

# La plateforme OKAPI

A. Beloued & S. Lalande



Projet Campus AAR ANR-13-CORD-0016-01



## Cadre des Humanités Numériques

Problématique de constitution et d'exploitation des patrimoines audiovisuels scientifiques pour la recherche, l'enseignement et la culture

## Réaliser une plateforme ouverte de description et de (re-)publication d'archives audiovisuelles

### Partager, capitaliser et valoriser des connaissances autour de corpus audiovisuels

Découper, annoter, analyser, rapprocher des contenus  
Valoriser ce travail par différents produits de médiation

### Adaptable à des contextes et usages variés

Communautés savantes (SHS, sémiotiques, etc.)  
Domaine culturel (valorisation à faible coût de fonds d'archives)

### Modèle-agnostique

Possibilités de définir et gérer plusieurs modèles de description et de publication

## Outiller la réalisation d'une base de descriptions/connaissances ouverte

- Produire des graphes sémantiques

## Utilisation des langages du Web sémantique pour la représentation des graphes et la réalisation des modèles

- Réutilisation de ressources de description existantes: ontologie RDFS/OWL et thésaurus SKOS
- CIDOC-CRM, thésaurus UNESCO, RAMEAU, THeSOz ...
- Utiliser les mécanismes d'extension de ces langages

## Utilisation d'un moteur d'inférence pour le processus de publication

- Publier = transformation de graphes

## Utilisation de bases de données type « triples store »

- Stockage de graphes
- Capacités d'inférences

# Architecture OKAPI (Open Knowledge-based Annotation and Publishing Interface)

The screenshot displays the OKAPI software interface. At the top, there is a menu bar with options: Fichier, Analyse, Individu, Recherche, Publication, Fenêtres, Aide. Below the menu bar is a 'Quick Access' section with icons for 'Analyse', 'Segmentation', and 'Fermer le Studio'. The main workspace is divided into several panes. On the left, a tree view shows a hierarchy: 'De l'universel au dialogue des cultures' > 'propos' > 'Sujet "Personnalités de référence"'. The central pane, titled 'Sujet "Personnalités de référence"', contains a form with several fields: 'intitulé : Sujet "Personnalités de référence"', '1) Description de la topique "Personnalité" : Kant, Emmanuel', 'Localisation du sujet dans un/plusieurs pays du monde : Allemagne', and 'Localisation du sujet dans une/plusieurs régions du monde :'. On the right, a panel titled 'Ajouter un Champ au Formulaire d'Analyse' lists various fields like 'beginAt', 'début de période', 'endAt', 'fin de période', 'latitude', 'lien wikipédia', 'locution', 'longitude', 'nom court', 'nom public', 'relation sémantique', 'comment ?', 'quoi ?', '2) Description de la topique "Tr"', and 'quand ?'. At the bottom, a 'Ligne de temps' (Timeline) pane shows a horizontal axis with a green bar labeled '0:00:00' and a list of thematic strategies: 'Strate thématique' (De l'universel au dialogue des cultures), 'Strate sonore', 'Strate visuelle', 'Strate rhétorique', and 'Strate thématique'. The timeline also shows a sequence of items: 'Introduc..', 'Présentation ..', and '"De l'universel au dialogue des cultures" par François Jullien'.

RDF  
Triples store



Cliopatria



Descriptions  
Publications  
Entités nommées  
Modèles & Ontol.

Sites  
Ontologie Sites

## Comment outiller de manière générique des besoins très différents ?

- Description liminaire genre Web: champs texte, tagging concept thésaurus
- Description détaillée par documentaliste: multiple champs, référentiels
- Analyse sémantique complexe : sémiotique (différents axes), analyse de films

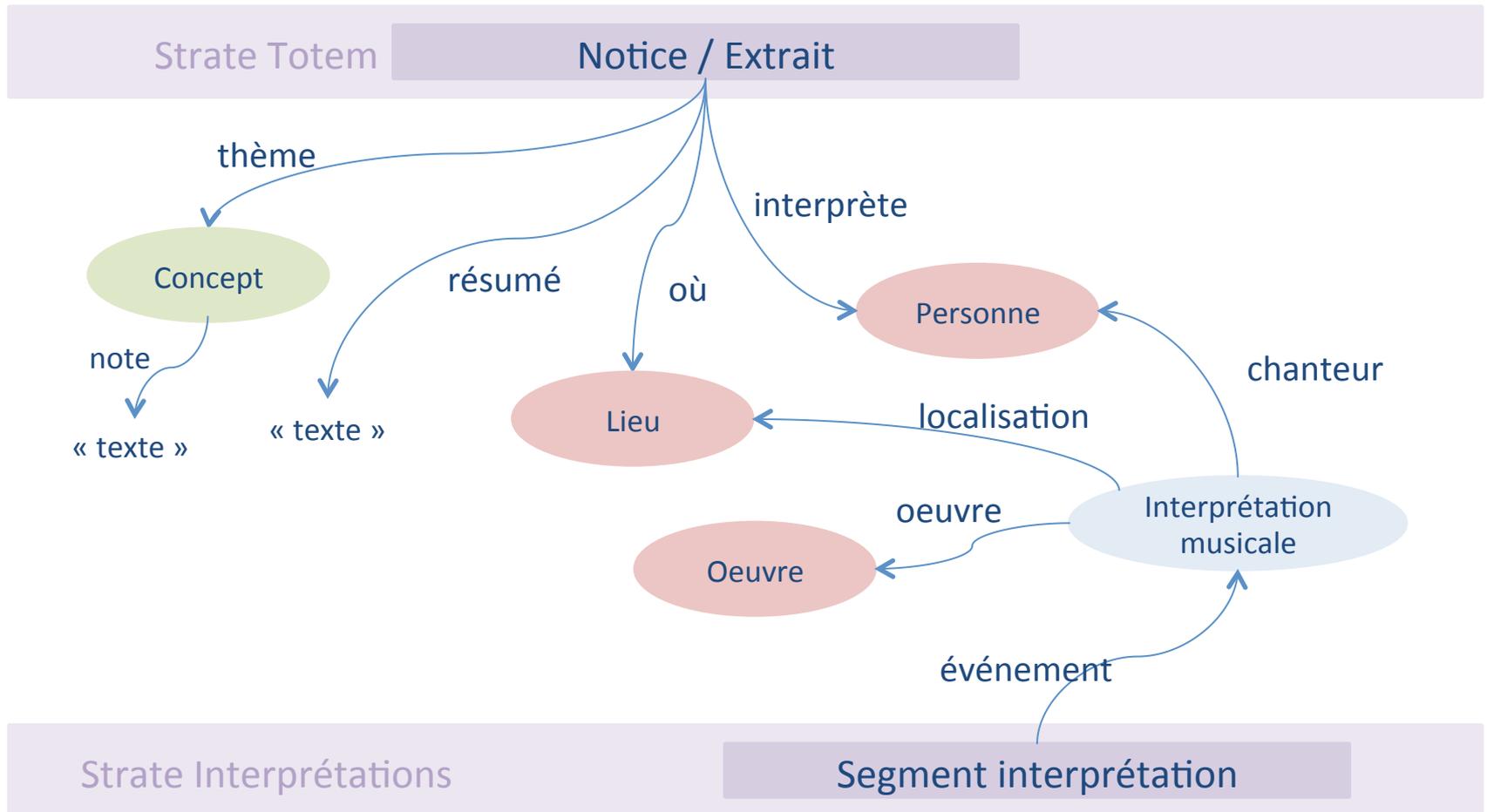
## Un modèle simple à 5 objets structurants:

- Analyse
- Strate
- Segment (ancre temporelle)
- Corpus
- Entité (racine des connaissances du domaine)

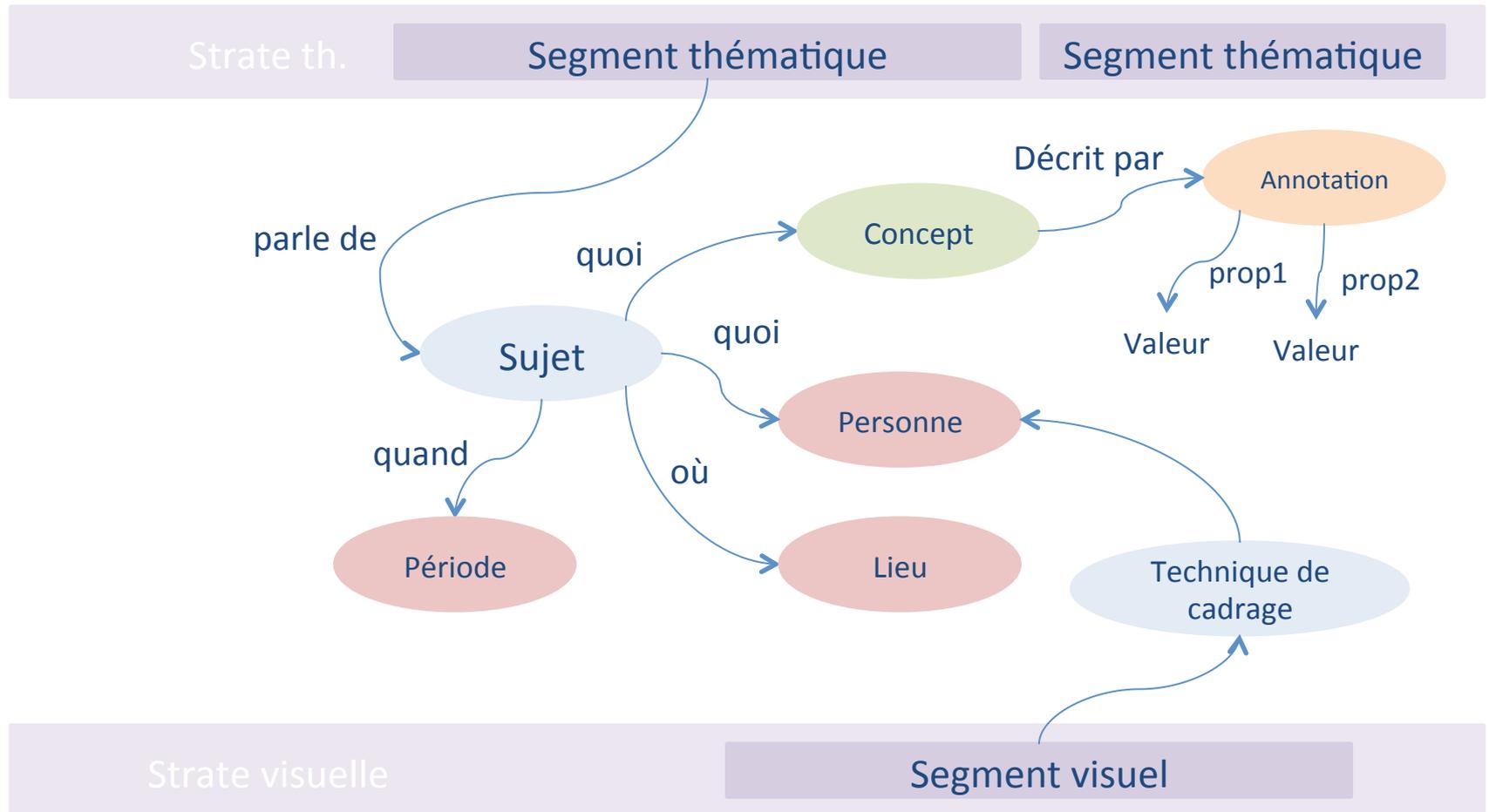
## Tout le reste est extension

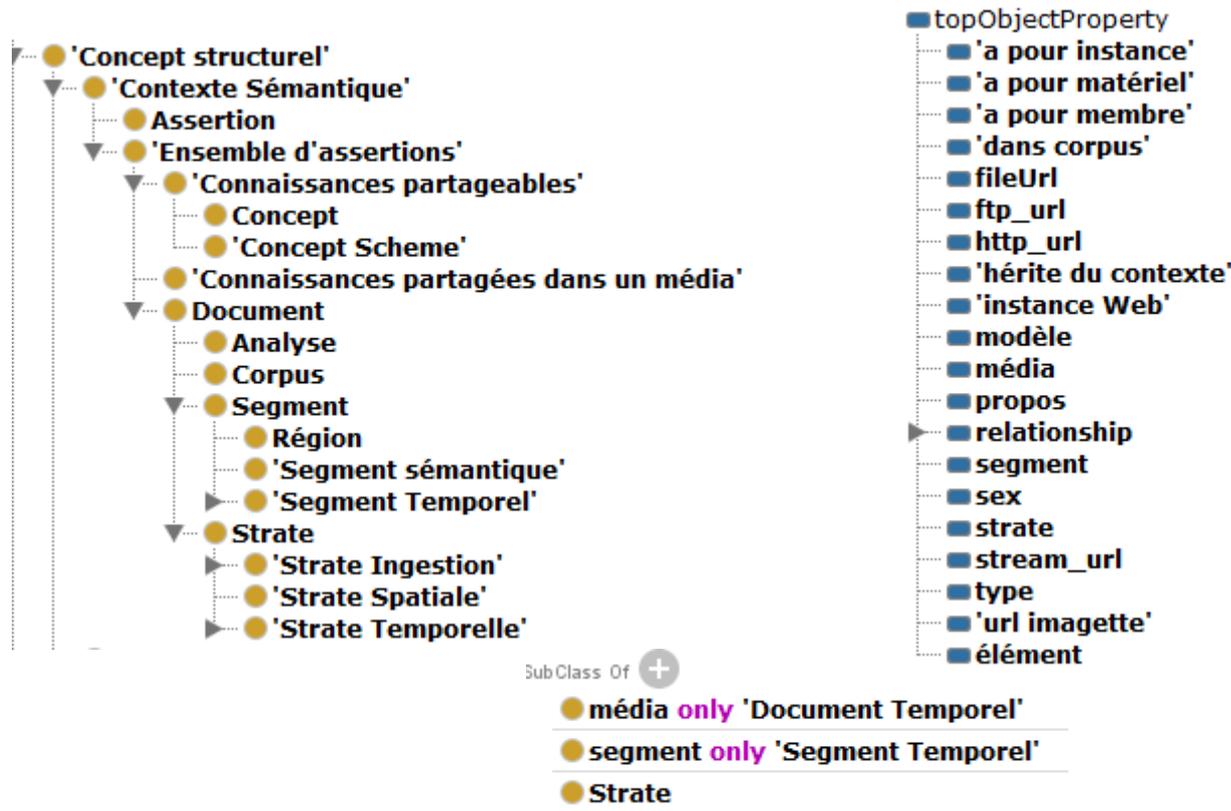
- définir ses propres strates, ses propres segments, son vocabulaire du domaine  
= modifier/étendre une ontologie OWL

# Modèle INA Totem ++ (Campus AAR)

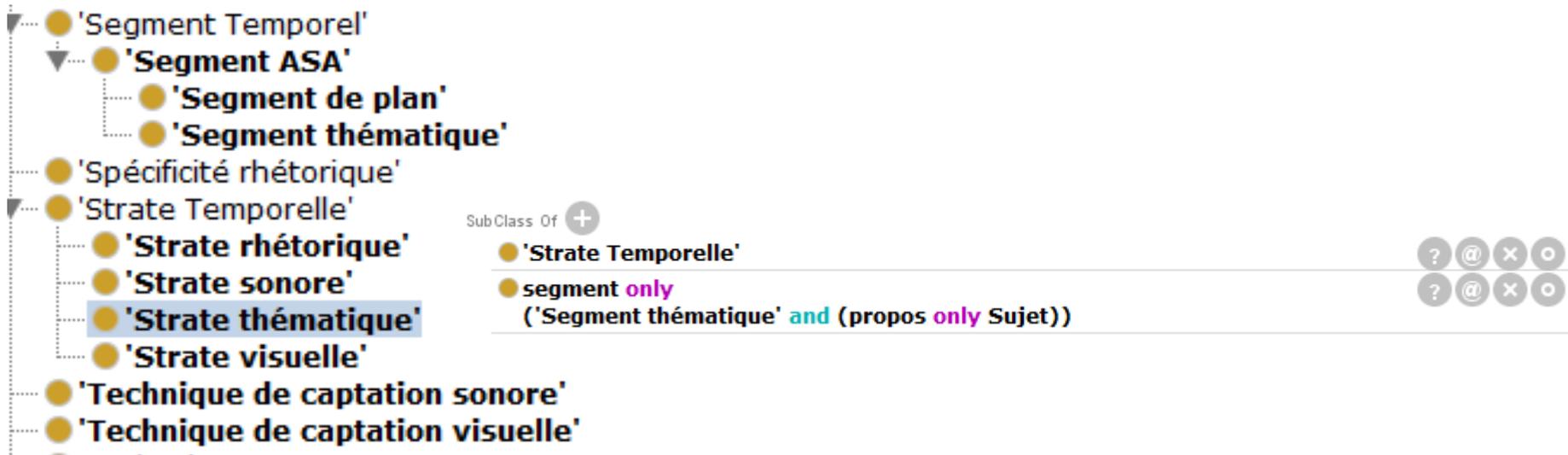


# Modèle ASA-SHS (version Campus AAR)

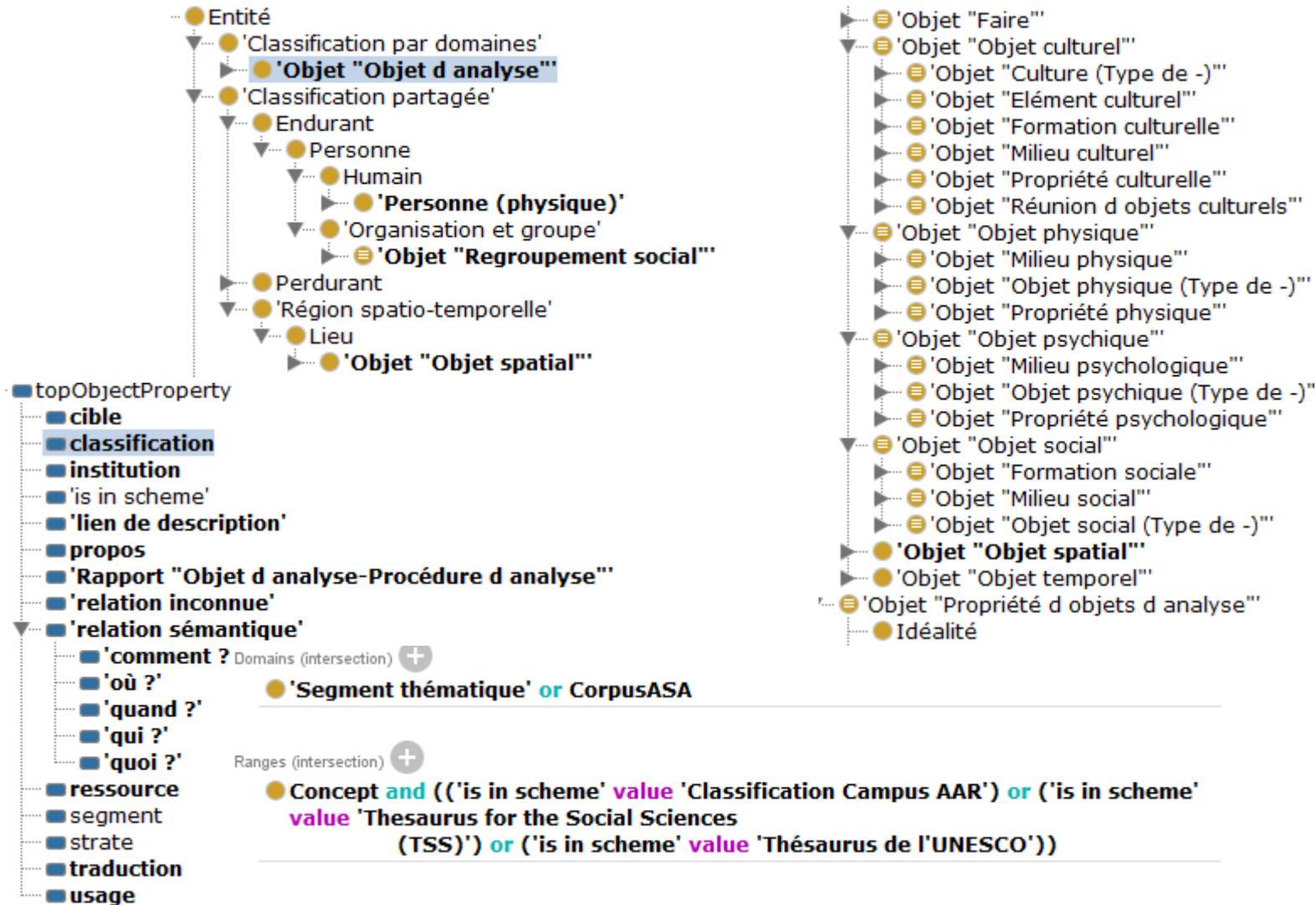




# Modèle ASA (structure)



# Modèle ASA (Domaine)



Ontologie Individus Annotations Attributs Relations Classes

Formulaire d'analyse

cherchez une classe

- Concept structurel
  - Contexte Sémantique
    - Assertion
      - Ensemble d'assertions
        - Connaissances partageables
          - Connaissances partagées dans un média
        - Document
          - Analyse
            - Analyse ASA
              - Analyse générale
                - Analyse de base d'un corpus d'objets média
                - Analyse de base de l'objet média
                - Analyse des constituants de l'objet média
              - Analyse spécialisée
                - Corpus
                - Segment
                - Strate
            - Instance de Media
              - Matériel
              - Média
    - Contributeur à la constitution de l'archive
      - Définir l'origine du champ des mots-clés
      - Définir le type du champ des mots-clés
    - DESCRIPTION PAR STANDARD
    - Droits et copyright relatifs à l'analyse
    - Droits et copyright relatifs à l'objet média
    - Entité
    - Expression
    - Facette sémantique

Campus AAR - Thesaurus ASA pour l'analyse d'une ressource audiovisuelle  
skos  
Modèle coeur Campus-AAR  
CAMPUS AAR - Activité d'indexation ASA (Atelier de Sémiotique Audiovisuelle)  
TYPES DE BASE  
CAMPUS AAR - Thésaurus ASA ( Atelier de Sémiotique Audiovisuelle)  
language  
CAMPUS AAR - Terme conceptuel du méta-lexique ASA (Atelier de Sémiotique Audiovisuelle)  
owlModel  
CAMPUS AAR - Bibliothèque ASA des schémas de description  
Campus AAR - Bibliothèque des schémas de définition du discours ARC  
rdfsModel  
Ontologie coeur ASA

dépend de :

- Campus AAR - Bibliothèque des sujets du discours ARC (Archives Rencontre des Cultures)
- Ontologie coeur ASA
- Campus AAR - Bibliothèque des schémas de définition du discours ARC
- Campus AAR - Bibliothèque des séquences d'analyse du discours ARC

Chercher un terme

intitulé :

- ARC

Saisir un texte libre

type :

- Ontologie
- Connaissances partageables

Chercher un terme

Ontologie : ARC

nouveau champ

- compatible avec
- incompatible avec
- individus différents
- individus similaires
  - individus similaires
  - intitulé alternatif
  - intitulé caché
  - intitulé préféré
- member de la liste notation
- note
  - scope note
  - definition
  - change note
  - editorial note
  - example
  - history note
  - valeur de la liste
  - version actuelle
  - version antérieure
- voir aussi
  - est défini dans

Ontologie Individus Annotations Attributs Relations Classes

↑ 0-3 | cherchez une classe

- > Analyste de l'objet média
  - a pour instance
  - a pour matériel
  - a pour membre
  - cible
  - classification
  - contre exemple
- > Contributeur à la constitution de l'archive
  - dans corpus
  - Définir l'origine du champ des mots-clés
  - Définir le type du champ des mots-clés
  - élément
  - fileUrl
  - ftp\_url
  - has member
  - has member list
  - has top concept
  - hérite du contexte
  - http\_url
  - Identifier la catégorie linguistique
  - Identifier le navs

Formulaire d'analyse

Re

x P > - cible :

UN(E) Skos Concept

Chercher un terme

x - intitulé :

classification

Saisir un texte libre

x P > - source :

UN(E) Segment thématique OU CorpusASA

Chercher un terme

x P > - type :

Relation

Chercher un terme



Lecteur vidéo

Player



Visionneuse d'images

Current time



0:04:05:7

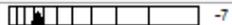


0:04:05:8



0:04:05:9

Ligne de temps



0:02:59:5



+ X	MediaLayer	
+ X	Strate thématique	Vie quotidienne à Hong-Kong Paysag., Ba., Ens., Mo., T., Lieux de c., Hugo Man-To Leung, Un, Kin's Kitchen .., Marché aux po., Poissons., L
+ X	Strate sonore	
+ X	Strate visuelle	
+ X	Strate rhétorique	

## Problématique

- Réutilisabilité de corpus de segments
- Mise en relation rhétorique potentielle de ces segments en contexte
- Valorisation sans éditorialisation

## Elaboration Requête sémantique complexe

- « Hyper-Formulaire »

[recherche.html](#)

## Offrir des modèles simples mais paramétrables

- Corpus (mini portail)
- Vidéo-lexique
- Parcours narratifs
- ...

## Permettre la réalisation de modèles

- Par extension des objets existants
- Par composition des objets existants
- et surtout... par paramétrage des comportements
- Le tout avec des fonctionnalités OWL/SPARQL ...

## Un ensemble d'objets hypermédias

- Pages, menus,
- ancres, zones de page
- textes, images,
- lecteur vidéo,
- objet carte,
- objet frise temporelle

Un graphe de ces éléments représente le site déployé

Transformation par xslt ->

HTML5 + angular JS

Amalia.JS

vis.JS



# Interface de paramétrage



composant graphique :

- Page principale
- Menu principal**
- Menu gauche
- Paragraphe
- Liens hypermédia
- Glossaire
- Nuage de mots-clés
- Carte Géographique
- Texte affiché sur la carte
- Timeline
- Texte affiché sur la Timeline
- Hiérarchie d'accès

éléments de l'ontologie :

- Propriété
- Classe**

Chercher un élément de l'ontologie

Classe	Intitulé affiché sur le site	Ordre d'affichage	Total
<ul style="list-style-type: none"> <li><input type="checkbox"/> Analyse et adaptation de l'objet média                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Traduction</li> </ul> </li> <li><input type="checkbox"/> Comportement basique                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Corrélation</li> <li><input type="checkbox"/> Disposition spatiale</li> <li><input type="checkbox"/> Ordre temporel</li> </ul> </li> <li><input type="checkbox"/> Concept structurel                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Contexte Sémantique                                     <ul style="list-style-type: none"> <li><input type="checkbox"/> Ensemble d'assertions   <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Connaissances partageables <b>Glossaire</b></li> <li><input type="checkbox"/> Document</li> </ul> </li> </ul> </li> <li><input type="checkbox"/> Conteneur thématique                                     <ul style="list-style-type: none"> <li><input type="checkbox"/> Texte éditorial</li> <li><input type="checkbox"/> Topique   <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Article</li> <li><input checked="" type="checkbox"/> Rubrique</li> </ul> </li> </ul> </li> </ul> </li> <li><input type="checkbox"/> Contributeur à la constitution de l'archive                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Auteur de l'objet média</li> </ul> </li> <li><input type="checkbox"/> Document audiovisuel                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Extrait vidéo</li> <li><input type="checkbox"/> Image</li> </ul> </li> <li><input type="checkbox"/> Entité                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Classification par domaines</li> <li><input type="checkbox"/> Classification partagée</li> <li><input type="checkbox"/> Sujet</li> </ul> </li> <li><input type="checkbox"/> Expression                             <ul style="list-style-type: none"> <li><input type="checkbox"/> Expression de la branche "Domaine d'analyse ASA"</li> <li><input type="checkbox"/> Expression pour la branche "Procédure ASA d'analyse"</li> </ul> </li> </ul>			
			1
			1
			40
			20
			14
			6
			68
			68
			68
			28
			40
			23
			14
			9
		1	5
		0	4
			11
			8
			34
			16
			18
			105
			98
			44
			7
		2	16
			14
			2



Rubrique Article Glossaire ▾ Sujets

Article

Contes quechuas de l'Ausangate, Pérou

Glossaire

La langue Quechua

L'utilisation de la littérature orale comme ressource pédagogique

Les contes en langue quechua

Apprentissage et mise en pratique de la tradition orale

Les chants waynos



TERMES DU GLOSSAIRE

Amérique du Sud

Anthropologie

Langues et cultures des Amériques

Quechua (langue)

# Composant Map

aPourTheme    générique ▼    a pour date de création    événement    a pour date de dernière modification

Personne physique ▼    noticeisolee

## aPourLieu



## date de diffusion



©Copyright

Les poupées folkloriques

December 11, 1975

### Les poupées folkloriques

Julie vient de recevoir par la poste un colis rempli de poupées de collection qui porte des costumes folkloriques. - Julie et Casimir demandent aux enfants à quelle région appartient chacune des poupées. - Julie chante "La chanson de Julie" Une scénette en papier découpé met en animation un canard et un castor.



DECEMBER 5, 1975  
Simons tel quel, première partie



DECEMBER 12, 1976  
Récital Dionne Warwick



# Publication à partir de données enrichies

chanson auteur ▼ a pour interprétation interprete ▼ oeuvre créée ▼ musique oeuvre person  
physicalperson place sketch work

## L ' Amérique

**paroles**

Mes amis, je dois m'en aller  
Je n'ai plus qu'à jeter mes clés  
Car elle m'attend depuis que je suis né  
L'Amérique

J'abandonne sur mon chemin  
Tant d[...]

[http://lyrics.wikia.com/Joe\\_Dassin:L%27Am%C3%A9rique](http://lyrics.wikia.com/Joe_Dassin:L%27Am%C3%A9rique)

**auteur**

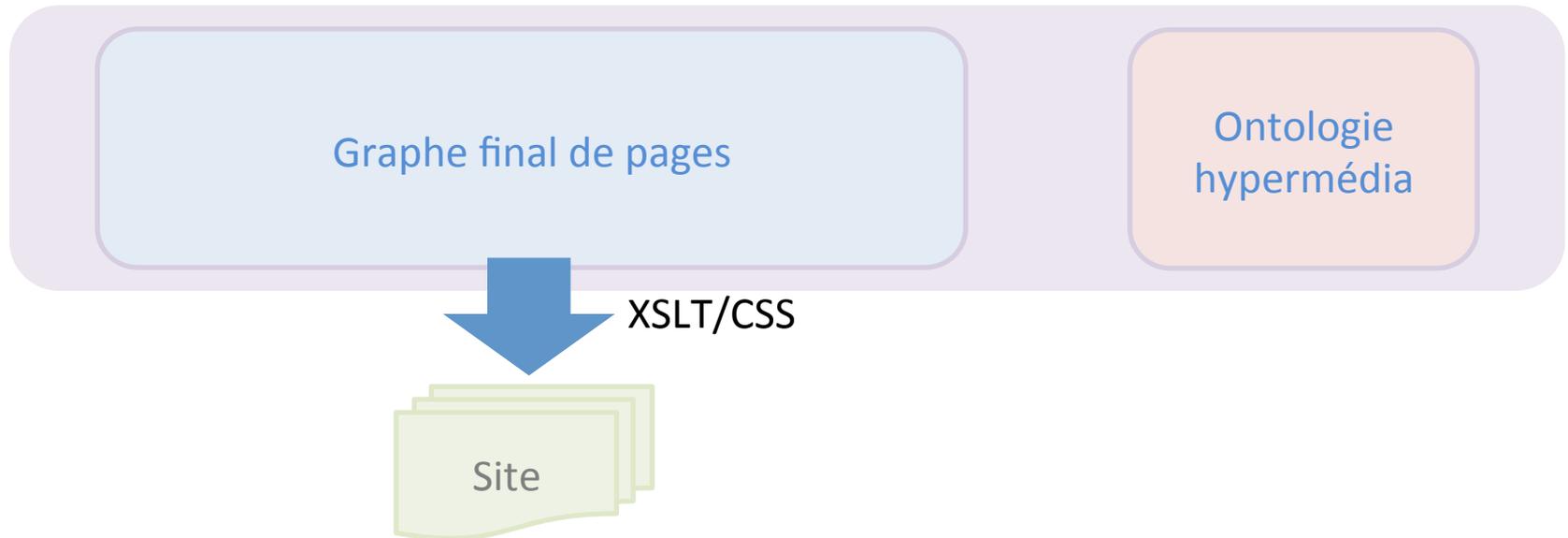
▼ compositeur  
Jeff Christie

▼ paroller  
Jeff Christie  
Pierre Delanoë

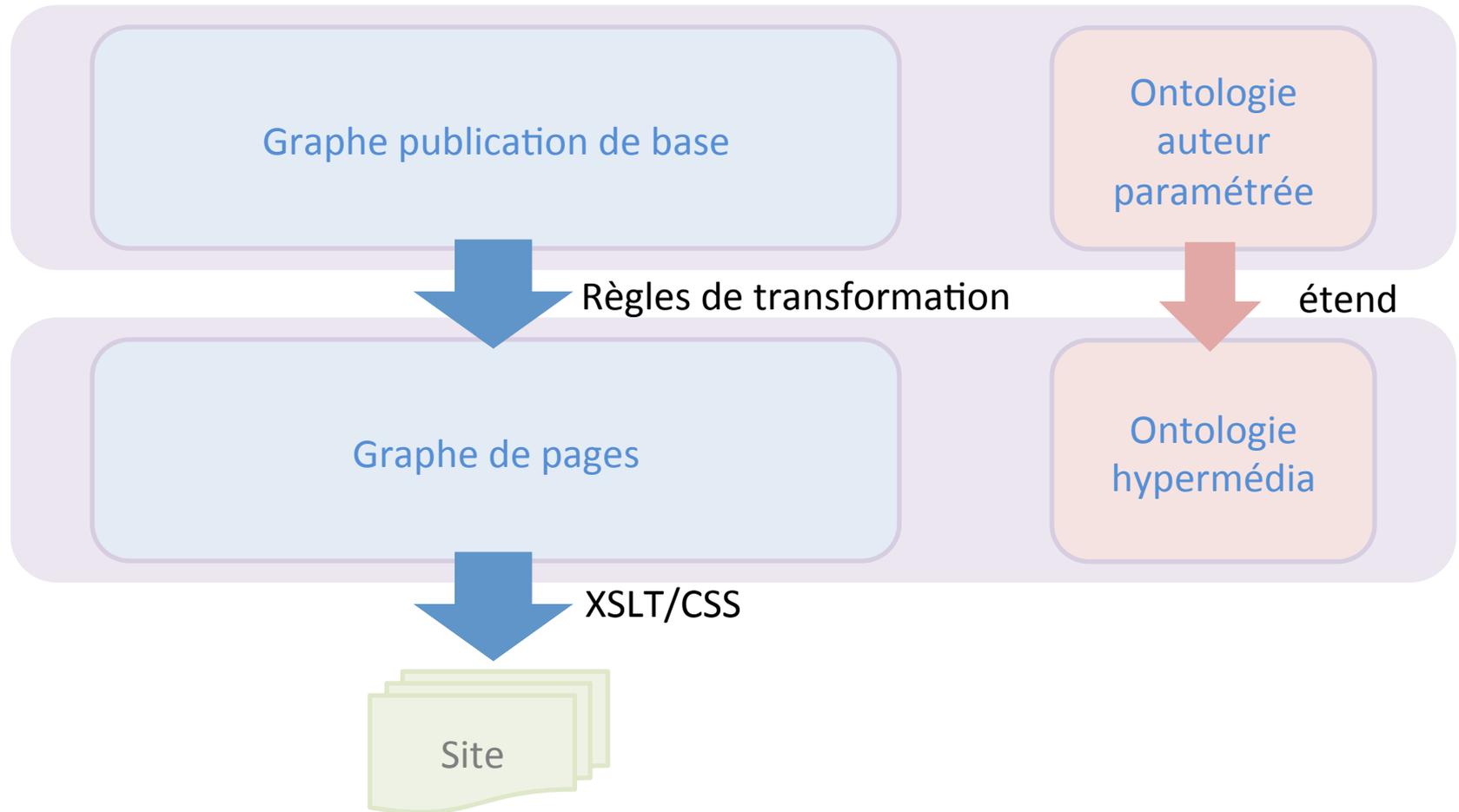
**a pour interprétation**

▼ Interprétation  
L'Amérique  
L'Amérique

# Processus de publication



# Processus de publication



## Une ontologie auteur extensible

- Étend l'ontologie de description : segment, corpus, média
- Objet « topique » : regroupement sémantique
- Un ensemble de relations entre objets
  - Positionnement relatif des objets, spatial, temporel
  - corrélation,
  - relations rhétoriques

## Un ensemble de règles de transformations paramétrables

- Calcul des pages, des affectations d'objets dans les zones, des accès...
- Prise en compte du paramétrage pour la ventilation des métadonnées de description
- Paramétrage par les ontologies/le graphe

## Possibilités de créer des règles spécifiques à un modèle

- Public très averti ;-)

## Modélisateur

- Sous-classement des éléments structurants notamment classe « topic »
- Mise en relation des nouvelles classes d'éléments : contraintes d'association

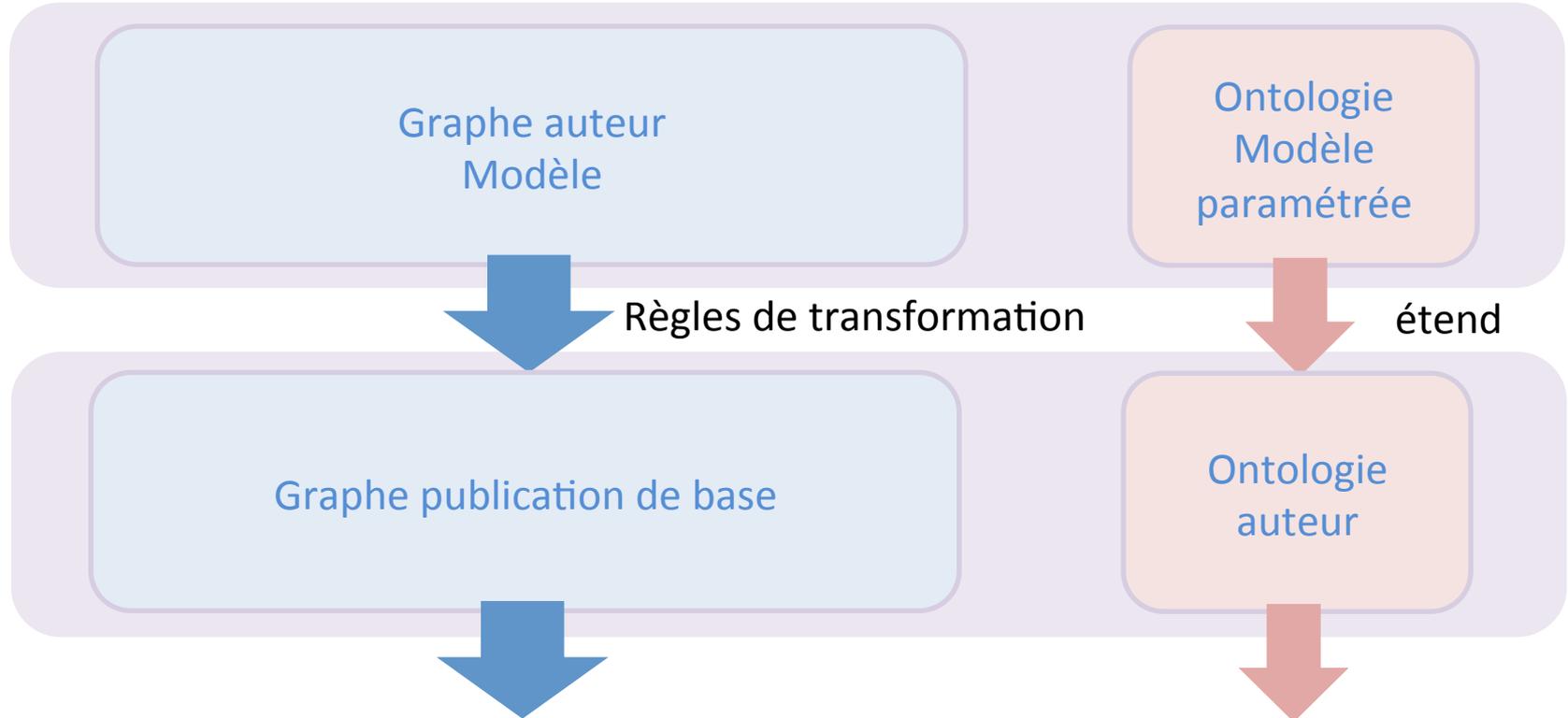
## **Paramétrage de l'ontologie de base dans le contexte du modèle**

- Annotation de relations rhétoriques (corrélation)
- Ventilation des métadonnées dans widgets logiques

## Auteur

- Construction graphe de publication
- Ventilation des métadonnées dans widgets logiques

# Processus de publication



# Vers Okapi Web...

The logo for 'ina' is displayed in white lowercase letters on a blue square background.

Timeline interface showing segments for various analysis layers:

- Strate sonore
- Strate rhétorique
- Strate thématique
- Strate thématique
- Strate visuelle
- Strate visuelle

Timeline markers: 0 m, 05:00 m, 10:00 m, 15:00 m, 20:00 m, 25:00 m, 30:00 m, 35:00 m, 40:00 m, 45:00 m, 50:00 m, 55:00 m, 01:00 h, 01:05 h, 01:10 h, 01:15 h, 01:20 h, 01:25 h, 01:30 h

Control buttons: back, zoom in (+), zoom out (-), zoom reset (i), forward, list, search, close.

**Tranches de vie à Hong-Kong**

Navigation icons: home, up, settings, play.

**SOMMAIRE**

- Tranches de vie à Hong-Kong
  - intitulé
  - Régime général "Creative Commons" relatifs à l'analyse
  - Droit d'auteur relatif à l'analyse
  - description
  - comment citer

Video player interface showing a title card:

Dans la collection "Road Movies"  
la Fondation  
Maison des Sciences de l'Homme  
et  
l'Equipe Sémiotique Cognitive  
Nouveaux Médias  
présentent

Logos: fmsH, escom

Navigation tabs: PROPRIÉTÉS, OBJETS EXISTANTS, NOUVEAUX OBJETS

Content: Propriété de classes

Vie quotidienne à Hong-Kong

▶ intitulé

▶ type

▼ date de réalisation

2007-04-09T15:13:10

▼ Droit d'auteur relatif à la production du document source

© ESCoM (Equipe Sémiotique Cognitive et Nouveaux Médias), FMSH (Fondation Maison des Sciences de l'Homme), France, 2007

▼ Droit d'auteur relatif au contenu du document source

© ESCoM (Equipe Sémiotique Cognitive et Nouveaux Médias), FMSH (Fondation Maison des Sciences de l'Homme), France, 2007

▼ Droit d'auteur relatif à la réalisation du document source

© BONNEMAZOU, Camille, ESCoM-FMSH, Hong-Kong, 2007  
 © STOCKINGER, Peter, Directeur de l'ESCoM, FMSH, Hong-Kong, 2007

▼ description

Le but principal de ces vidéos n'est pas de constituer des documentaires sui generis mais de fournir un matériel de travail pratiquement brut pour chercheurs, enseignants ou toute personne désireuse de s'initier, de mieux connaître, de « toucher avec ses yeux » la réalité de tous les jours d'un hongkongais.

▶ Régime général "Creative Commons" relatifs au document source

▶ comment citer



PROPRIÉTÉS    OBJETS EXISTANTS    NOUVEAUX OBJETS

▶ Propriété de classes

▼ comment ?

Aspects rhétoriques

Points de vue relatifs au sujet

locution

nom public

nom court

usage

▶ quoi ?

▼ quand ?

Le sujet par siècle

Le sujet par périodes ou époques

où ?

qui ?

lien wikipédia

Type

beginAt

endAt

sHt shat

SOMMAIRE   CONTRIBUTEUR   RESSOURCE   **SUJET**

propos

Sujet "Villes, villages et autres localités du Monde"

- intitulé
- type
- Localités
  - Ville
- Le sujet par année
  - 2007
- Thèmes plus spécifiques à explorer
  - Thème de discours
- Le sujet par pays
  - Chine (République populaire)
- relation sémantique

← Précédent

fresh  
escom

Dans la collection "Road Movies"  
la Fondation  
Maison des Sciences de l'Homme  
et  
l'Equipe Sémiotique Cognitive  
Nouveaux Médias

00:00:00.00 / 01:33:20.17

OBJETS EXISTANTS   NOUVEAUX OBJETS

Locaux

- Sujet "Aliments du monde"
- Sujet "Cuisines du monde"
- Sujet "Villes, villages et autres localités du Monde"

Saisir des mots clés

- Accès >
- 1- Sujet >
- Sujet "Domaine de recherche en SHS" ( 276 )
  - Sujet "Histoire du monde" ( 45 )
    - Sujet "Histoire des lieux" ( 5 )
    - Sujet "Récit historique" ( 3 )
    - Sujet "L'histoire par dates" ( 5 )
    - Sujet "Personnalité historique" ( 1 )
    - Sujet "Événement marquant" ( 4 )
    - Sujet "L'histoire par époques" ( 12 )
    - Sujet "Histoire des peuples" ( 4 )
    - Sujet "Histoire par pays ou régions" ( 12 )
    - Sujet "Servage et esclavage dans l'histoire" ( 3 )
    - Sujet "Histoire des mouvements sociaux" ( 2 )
    - Sujet "L'histoire par siècles" ( 10 )
  - Sujet "Culture religieuse" ( 15 )
  - Sujet "Littérature, arts et médias" ( 4 )
  - Sujet "Géographie, urbanisme et architecture" ( 22 )
  - Sujet "Agriculture, industrie et services" ( 7 )
  - Sujet "Diversité culturelle, patrimoine et mémoire" ( 47 )
  - Sujet "Monde moderne et contemporain" ( 78 )
  - Sujet "Monde social" ( 84 )
  - Sujet "Territoire, pays et peuple" ( 86 )
  - Sujet "Vie et vision politiques" ( 49 )
  - Sujet "Violence, domination et discrimination" ( 29 )
  - Sujet "Interrogation philosophique" ( 3 )
  - Sujet "Langue, culture et société" ( 4 )
- Sujet "Médecine et pharmacie" ( 1 )
- Sujet "Sciences de la nature" ( 8 )

Affichage sous forme de liste

Sujet "Histoire par pays ou régions"(4) X

Sujet "L'histoire par siècles"(4) X

## Résultats

0-4 / 4 items trouvés

Les funérailles princières en Europe, XVIe-XVIIIe siècle : le grand théâtre de la mort

De D. Pedro I du Brésil à D. Pedro IV de Portugal, avec un bref séjour à Paris

Colloque « Migrations et dynamiques urbaines » : Visibilités et invisibilités migrantes (2/2)

L'Angola entre recherche brésilianiste et recherche africaine - A l'occasion de la publication du numéro 1 de la revue Brésil(s) : «Vies d'esclaves»

Strate thématique

Strate visuelle

Strate visuelle



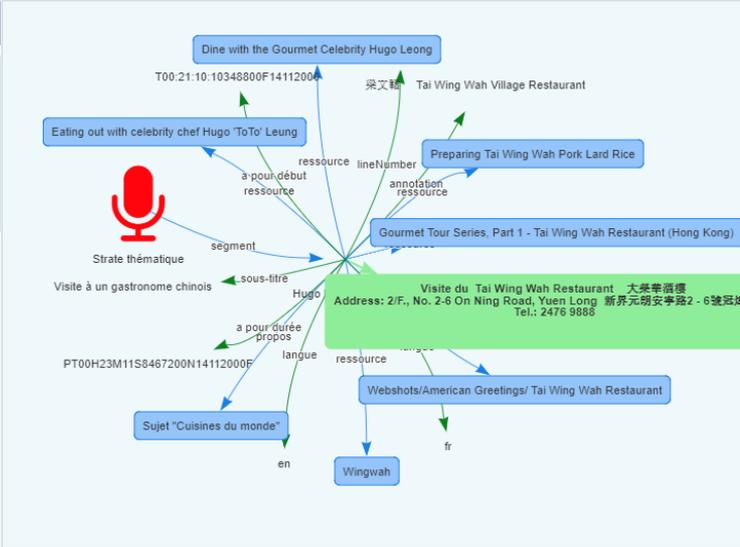
## Hugo Man-To Leung



SOMMAIRE RESSOURCE SUJET

propos

Sujet "Cuisines du monde"



OBJETS EXISTANTS NOUVEAUX OBJETS

- Locaux
- Sujet "Aliments du monde"
  - Sujet "Cuisines du monde"
  - Sujet "Villes, villages et autres localités du Monde"

← Précédent

## Portage complet en technologie Web

- En cours d'achèvement sur l'ensemble des services
- Plus de téléchargement d'outils, compatibilité Linux/Mac
- Navigation/saisie simplifiée : moins d'hyperliens
- Nouvelles fonctionnalités :

## Perspectives à court/moyen terme :

- Amélioration interfaces de saisie des divers types de connaissances : CNL, saisie voix via google API
- Génération de portails
- Amélioration structuration des processus de transformation
- Enrichissement palette de modèles de publication (projet MemoMines) et Widgets

## Partenariat

### INALCO (Peter Stockinger)

**P8 Paragraphe** : scénarios de valorisation alternatifs, interface d'annotation collaborative

IDEFI-CREATIC

### CCSD (Archives HAL)

Merci !!