

# **L'indexation de documents audiovisuels dans ViSA**

Andrée Tiberghien

UMR ICAR

Université de Lyon, CNRS, INRP

# Plan

- Présentation de ViSA
- L'indexation de documents
- L'indexation des unités vidéo dans ViSA

# Présentation de ViSA

# Base de données **ViSA** **V**idéo de **S**ituations d'enseignement et d'**A**pprentissage

## Objectifs

- mettre à la disposition de la communauté des scientifiques des enregistrements vidéos de situations d'enseignement et de formation avec les documents qui leur sont associés.
- collecter de nouveaux enregistrements vidéographiques ;
- « outiller » les communautés de chercheurs et de formateurs pour analyser, gérer et traiter les enregistrements vidéos.
- Jouer un rôle patrimonial concernant les pratiques de l'école.

# Caractéristiques des enregistrements vidéo de situations de classe

## *Caractéristiques matérielles des enregistrements*

- très longues durées
  - Entre quelques heures souvent des dizaines d'heures voire des centaines pouvant concerner les mêmes acteurs (professeur, élèves, classes)
- séquençage imposé par l'institution « école » (durée des séances) et non par les chercheurs
- associées à d'autres données
- souvent réalisés avec des caméras amateurs (ou semi professionnelles) par les chercheurs eux-mêmes ; caméras quasi-fixes pour des raisons méthodologiques (analogique puis numérique DV)

# Caractéristiques des enregistrements (2)

## *Caractéristiques légales*

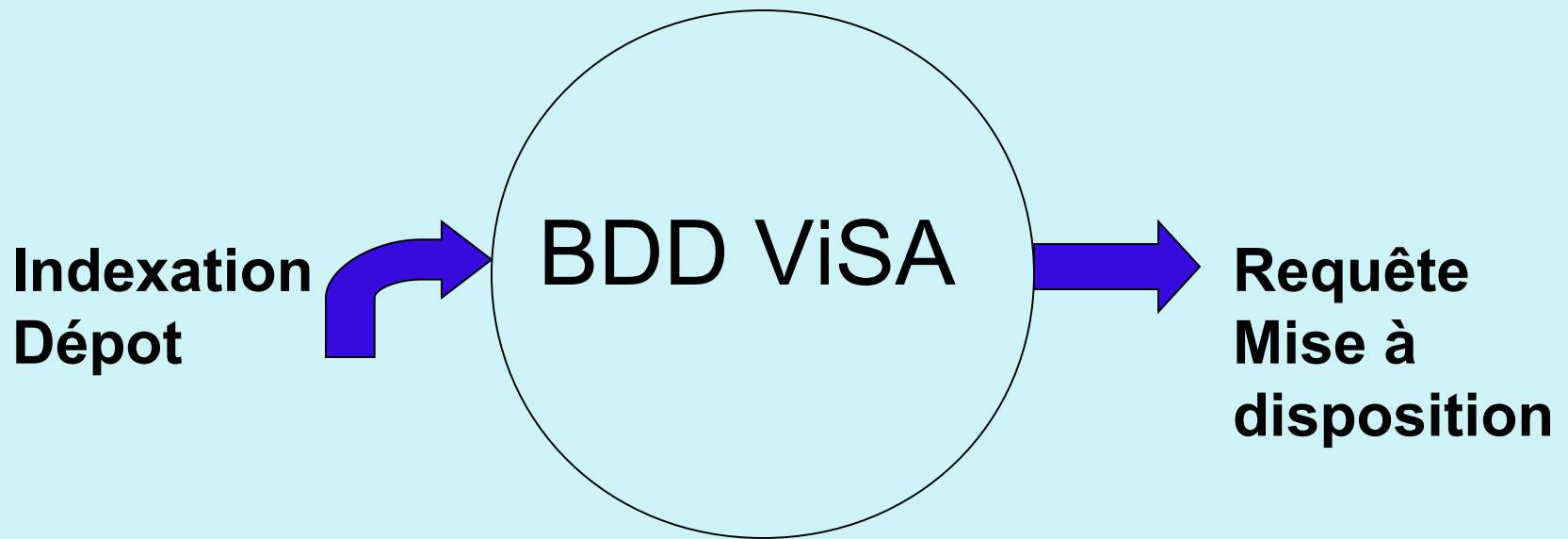
- Des autorisations doivent être établies pour filmer
  - dans les classes
  - des personnes (majeures et mineures)

## *Caractéristiques éthiques*

Des jugements sur les personnes filmées peuvent être fréquents et constituer une atteinte au respect de la personne (problèmes importants pour la formation de maîtres)

 **Actuellement ViSA est limité aux chercheurs**

## Instrumentation: vie de la communauté scientifique



# Choix de la structure de la base ViSA

Un chercheur qui fait des requêtes doit pouvoir choisir diverses unités vidéo ***qui ont été réalisées dans différentes recherches*** par ***différents*** chercheurs.

Il est donc nécessaire que ***la requête porte sur les unités vidéo*** et non sur la recherche qui a permis de constituer les enregistrements vidéo ; la base doit donc pouvoir les dissocier.

Les unités vidéo doivent ***être dissociées de la recherche mais documentées par elle***



# Usage envisagé de ViSA

## Corpus d'une recherche déposée

Cas d'une séquence d'enseignement de 3 mois à raison de 1h+1h30 par semaine

12x3h30 = 42 heures avec 1 caméra



Enseignement : 84 heures

Écrits élèves (photo): 100 à 500 photos

Écrits professeurs

Entretiens (audio ou vidéo)

Données classe

Données établissement

# Les documents associés aux vidéos et à l'ensemble de la recherche

## Données primaires

- copies cahier d'élèves (photo, tiff..),
- notations des élèves (pdf),
- données classe, établissement,
- etc.

## Données secondaires

- Transcription
- Fichier d'analyse de vidéo (xml)
- Article de recherche (pdf)

# Données primaires / secondaires

- les **données primaires** sont constituées par les **enregistrements**, ayant un lien le plus proche possible avec l'événement documenté. Elles comprennent aussi les autres objets recueillis dans le contexte de l'action, comme les documents lus ou écrits durant l'action enregistrée, les objets manipulés, les images consultées, etc. Elles comprennent aussi les traces informatiques laissées par l'activité.
- les **données secondaires** sont constituées par la série de descriptions, transcriptions, annotations qui viennent enrichir les données primaires et qui sont souvent fournies après coup et sur la base des données primaires. Elles comprennent aussi les méta-données, les conventions de transcription, les autorisations des participants, etc. (Baude et al. 2005, p.28)

Baude, O., Blanche-Benveniste, C., Cordereix, P., de Lamberterie, I., Goury, L., Jacobson, M., et al. (2005). Guide des bonnes pratiques pour la constitution, l'exploitation, la conservation et la diffusion des corpus oraux

# Exemple

17/02/03 10h:00	Lundi	1:30	2 <b>Videos</b> + texte : partie1meca2.doc + 2 photos du cahier de O et 2 photos du cahier de M + grille competence (competencesmecal.doc) + Transcriptions	TP1a
21/3/03 14:00	Vendredi	0:30	1 <b>Vidéo</b> + transcription	Entretien DS1 1 élève

**3 vidéos à indexer + documents associés à déposer**

# Exemples de corpus déposés

	Niveau	Disciplines	Nbre séances (nbre vidéo par séance)
Corpus 1	CP (grade 3)	Lecture	10 (2)
	CP	Math	10 (2)
	CE2	Lecture	10 (2)
	CE2	Math	10 (2)
Corpus 2	CM1	Math	8 (2)
	CM1	Science	8 (2)
	6ème	Math	12 (1)
Corpus 3	4ème	Math	7 (2)
	3ème	Math	7 (2)

# Requête d'un utilisateur

	Niveau	Disciplines	Nbre séances (nbre video par séance)
Corpus 1	CP (grade 3)	Lecture	10 (2)
	CP	<b>Math</b>	10 (2)
	CE2	Lecture	10 (2)
	CE2	<b>Math</b>	10 (2)
Corpus 2	CM1	<b>Math</b>	8 (2)
	CM1	Science	8 (2)
	6ème	<b>Math</b>	12 (1)
Corpus 3	4ème	<b>Math</b>	7 (2)
	3ème	<b>Math</b>	7 (2)

# Corpus re-construit (1)

	<b>Niveaux</b>	<b>Disciplines</b>	<b>Nbre séances vidéo choisies</b>
	CP	Math	5 (vidéo camera prof)
	CE2	Math	“
	CM1	Math	“
	6ème	Math	“
	4ème	Math	“
	3ème	Math	“

## Corpus reconstruit (2)

*Corpus pertinent pour l'utilisateur de la base ViSA*

peut être composé d'enregistrements vidéo de plusieurs corpus issus de recherches différentes déposés dans la base



- Les séries d'enregistrements vidéo des corpus initiaux doivent être décomposées
- Chaque enregistrement doit rester informé par l'ensemble de la recherche



# Dépôt d'un corpus dans ViSA

- Quelle procédure?

## Collecte données

Bande mini DV, ..  
(Tournage avec camera)

Documents associés  
primaires (textes, photos, ...)

## Premiers traitements techniques et scientifiques

Numérisation, compression  
Transcription, synopsis, ..

Stockage temporaire (accessibilité privée limitée au(x)  
chercheur(s) avec indexation limitée)  
Utilisation des outils d'analyse disponibles dans ViSA

## Traitements conduisant à d'autres documents secondaires et aux rapports et publications

Indexation finale dans ViSA. Demande de validation  
Signature du contrat de dépôt, mise à disposition des  
autorisation pour les enregistrements vidéo

Stockage pérenne, Archivage CINES  
Accessibilité via l'interrogation de la base ViSA (suivant  
les droits donnés par le chercheur)

### Première possibilité

Autorisation dépôt par le conseil  
scientifique et par le gestionnaire  
du site : login, mot de passe

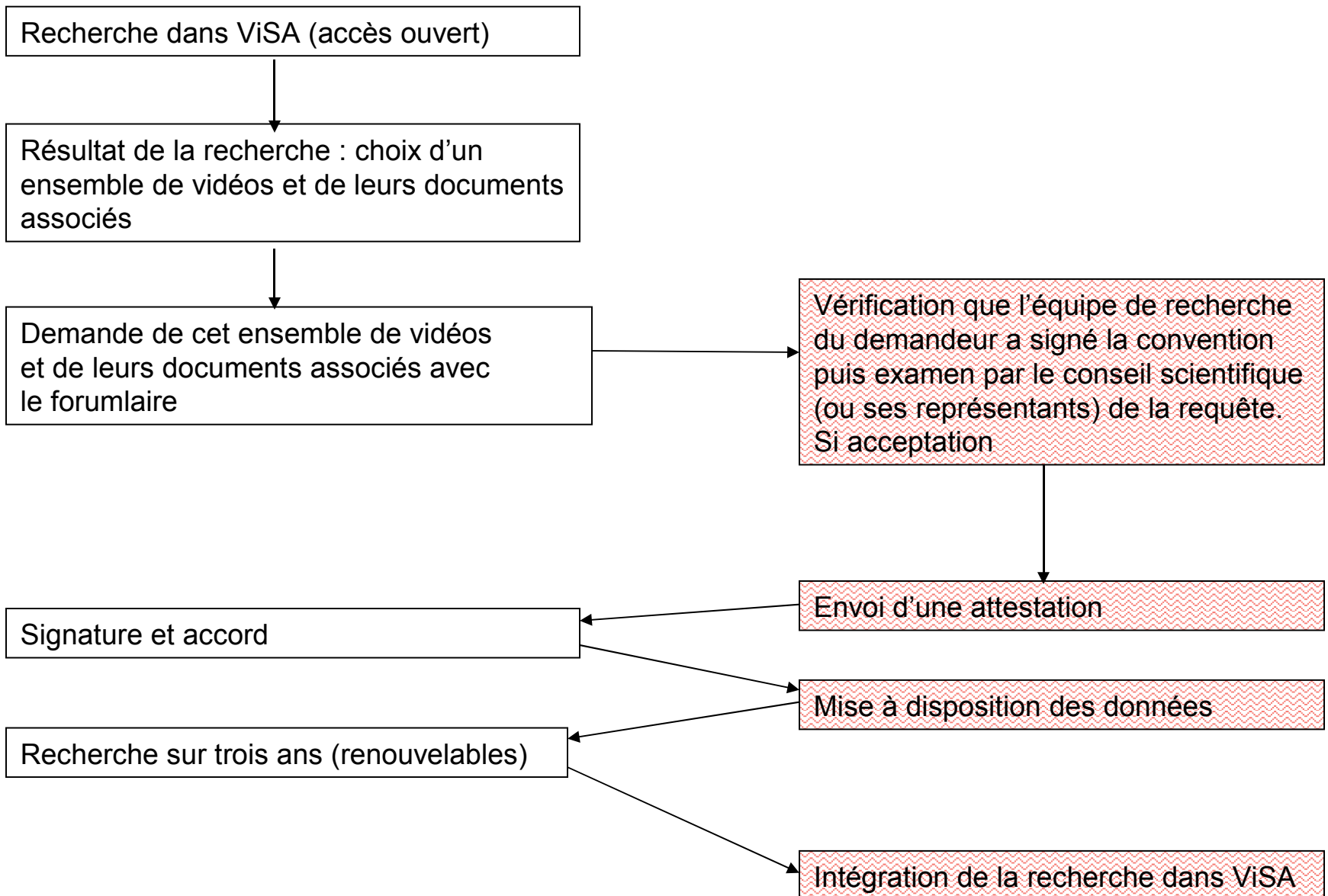
### Deuxième possibilité

Autorisation dépôt par le conseil  
scientifique et par le gestionnaire  
du site : login, mot de passe

Validation par le responsable  
éditorial



# Requête



# Choix de l'unité documentaire

## Unité documentaire

Un enregistrement vidéo **pris en un temps et un lieu donnés.**

Cette unité est aisée à définir dans le cas d'enregistrements pris lors de séances d'enseignement. Dans ce cas, une unité est l'enregistrement de la séance par **une** caméra ; s'il y a plusieurs caméras dans la salle de classe, chacun des enregistrements donne lieu à une unité vidéo.

Cette unité documentaire **est située** dans la recherche qui l'a produite puis dans toutes les recherches ultérieures qui l'utilisent.

# Commentaires

Avec la base de données ViSA

- « La vidéo est une ressource, un composant »  
On décompose un « ensemble » (une recherche avec son corpus) pour l'intégrer dans une nouvelle recherche avec **d'autres auteurs**
- La base de données devient un « **instrument** » partagé par une (ou plusieurs) communautés

# Commentaires (suite)

Le choix de l'unité documentaire converge avec la « réalité **spatio-temporelle** » de la majorité des situations filmées dans l'enseignement (de l'ordre de l'heure)

Cet instrument (ViSA) suppose une stabilisation au moins provisoire de l'indexation que l'on va utiliser (Habert, séminaire 2007).

# L'indexation de documents

# Quelle indexation?

## Deux perspectives différentes

1. donner les informations nécessaires au chercheur sur chaque vidéo pour lui permettre d'en construire une signification suffisante pour la sélectionner (ou non) et pour ensuite en construire une autre signification grâce à une nouvelle recherche (**indexation**) ;
2. permettre de faire partager au sein de la communauté scientifique la (les) analyse(s) d'une vidéo construite par un chercheur.

## Complémentarité de ces perspectives



# Quelques points généraux sur l'indexation

- L'indexation (traditionnelle) est le processus selon lequel le contenu d'un document est analysé pour être ensuite **reformulé** dans une forme permettant d'accéder au contenu et de le manipuler. (Bachimont, 2007 p.196)

Bachimont, B. (2007). *Ingénierie des connaissances et des contenus. Le numérique entre ontologies et documents*. Paris: Lavoisier.

# L'indexation des unités vidéo

« La principale difficulté présentée par l'image ou par tout signe qui montre est que l'image montre sans préciser ce qu'elle montre, ***l'image ne prescrit pas par elle-même une signification***»  
(Bachimont, 2007 p. 202)

# Le cas de la langue

Chaque unité linguistique signifie en fonction des identités et des différences qu'elle entretient avec les autres unités linguistiques,

« dans l'intérieur d'une même langue, tous les mots qui expriment les idées voisines *se limitent réciproquement*: des synonymes comme redouter, craindre, avoir peur n'ont de valeur propre que par leur opposition, si redouter n'existait pas, tout son contenu irait à ses concurrents. » (Saussure, 1985)

(Bachimont, 2007; p.203)

# Comparaison image - langue

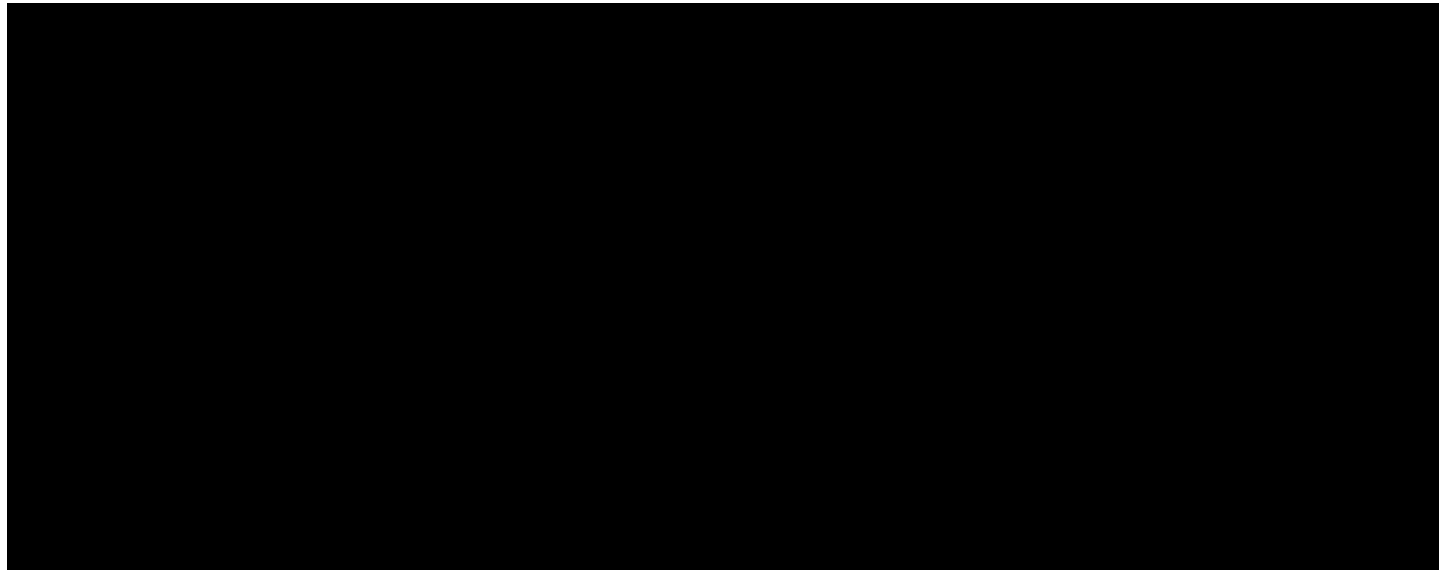
Le concept [mot] prescrit une signification car il appartient à un système d'oppositions et de différences qui lui donne hors contexte un contenu.

La situation est très différente des images et des sons. *Il n'existe pas de système fonctionnel pour les images et les sons, c'est-à-dire de système permettant de prescrire a priori une signification à une image ou à l'un de ses segments. C'est pourquoi il faut soigneusement distinguer les descripteurs d'une image ou d'un son des index permettant de la qualifier.*



- Un descripteur est une information *extraite d'un document ou d'un contenu* par des moyens automatiques. L'information associée à un descripteur est donc proche du contenu physique du document;
- Un index est une *position ou une valeur donnée* dans un système d'interprétation associé à *l'exploitation d'un ensemble de documents*, c'est-à-dire un fonds documentaire.

(Bachimont, 2007; p.203)



« Les documents audiovisuels *ne peuvent signifier par eux-mêmes*, ils sont plongés dès leur conception dans un univers textuel qui construit le sens dont ils seront la manifestation audiovisuel. [...]. Finalement, l'archivage documente les contenus pour *leur associer une signification* permettant de les retrouver et de les utiliser ».Bachimont, 2007 p. 204



# Différence Descripteur/Index

Distinction entre :

- information extraite d'un document (*descripteur*)
- position ou valeur donnée dans un *système d'interprétation* associé à l'exploitation d'un ensemble de documents (*Index*)

« Toute indexation doit donc être respectueuse de l'objet, mais *créative* dans la description. Faisant fi d'une objectivité qui serait donnée, elle doit la rechercher à *travers la multiplicité des usages* qu'elle rend possible et qui lui renvoie en miroir la richesse interprétative du contenu indexé. » (Bachimont, 2007; p. 236)

# L'indexation des unités vidéo dans ViSA

## Le module indexation

- a été conçu par un groupe de travail avec comme « expert documentaliste » Cyril Mathevon (INRP)
- a été développé par Olivier Morel sous la responsabilité de Yvonnick Fesselier (INRP).



# Accès à ViSA pour indexer

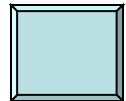
## Trois étapes

1. Inscrire l'équipe de recherche (et faire signer la convention)

Login et Mot de passe pour **TOUS les chercheurs de l'équipe**

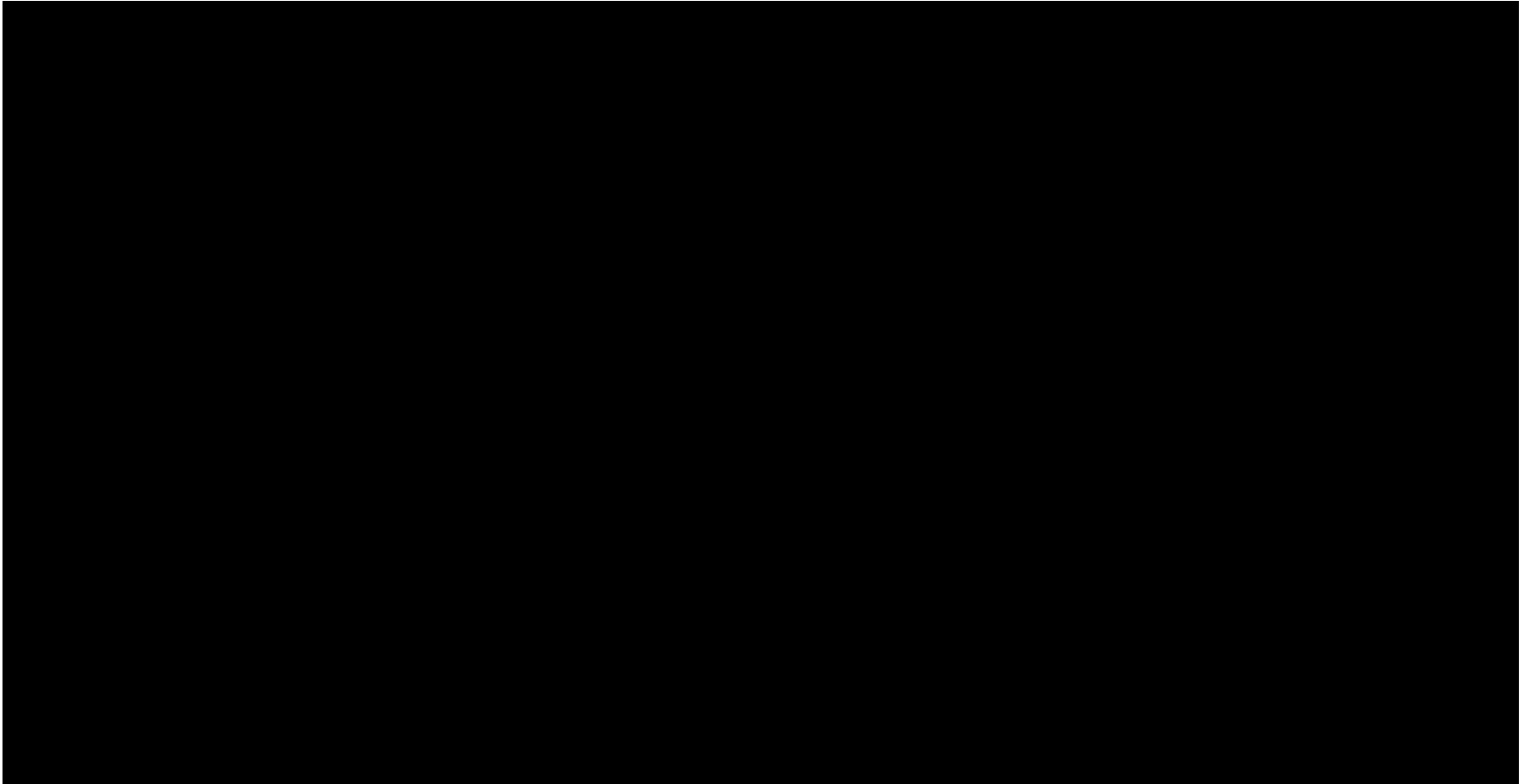
1. Inscrire et présenter la recherche (demande auprès du Conseil scientifique) avec les noms des chercheurs

Login et Mot de passe **INDIVIDUEL**



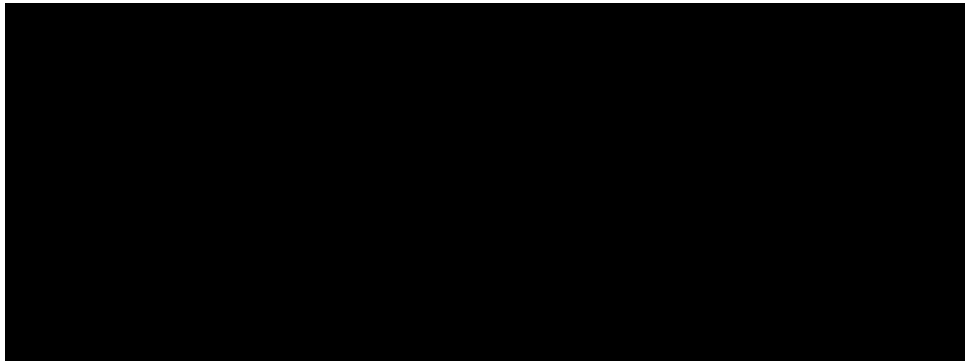
1. Indexer les unités vidéo

Inscrire une recherche  
(avec le mot de passe de l'équipe)

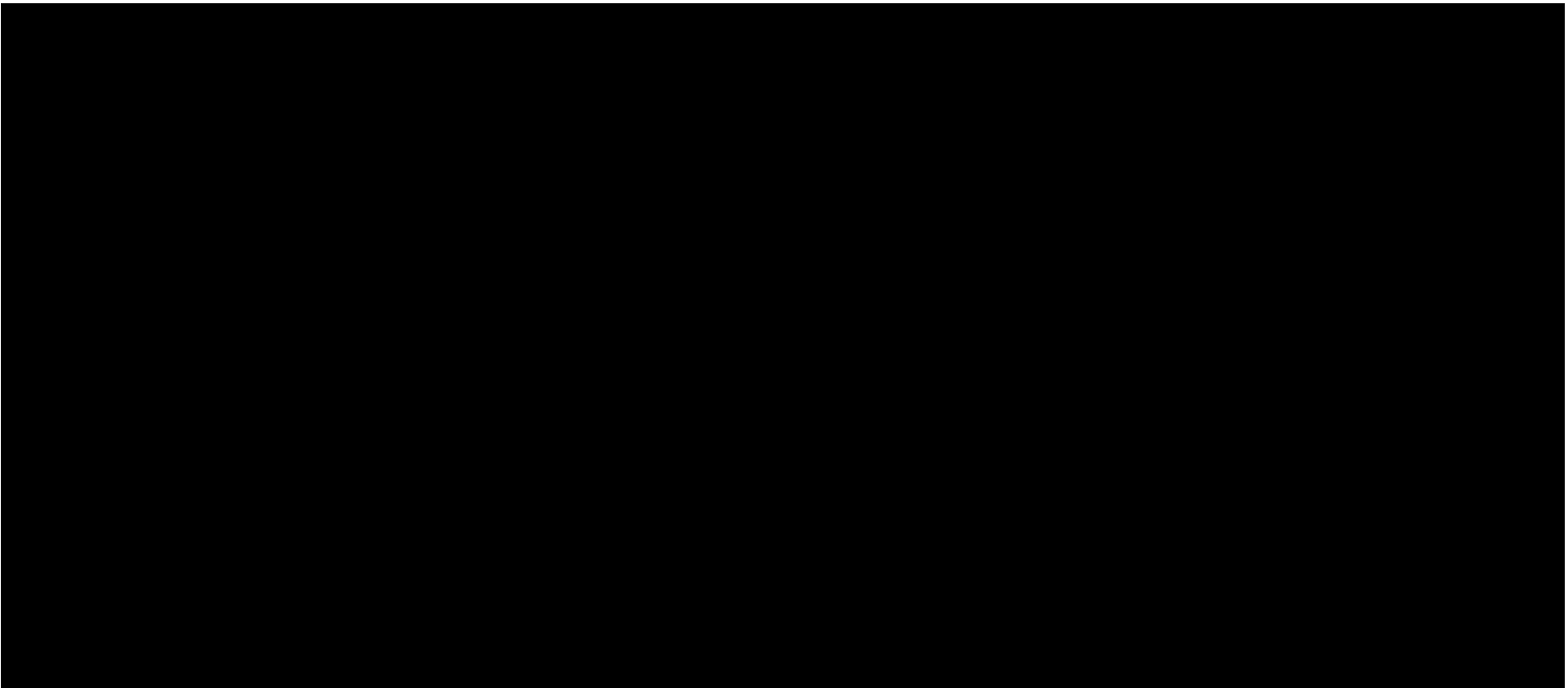


# Inscription Recherche

## Suite écran

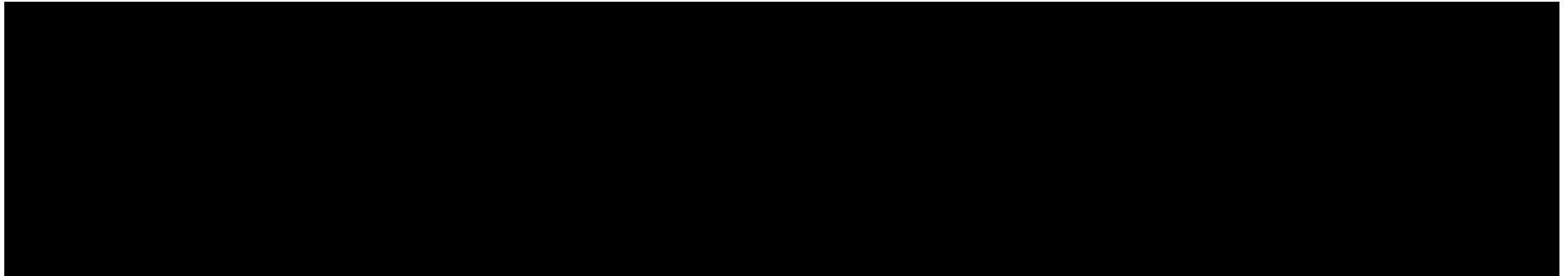


# Liens recherche-unités vidéo



# La structure de l'indexation d'une unité vidéo

## La page écran

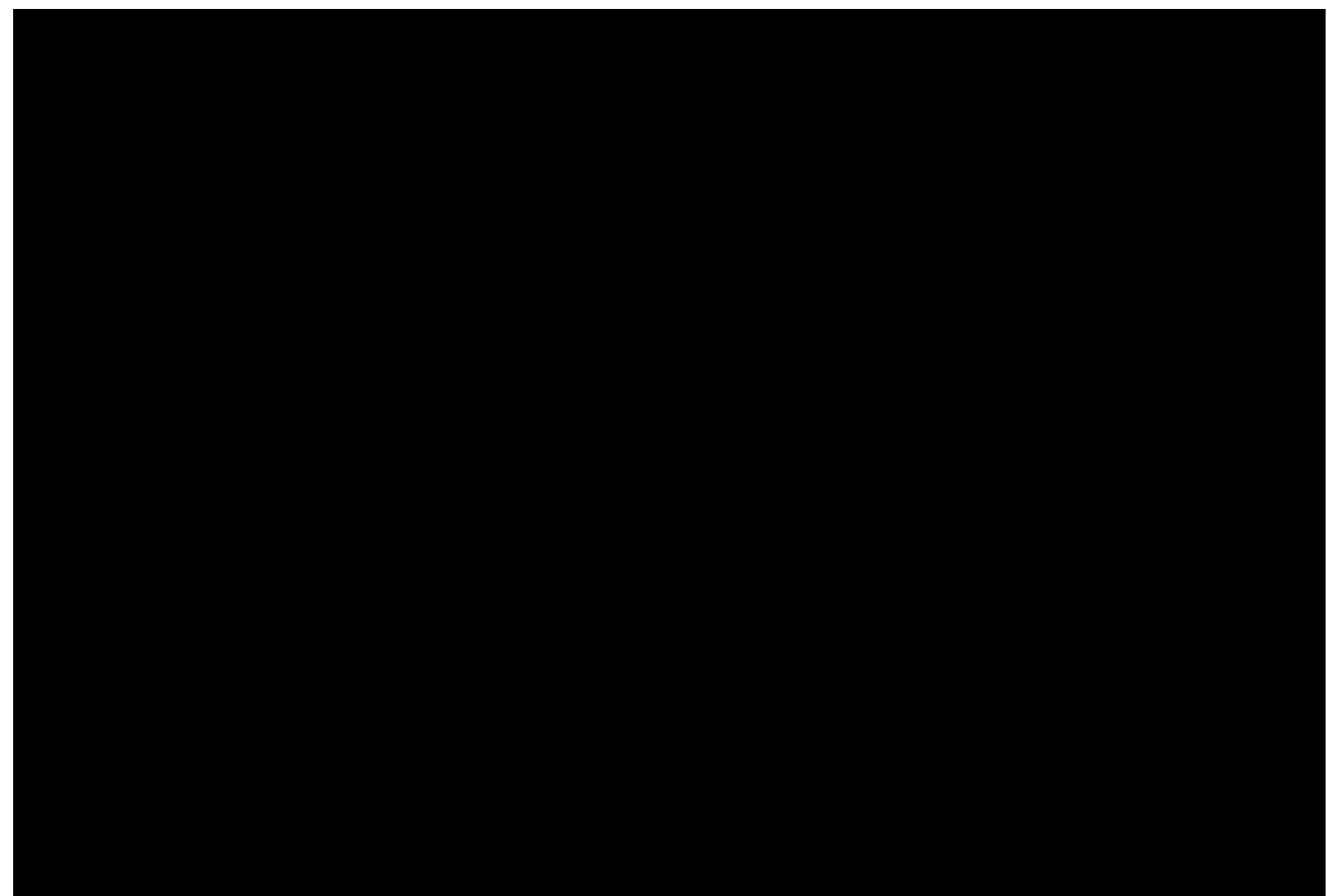


Basée sur Dublin Core (C. Mathevon, INRP)

<b>Descriptif de la vidéo</b>	<b>Conditions de tournage</b>	<b>Situations d'enseignement</b>	<b>Format</b>	<b>Cadre juridique</b>
Titre de l'unité vidéo Auteurs / contributeurs Mots clés Langue parlée Description	Date (jour/heures) Matériel de prise de vue et de son Prise de vue et montage Personnes filmées Situation filmées Lieu de tournage	Classe (Nbre d'élèves, filles garçons, etc.) Niveau Discipline Programme officiel Etc.	Durée Format du fichier vidéo Qualité de la vidéo Localisation du document primaire	Condition d'usage de la vidéo, etc.

# Descriptif de l'unité vidéo

<b>Identifiant</b>	
<b>Titre</b>	
<b>Titre alternatif</b>	
<b>Fonds</b>	
<b>Créateur</b>	<b><i>Auteur</i></b> (personne(s) physique(s) qui ont la triple responsabilité scientifique, morale, juridique de l'enregistrement)
<b>Contributeur</b>	
<b>Sujet</b>	<b><i>Description</i></b>
<b>Sujet en anglais</b>	





# Conditions de tournage (suite)



L'indexation est essentielle si l'on veut que les requêtes puissent être efficaces pour les chercheurs.