

Conditor, vers un référentiel national de la production scientifique française

Une recommandation de BSN3 (signalement)
Construire un référentiel de la production scientifique française
en s'appuyant sur des référentiels communs
aux universités et organismes de recherche
(unités, thématiques, auteurs,...)

Journée référentiels Couperin
3 juillet 2015

Annie Coret (CNRS-DIST)
Alain Zasadzinski (CNRS-INIST)

Sommaire

❑ 1 Conditor

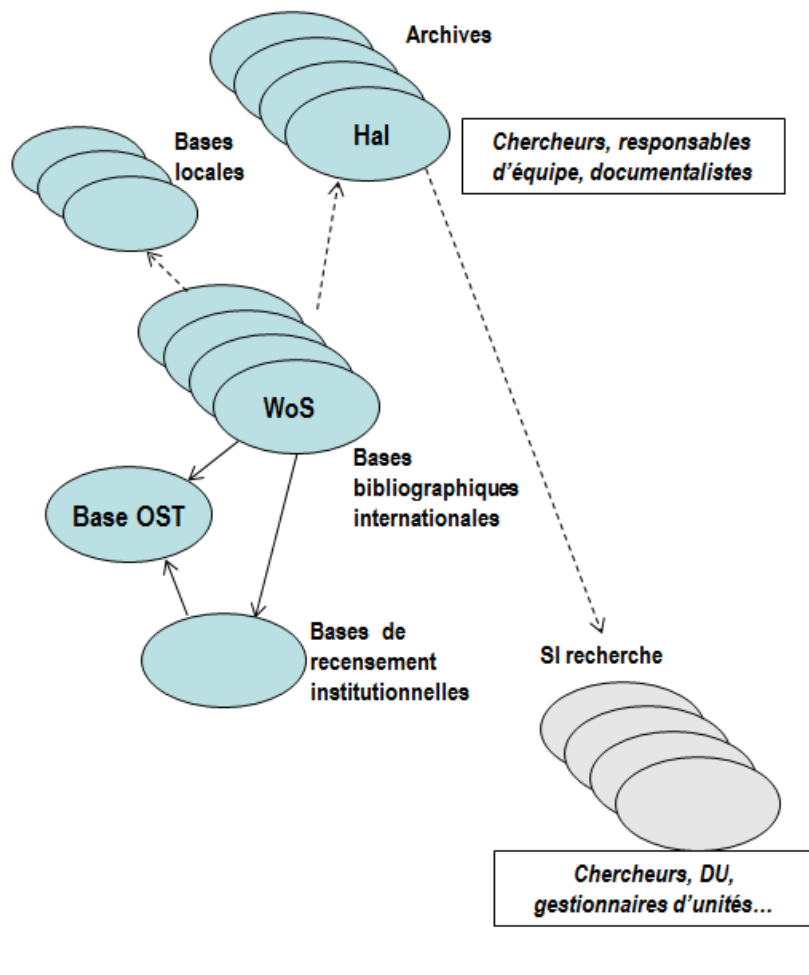
- Origine du projet
- Un référentiel national
- Le paysage
- Grands principes de construction
- Méthode de constitution du corpus (expérimentation)

❑ 2 Les « autres » référentiels

- Utilisés pour l'expérimentation
- Constitués dans le cadre de l'expérimentation
- En cours d'étude ou à étudier pour la suite
- Apport du RNSR

❑ 3 En guise de conclusions

1 Conditor : origine du projet



Référentiels « IST »

Adresses, thématiques, auteurs, revues, colloques...

De multiples dispositifs, acteurs, saisies, travaux de recensement... pour différents usages

Référentiels de gestion

Structures, personnel, projets ANR...

Pas de vision complète Bases internationales insuffisantes

« **Savoir ce qu'on a en magasin** » (président de l'université Paris Descartes en juin 2014)

Limiter les multiples travaux de recensement et saisie

1 Conditor : un référentiel national

Un « pot » commun de métadonnées :

- ❑ **références bibliographiques**
 - Titre, auteurs, titre de revue, ISSN,...
- ❑ **structures ESR et tutelles**
- ❑ **identifiants numériques d'articles et auteurs (DOI...)**
- ❑ **lien vers le texte intégral**
- ❑ **+ financements, données liées, classifications / indexations...**

- ❑ **« exhaustif » (tous les « objets numériques » pertinents)**
 - ❑ **Typologie large : article, ouvrage, thèse, rapport,...**
 - ❑ **Utilisation des réservoirs ESR et des bases ou archives internationales**
- ❑ **« riche » (agrégation de données et utilisation de référentiels)**
- ❑ **« interopérable »**
- ❑ **« intégrable » dans les dispositifs de l'ESR**

1 Conditor : un référentiel national

~~Archive~~

~~Outil décisionnel~~

~~Portail~~

~~Systeme d'Information
Recherche~~

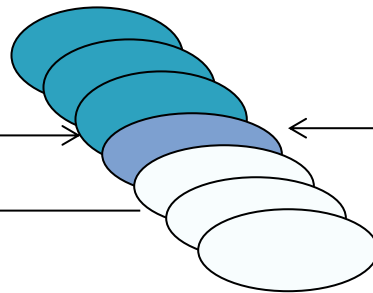
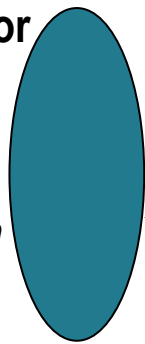
des métadonnées + un applicatif pour sa gestion + une équipe d'administration

Utilisateurs « directs »

Utilisateurs « indirects »

Conditor

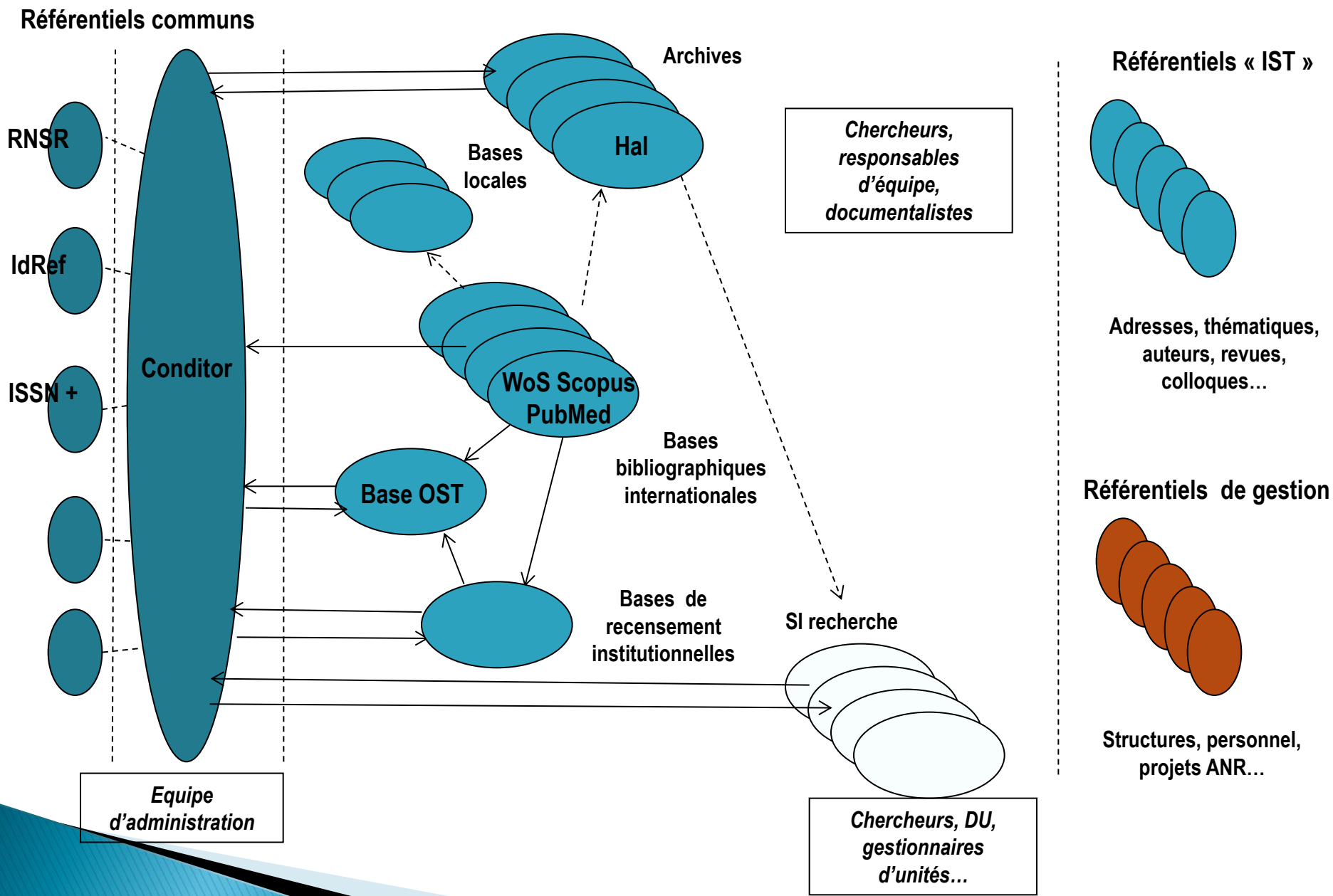
L'équipe
d'administration



Chercheurs,
responsables d'équipes,
producteurs d'indicateurs,
décideurs...

Archive nationale
Archives institutionnelles
Bases bibliométriques
Outils décisionnels
SI Recherche...

1 Conditor dans le paysage



1 Conditor : grands principes de construction

Privilégier le **pragmatisme**

Miser sur la **complémentarité** et la **mutualisation**

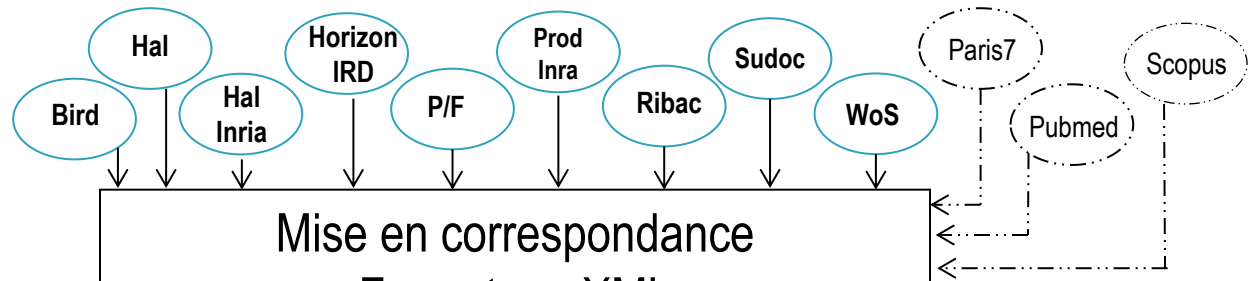
- ❑ Utiliser les données dont on dispose
- ❑ S'organiser progressivement pour minimiser les efforts de chacun
- ❑ Avancer en faisant
 - => Une expérimentation pour apprécier concrètement :
 - les difficultés **de tous ordres** à surmonter
 - les bénéfices **potentiels** pour chaque partenaire
 - l'implication nécessaire de chacun

Participants à l'expérimentation :

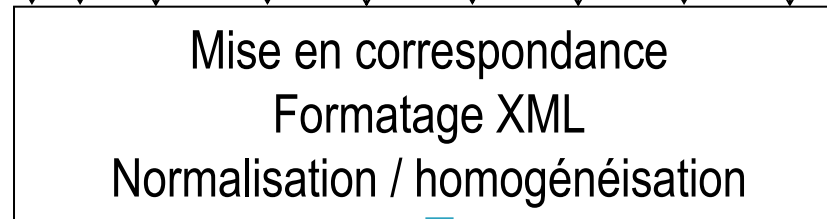
- Abes
- CNRS (Dist, Inist, CCSD, INSHS, SAP2S)
- Inra
- Inria
- IRD
- Ministère de la recherche (RNSR)
- Université de Bordeaux
- Université Paris Dauphine

**Equipe
multi-partenaires
et multi-métiers**

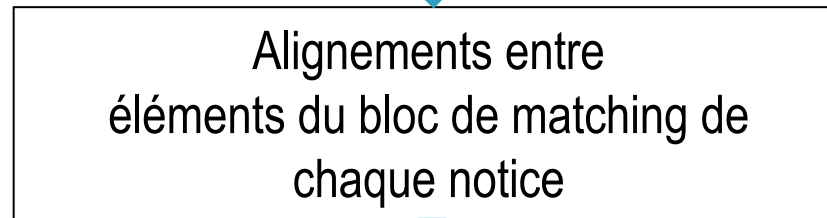
1 Conditor : méthode de constitution du corpus (expérimentation)



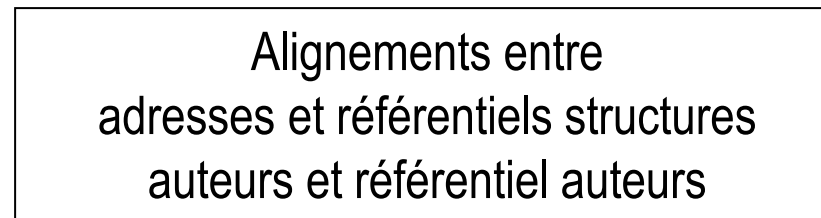
Etape 1
Préparation des
Métadonnées



Etape 2
Détection des
doublons



Etape 3
Enrichissements à l'aide
des référentiels



- « Bloc de matching »
- Identifiants
 - Titres du document
 - Auteurs
 - Sources
 - Collations
 - Adresses
 - Types de document

Référentiels utilisés

- RNSR
- Structures CNRS (Labintel)
- IdRef



Notice finale

- Méta données sources
- Bloc de matching
- Enrichissements

2 Référentiels utilisés pour l'expérimentation

Référentiels	Notice Conditor	Chaînes de caractères recherchées	Métadonnées ajoutées
RNSR	Adresses	Code ou sigle ou libellé unité + ville ou tutelle	Identifiant national de structure et tutelles
Référentiel de structures CNRS	Adresses	Code ou sigle ou libellé unité + ville ou tutelle	Code unité CNRS et tutelles
IdRef Référentiel des autorités de l'Abes	Auteurs	Nom et prénom	Identifiant auteur IdRef (à valider)
Registre ISSN	ISSN ou nom de revue	ISSN ou nom de revue	ISSN associés, ISSN de liaison, ...
Codes pays ISO 3166	Codes pays	Code pays	Code pays verbalisé

2 Référentiels constitués dans le cadre de l'expérimentation

Référentiels	Notice Conditor	Chaînes de caractères recherchées	Métadonnées ajoutées
Tutelles	Affiliations Tutelles	Alignement et homogénéisation des formes d'écritures des tutelles	Forme unifiée de chaque établissement
Départements et régions de France (wikipédia)	Adresses Affiliations	Code « département » dans les codes postaux	Départements et régions

2 Autres référentiels en cours d'étude ou à étudier

Référentiels	Notice Conditor	Chaînes de caractères recherchées	Métadonnées à ajouter
Projets Cordis (FP6 et FP7)	Résumés + Champs spécifiques	Identifiant, sigle et nom de projet	Identifiants RCN, sigle, nom du projet, nom de programme FP, pays coordinateur, dates ...
Annuaire des écoles doctorales (data.gouv.fr)	Thèses	Nom d'école doctorale	Code école doctorale
Etablissements ESR (data.gouv.fr)	Tutelles	Nom d'établissement (tutelles)	Code UAI

Autres référentiels à étudier en vue d'enrichissements éventuels des notices :

- Identifiants Auteurs : ResearcherID / ORCID / ISNI / VIAF/ scopus_auth-id / idHal... ?
- Identifiants Projets : ANR... ?
- Identifiants Articles : DOI (PMID, UT, scopus_EID, IsteX_ID, Refdoc_id ...)... ?
- Revues : JCR, heloise, SherpaRomeo, APC (à créer ?) ... ?
- *Colloques* : à créer ?
- Classifications thématiques : classifications nationales et internationales (MeSH, LCSH, Rameau, Dewey, ERC, JEL, Arxiv, HAL ...), WoS categories, Scopus catégories ... ?

2 Apport du RNSR

Mise à disposition d'une notice Conditor potentiellement pertinente pour un partenaire

Notice Conditor

Métadonnées Hal

U1016 INSERM, **UMR8104** CNRS, Université
Paris V - Paris Descartes, Institut Cochin, 22 rue
Méchain, 75014 **Paris**, FRANCE

Métadonnées Pascal/Francis

Inserm, **U1016**, Institut Cochin, Paris, France
and Cnrs, **UMR8104**, **Paris**, FRANCE

Éléments « matchant » avec le RNSR

Item RNSR

<code>**200217519N**

<LIBELLE>**CENTRE DE RECHERCHE INSTITUT COCHIN**

<SIGLE>**UM3**

<VILLE>**PARIS**

<LABO>

<CODE_UNITE>**U1016**

<TUTELLE_SIGLE>**INSERM**

<LABO>

<CODE_UNITE>**UM3**

<TUTELLE_SIGLE>**CHR CHU**

<LABO>

<CODE_UNITE>**UM3**

<TUTELLE_SIGLE>**INRA**

<LABO>

<CODE_UNITE>**UM3**

<TUTELLE_SIGLE>**PARIS 5**

<LABO>

<CODE_UNITE>**UMR8104**

<TUTELLE_SIGLE>**CNRS**

Enrichissements potentiels

Notice présente dans 2 réservoirs ESR : Hal, Pascal/Francis

AMP-activated protein kinase and metabolic control. [More like this](#)

Auteurs Andreelli, Fabrizio ; Viollet, Benoit ;

Source Handb. exp. pharmacol. 203:303-330, 2011.

Journal Handbook of experimental pharmacology ;

ISSN 0171-2004 - Print ; 1865-0325 - Online ;

ISSN liaison 0171-2004 ;

Type Article

id-conditor 091691-conditor

Origine HAL INIST

Pays GERMANY

Editeur

Springer, Heidelberg.

Financement

ANR-06-PHYSIO-026: AMP-activated et S6 kinases: roles opposés dans l'adaptation du muscle squelettique à l'état nutritionnel et à l'exercice.

EXGENESIS: Health benefits of exercise: identification of genes and signalling pathways involved in effects of exercise on insulin resistance, obesity and the metabolic syndrome.

Association Française contre les Myopathies (AFM) Association pour l'Etude des Diabètes et des Maladies Métaboliques (ALFEDIAM) Institut Benjamin Delessert.

Adresse affiliation HAL

Assistance publique - Hôpitaux de Paris (AP-HP), Hôpital Pitié-Salpêtrière, Université Pierre et Marie Curie (UPMC) - Paris VI, Service de diabétologie Pitié-Salpêtrière, 47-83, Boulevard de l'Hôpital 75651 Paris Cedex 13, FRANCE

U1016 INSERM, UMR8104 CNRS, Université Paris V - Paris Descartes, Institut Cochin, 22 rue Méchain, 75014 Paris, FRANCE

Adresse affiliation Inist

Department of Diabetology, Pitie-Salpetriere Hospital AP-HP, Univ Pierre et Marie Curie-Paris 6, Paris, FRANCE.

Fakultät Lebenswissenschaften, Molekulare Pharmakologie und Toxikologie, Technische Universität Braunschweig, Beethovenstrasse 55, 38106 Braunschweig, GERMANY.

Inserm, U1016, Institut Cochin, Paris, France and Cnrs, UMR8104, Paris, FRANCE.

Univ Paris Descartes, Paris, FRANCE.

Structure d'affiliation (équipe, labo, ...)

CENTRE DE RECHERCHE INSTITUT COCHIN - RNSR: 200217519N [2002] U1016 (INSERM) ; UM3 (CHR CHU - INRA - PARIS 5) UMR8104 (CNRS)

INSTITUT COCHIN - CNRS: UMR8104

Tutelles

AP-HP; CNRS; INRA; INSERM; UNIV PARIS 5 - PARIS DESCARTES; UNIV PARIS 6 - PIERRE ET MARIE CURIE;

Liens externes

HAL [inserm-00597010](#)

INIST 24091751

DOI [10.1007/978-3-642-17214-4_13](#)

pdf [[AMPK_metabolic_control_HAL_.pdf](#), [inserm-00597010_edited.pdf](#), [index.xhtml](#),]

IDREF

ANDREELLI, Fabrizio [035798025](#)

LABO [UMR8104](#);

Sources HAL et base Pascal/Francis

Information présente dans Hal mais pas dans la Pascal/Francis

Curation et enrichissement

Adresse présente dans la base Pascal/Francis mais pas dans Hal

Information ajoutée grâce au RNSR

3 En guise de conclusions

Un référentiel national **décrivant** la production scientifique est **faisable**

Un « pot » commun est **nécessaire** pour mutualiser les efforts

- ❑ Conditor contribuera à l'**amélioration** des référentiels existants
- ❑ Conditor favorisera l'**interopérabilité** entre référentiels
- ❑ Conditor contribuera à la **mise en place** de référentiels communs quand ils n'existent pas (colloques et APC par exemple)

Si on veut qu'un référentiel national existe et s'améliore, il faut :

- ❑ **l'utiliser et l'intégrer à son système,**
- ❑ **contribuer à son enrichissement et à son évolution...**

Merci de votre attention

La suite des travaux : des résultats opérationnels à l'horizon 2 ans

Macroplanning (après organisation du projet)



Négociation des ressources et des conventions

Constitution d'un corpus de notices sur plusieurs années

Définition des métadonnées à gérer et de la stratégie de construction puis mise en œuvre dont formatages, ...

Corpus 0

Corpus 1

Conception et développement itératif de l'applicatif

Mise en place progressive des fonctionnalités

Recherche, modification, suppression, validation, mise à disposition

Collecte, formatage, dédoublonnage
Enrichissements via les référentiels
Industrialisation des échanges

Conception et mise en place du service Conditor

Organisation de l'équipe d'administration et de la structure pérenne de régulation, gestion de la montée de puissance...