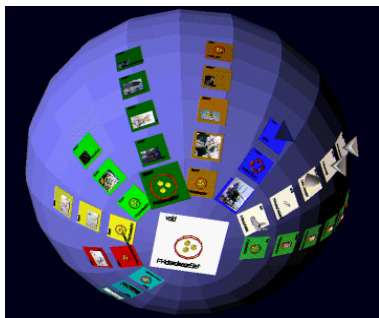
16:53

And handling content



*Reasoning
Representing
Manipulating*

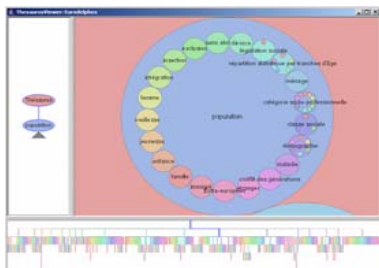
Bla bla bla Bla bla bla
Bla blaBla bla bla Bla bla bla
Bla bla
Bla bla bla Bla bla bla
Bla bla Bla bla bla Bla bla bla



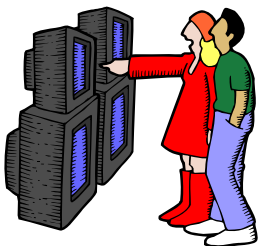
Visualizing



Extracting



Interpreting



Conclusion



- Opportunity coming from the shift to content distribution.
- Paradigm shift in documentation and archiving;
- Needs for tools to represent content and meaning in order to repurpose it in a genealogical fashion

A question for the Semantic Web ?