

# CASTORE



## Plateforme logicielle libre de bibliothèque numérique de littérature grise



# CASTORE

Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
Positionnement  
Architecture technique

# CASTORE

## Bases du projet

La plateforme - démonstration

Castore aujourd'hui

Positionnement

Architecture technique

- Enjeux
  - Favoriser l'accès au savoir en écartant les contraintes de lieux et de temps
- Enseignement Recherche (Administration / Vie Courante)
  - Valoriser la production de documents
  - Diffuser
  - En assurer la conservation
- Répondre à des besoins des instituts d'enseignement et de recherche
  - Capitaliser
  - Diffuser
- Maîtrise locale
  - De la production documentaire
  - De sa diffusion

# Besoins

- "Les services de documentation de l'enseignement supérieur doivent aujourd'hui choisir entre deux options : devenir exclusivement des lieux d'aiguillage vers des réservoirs de ressources payants dont la maîtrise leur échappe, ou parallèlement contribuer à la construction des bibliothèques numériques d'accès placées sous leur contrôle".  
(Une ambition universitaire légitime, février 2003, Université de Lyon 2)
- "Beaucoup d'universitaires ressentent le besoin de rappropriation collective de l'information produite dans leurs laboratoires et concluent à la nécessité de se doter d'un entrepôt numérique pour archiver et diffuser leurs productions. La bibliothèque intervient ici pour assurer non seulement la qualité et la normalisation de l'archivage mais également le prolongement de la démarche à l'échelle d'un réseau national et international. Son rôle de fournisseur d'information se trouve ainsi enrichi."  
(Grisemine, une bibliothèque numérique de littérature grise, Marie-France Claerebout, Jean-Bernard Marino, GL5, dec 2003)

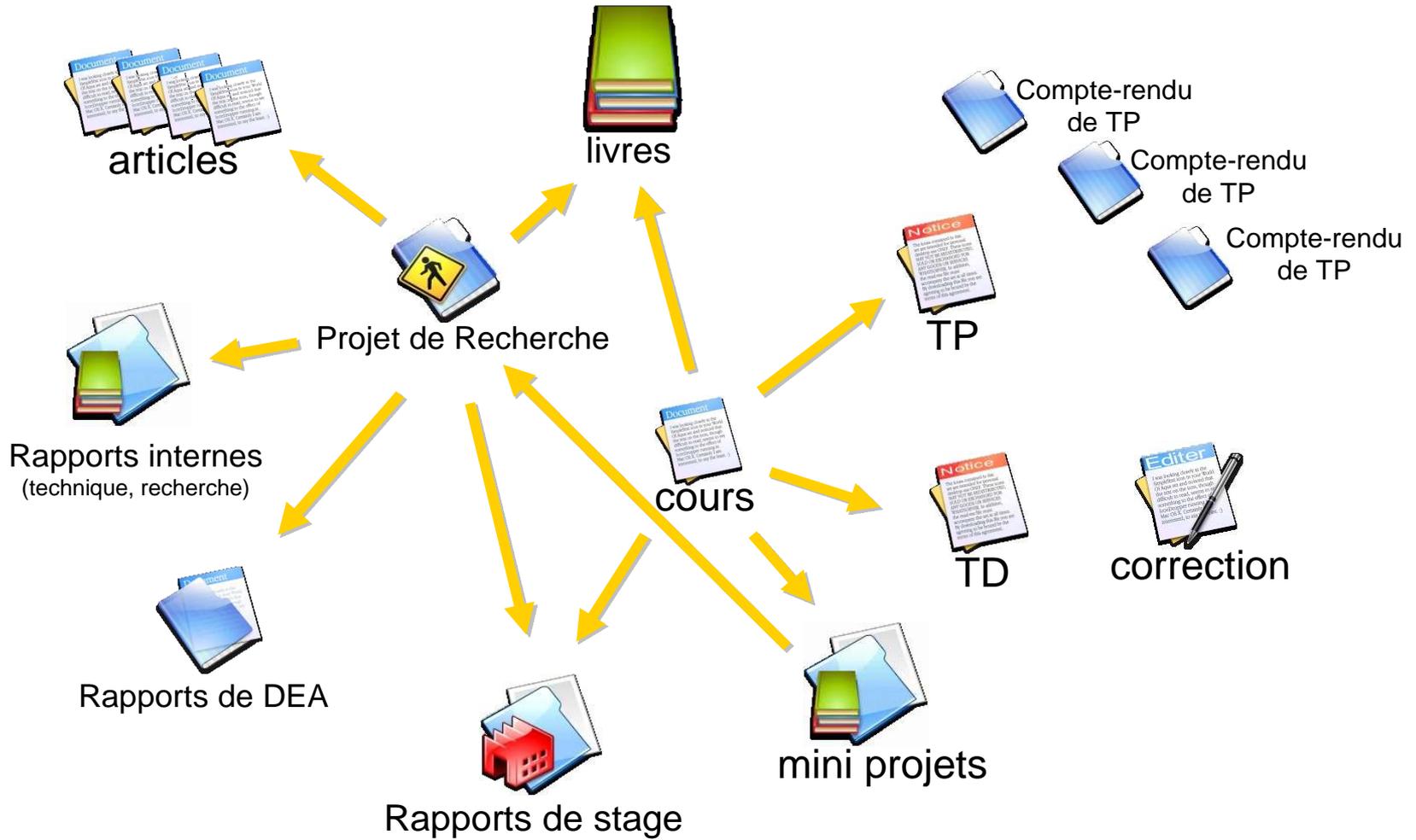
# Étude des Besoins (1)

- La place centrale du document dans les activités de recherche comme d'enseignement
- L'omniprésence du document numérique :
  - gain de temps (accès rapide à l'information, évaluation immédiate de la pertinence des résultats, limitation des déplacements)
  - facilité d'accès à l'information
  - facilité de recherche dans le texte
- Des difficultés dans la recherche d'information  
→ système D

# Étude des Besoins (2)

- Disposer de plus de ressources sous format numérique
  - de l'information scientifique
  - en texte intégral
  - aisément accessible
  - rapidement accessible
  - disponible dans des outils ciblés pour
    - plus de fiabilité
    - éviter de perdre son temps à trier la masse de résultats inégaux fournis par les moteurs généralistes
- Avoir une meilleure appréciation de la teneur d'un document avant sa visualisation

# Les documents...



# Les documents...



articles



livres



Compte-rendu  
de TP



Compte-rendu  
de TP



Compte-rendu  
de TP



Projet de Recherche



TP



Rapports internes  
(technique, recherche)



COURS



TD



correction



Rapports de DEA

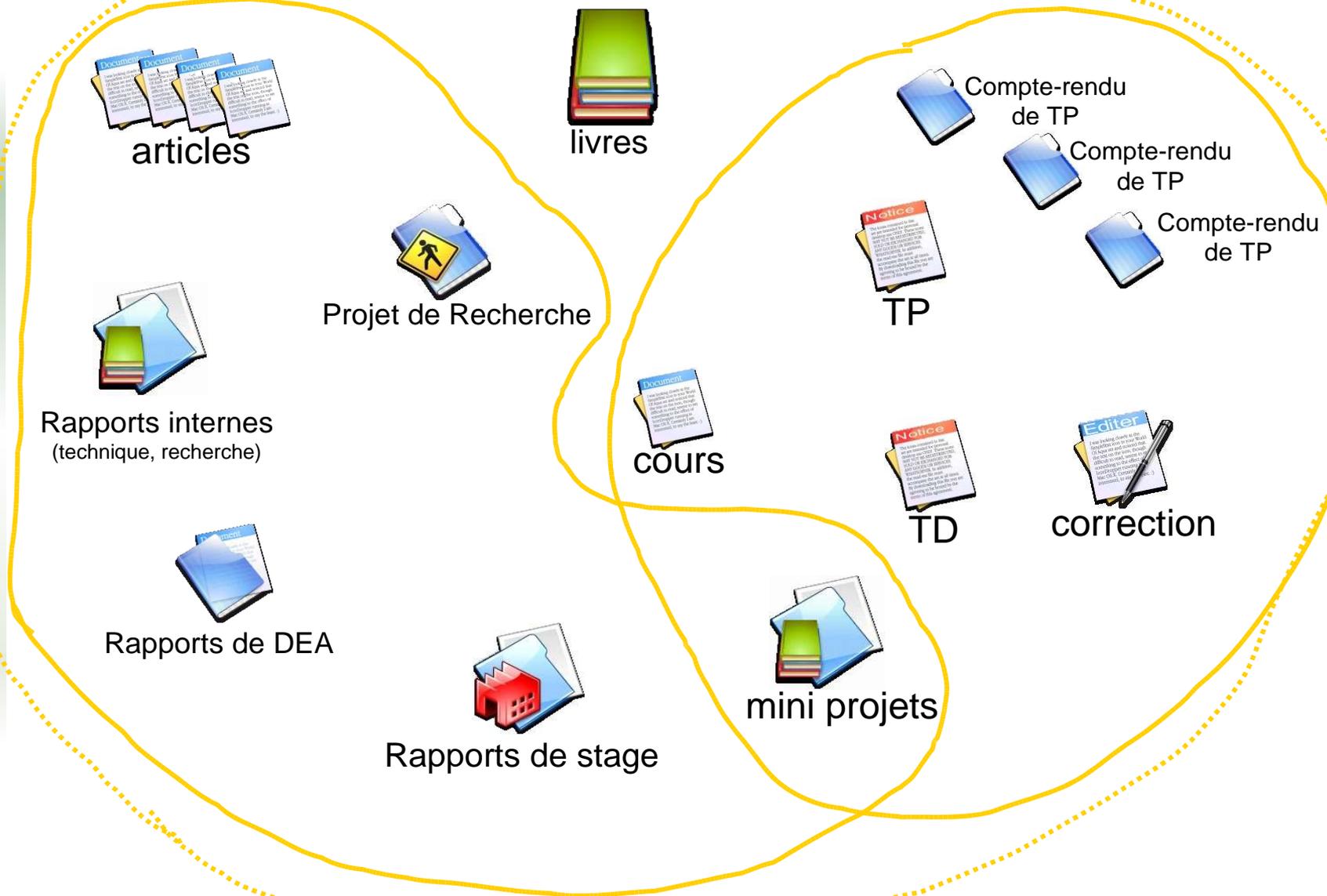


Rapports de stage



mini projets

# Les documents...



# Les documents...



articles



livres



Compte-rendu  
de TP



Compte-rendu  
de TP



Compte-rendu  
de TP



Projet de Recherche



TP



Rapports internes  
(technique, recherche)



COURS



TD



correction



Rapports de DEA



Rapports de stage

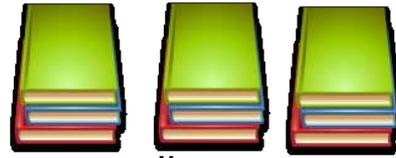


mini projets

# Les documents...



articles



livres



Compte-rendu  
de TP



Compte-rendu  
de TP



Compte-rendu  
de TP



Projet de Recherche



TP



Rapports internes  
(technique, recherche)



COURS



TD



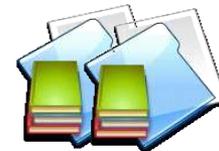
correction



Rapports de DEA

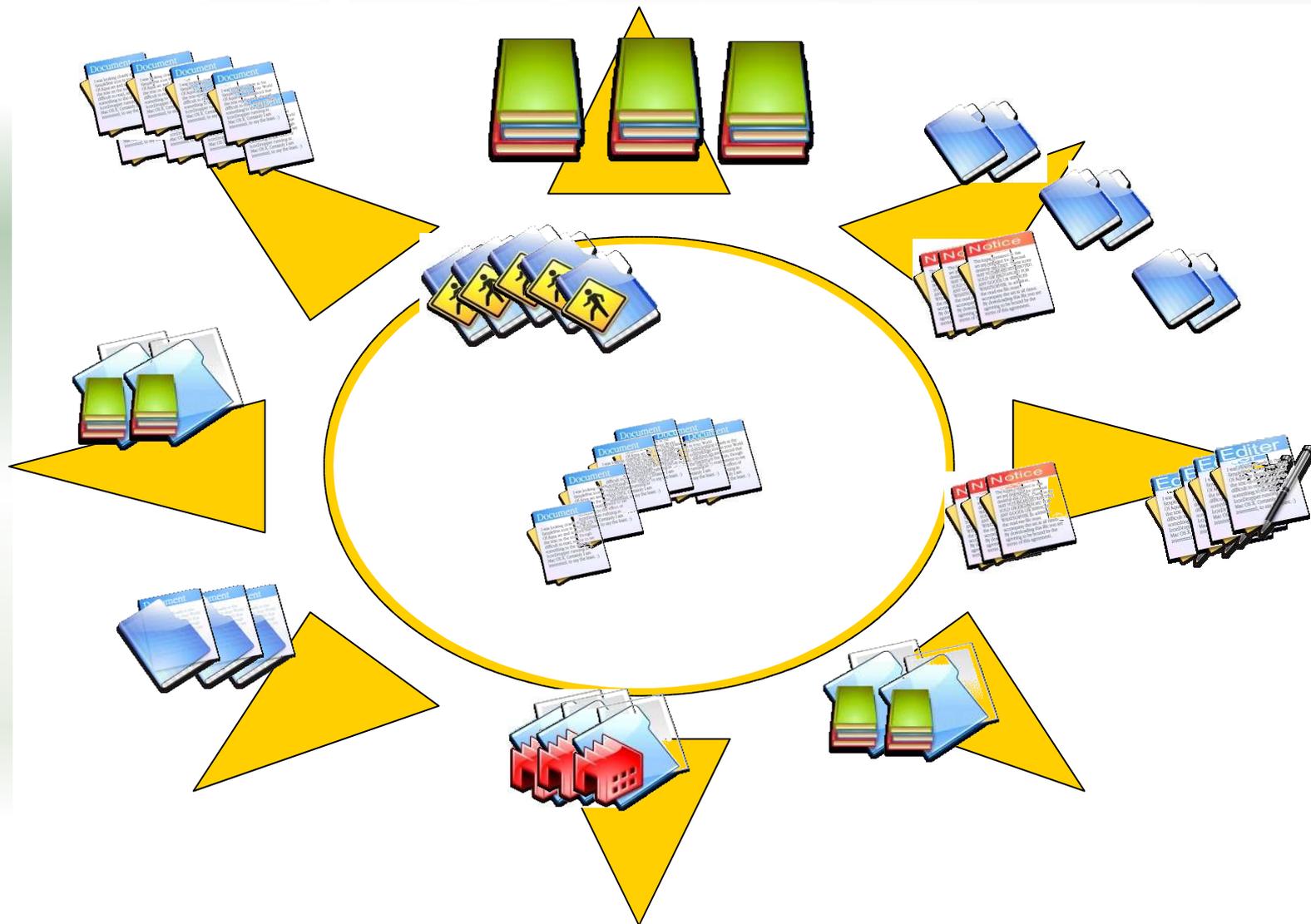


Rapports de stage

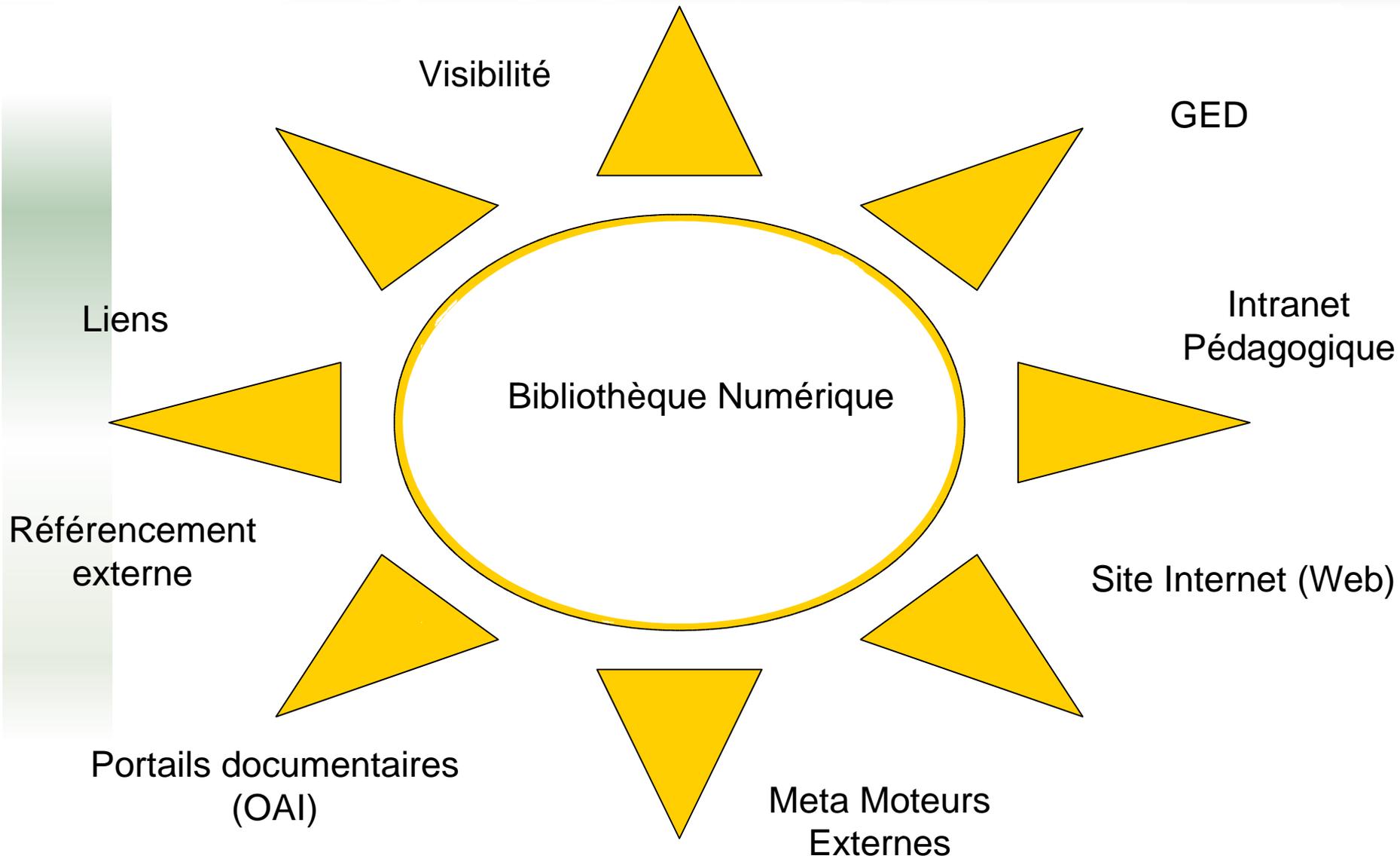


mini projets

# Les documents...



# Les documents...



- Chaîne documentaire minimale
  - soumission/diffusion
- Limité la charge de travail
  - Pas d'intervention manuelle sur les documents (gros volumes)
  - Installation / Configuration / Administration (taille centres)
- Respect des normes documentaires
  - OAI, Documents en XML et Metadonnées en DC
- Intégration dans l'existant
  - Architecture à composants (J2EE)
  - Multimédia, clients lourds, etc.

# CASTORE

Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
Positionnement  
Architecture technique

# CASTORE

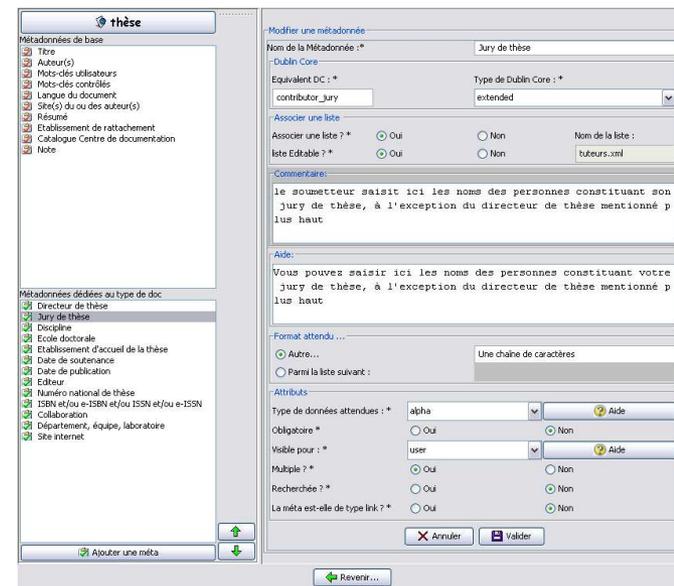
Bases du projet  
**La plateforme - démonstration**  
Castore aujourd'hui  
Positionnement  
Architecture technique

# Plateforme Castore

- Documents transformés en XML
- Méta données XML en Dublin Core
- Gestion des droits sur le document (licence et droit d'accès)
- Utilisation à travers un client léger (tout navigateur web)
- Simple et ergonomique
- État de l'art des technologies du Web

# Types de documents

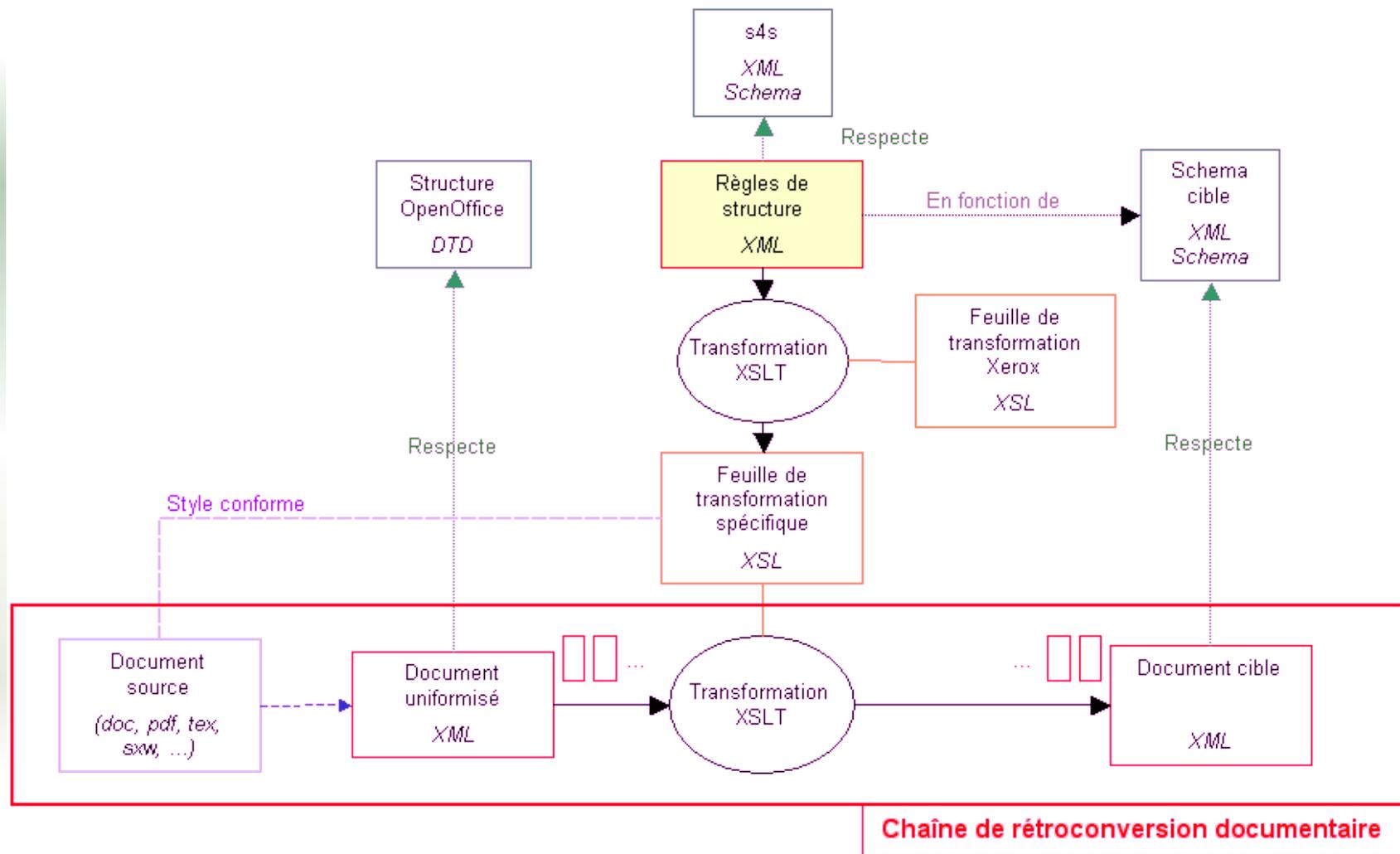
- Classés en catégories
- Chaîne de validation
  - Séquentielle
  - Par personne ou par groupe
  - Différents niveaux de validation
- Metadonnées associées
- Tout est configurable.



The screenshot displays a web-based interface for configuring metadata for a document titled 'thèse'. The interface is divided into several sections:

- Métadonnées de base:** A list of standard metadata fields including Titre, Auteur(s), Mots-clés utilisateurs, Mots-clés contrôlés, Langue du document, Site(s) du ou des auteur(s), Résumé, Etablissement de rattachement, Catalogue/Centre de documentation, and Note.
- Métadonnées dédiées au type de doc:** A list of specific metadata fields for the 'thèse' type, such as Directeur de thèse, Jury de thèse, Discipline, Ecole doctorale, Etablissement d'accueil de la thèse, Date de soutenance, Date de publication, Editeur, Numéro national de thèse, ISBN et/ou e-ISBN et/ou ISSN et/ou e-ISSN, Collaboration, Département, équipe, laboratoire, and Site internet.
- Modifier une métadonnée:** A configuration panel for a specific metadata field, 'Jury de thèse'. It includes:
  - Nom de la Métadonnée: \*** Jury de thèse
  - Equivalent DC: \*** Dublin Core
  - Equivalents DC: \*** contributor\_jury
  - Type de Dublin Core: \*** extended
  - Associer une liste? \*** Oui (selected)
  - Associer une liste? \*** Non
  - Associer une liste? \*** Oui (selected)
  - Associer une liste? \*** Non
  - Nom de la liste:** tuteurs.xml
  - Commentaire:** Le soumetteur saisit ici les noms des personnes constituant son jury de thèse, à l'exception du directeur de thèse mentionné plus haut.
  - Aide:** Vous pouvez saisir ici les noms des personnes constituant votre jury de thèse, à l'exception du directeur de thèse mentionné plus haut.
  - Format attendu...:** Autre... (selected), Une chaîne de caractères
  - Attributs:** Type de données attendues: \* alpha, Aide; Obligatoire: \* Oui (selected), Non; Visible pour: \* user, Aide; Multiple: \* Oui (selected), Non; Recherchée: \* Oui (selected), Non; La méta est-elle de type link: \* Oui (selected), Non.
  - Buttons: Annuler, Valider

# Conversion des documents



# Règles de structuration

- Typographie et/ou nom de style -> élément XML

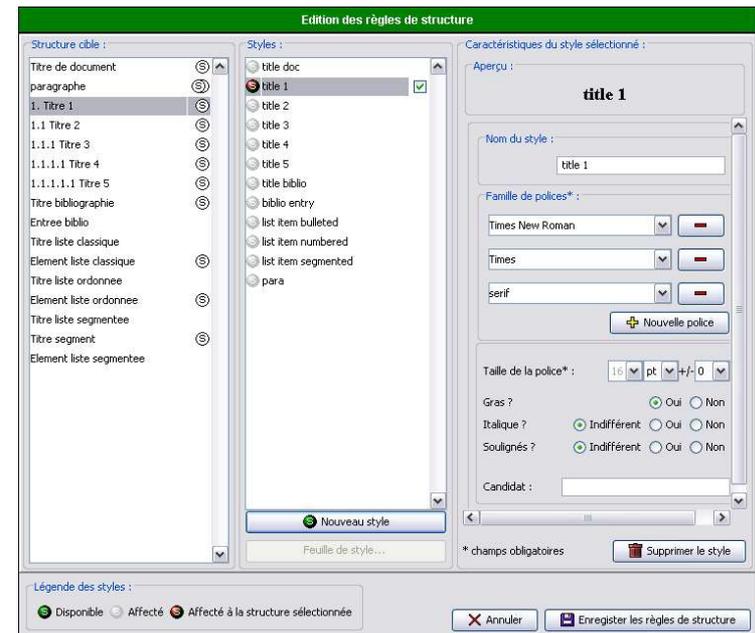
```

<structure.rule id="r.1">
<criteria scope="unique" id="r.1.c">
<fp>
  <font-family>
    <font-name> Times New Roman</font-name>
    <font-name> Times</font-name>
  </font-family>
  <font-size>
    <value> 16 </value>
    <unit> pixel </unit>
  </font-size>
  <font-weight> normal </font-weight>
</fp>
<source-style-name> title 1 </source-style-name>
</criteria>

<pivot.element> title </pivot.element>
<target.element> sect1 </target.element>

</structure.rule>

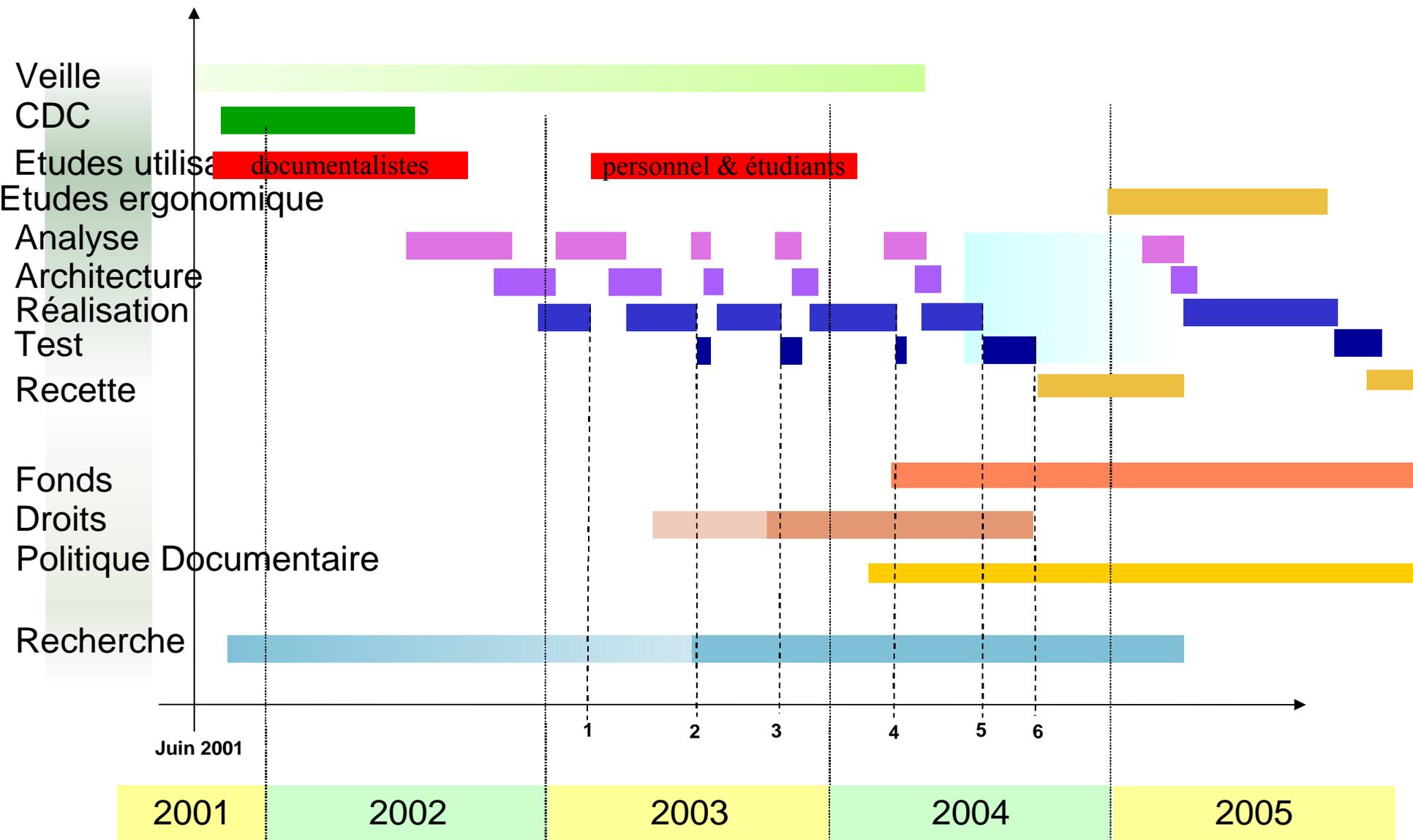
```



# Principales Étapes

- Étude depuis juin 2001
- Développement depuis septembre 2002
- Mise en logiciel libre le 1<sup>er</sup> janvier 2004 de la version 1.0
- Mise en production le 1<sup>er</sup> mars 2004 à l'EMN
  - Rapports de projets 2<sup>ème</sup> et 3<sup>ème</sup> années
  - Cours aux élèves

# Planning : réalisations



# CASTORE

Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
Positionnement  
Architecture technique

# CASTORE

---

Bases du projet  
La plateforme - démonstration  
**Castore aujourd'hui**  
Positionnement  
Architecture technique

# Principales réalisations 2005

2005

- Etude ergonomique de mars à août 2005
- Ajout des modules :
  - Administration documentaire
  - Support du protocole OAI-PMH
  - Support de JDO en libre
  - Nouvelle chaîne de retro conversion XML
- Travail de design

# Futures réalisations 2006

- Gestion de documents multimédias
- Ajout des modules :
  - Administration documentaire (WF)
  - Nouvelle chaîne de retro conversion XML
    - OpenOffice 2.0, médias
    - PDF
- Services Web et Portlets
- Import de métadonnées
- Gestion multi-Fonds

- Plateforme stabilisée
  - Version 1.0.2 le 30 mai 2005
- Distribution de la version 1.1.4
  - Université de Nantes
  - Adullact.org
- Extension du projet
  - Partenaires industriels
  - Projet UVPL

# Partenaires du projet

- Xerox
- Code Lutin
- Dynix
- Université de Nantes
- UVPL
  - Université d'Angers
  - Université du Mans
- etc...

## ■ Adullact

<http://adullact.net/projects/castore/>

## ■ Diffusion : 1.1.4 (sources + war + exe)



The screenshot shows the Adullact project page for Castore. The page features a search bar at the top with the text "Le projet entier" and a "Rechercher" button. Below the search bar is a navigation menu with tabs for "Accueil", "Ma page", "Arbre des projets", "Boite à outils", "Demande d'aide", and "Bibliothèque Numérique Castore". A secondary menu includes "En bref", "Forums", "Suivi", "Listes", "Documents", "Annonces", "Sources", and "Téléchargements". The main content area contains a description of Castore as an open-source digital library platform, a list of project details (category, environment, state, functions, license, etc.), and a list of team members under the heading "Equipe-Projet". At the bottom, there is a section for "Derniers fichiers publiés" with columns for "Paquet", "Version", "Date", "Remarques / Surveillance", and "Téléchargement".

**GFORGE ADULLACT** Le projet entier   Rechercher par avancées [S'identifier](#) [Nouveau compte](#)

Accueil Ma page Arbre des projets Boite à outils Demande d'aide **Bibliothèque Numérique Castore**

En bref Forums Suivi Listes Documents Annonces Sources Téléchargements

Castore est une plateforme open-source de bibliothèque numérique pour Conserver, Valoriser et Diffuser le patrimoine documentaire de votre institut. Basée sur une gestion complète en XML, elle permet de stocker, annoter et décrire vos documents.

- Catégorie: Gestion de contenus, Partage des documents, Entrepôts de données
- Environnement: Web Environment
- Etat d'avancement: S - Production/Stable
- Fonctions système: File Sharing, Internet, Archiving
- Licence: CeCILL (licence française compatible GPL)
- Métiers: Administrations, Education, Informatique
- Programming Language: Java
- Public concerné: Administratifs, Facilitateurs, Utilisateur final
- Système d exploitation: OS Independent

Enregistré le : 12/09/2005 16:25  
Taux d'activité : 96%  
[Voir les statistiques d'activité du projet.](#)

**Equipe-Projet**  
Administrateurs :  
Cédric Dumas  
Olivier Grouhan  
Stéphane Bouchet  
Développeurs :  
Emilie DIETRICH  
Peggy Bardon  
Thomas MORIN

[\[Voir les membres\]](#)  
[\[Demander à rejoindre le projet\]](#)

**Derniers fichiers publiés**

Paquet	Version	Date	Remarques / Surveillance	Téléchargement
--------	---------	------	--------------------------	----------------

# CASTORE

Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
Positionnement  
Architecture technique

# CASTORE

---

Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
**Positionnement**  
Architecture technique

- Littérature grise
  - Plateformes libres
    - Composants
      - Dspace (MIT)
      - dLibra (WBC)
      - Archimède (U. Laval)
    - SDX (Cocoon)
      - Cyberdoc (UL2)
    - Script
      - CDSWare (CERN)

# Avantages

- **Pédagogiques**
  - Reconnaissance
  - Travaux des anciens et des nouveaux
- **Scientifiques**
  - Continuité des projets
  - Cohérence et qualité de l'information
  - Gestion électronique des thèses
- **Institutionnels**
  - Visibilité externe
  - Diffusion des résultats scientifiques
- **En 4 points**



- La conservation des thèses
- La prise en compte du patrimoine documentaire des instituts
- La publication libre des résultats scientifiques (Budapest Open Access Initiative )
- L'enjeu des centres de ressources autour de la connaissance

# Fédérer

2

- Un centre de ressource documentaire
  - Géré par les documentalistes
  - Centralisant la production
- Utilisé par toutes les composantes
- Animé par les cellules TICE

3

- Nouveaux outils, nouveaux usages  
nouveaux enjeux
- Évolution globale de l'impact des documents numériques :  
documentation d'entreprise, doc. technique, Web sémantique, etc.
- Besoin de formation

4

- Formations des élèves
  - Utilisation des documents numériques
  - Production des documents numériques
- Sensibilisation à la conservation du patrimoine scientifique auprès des enseignants chercheurs

# CASTORE

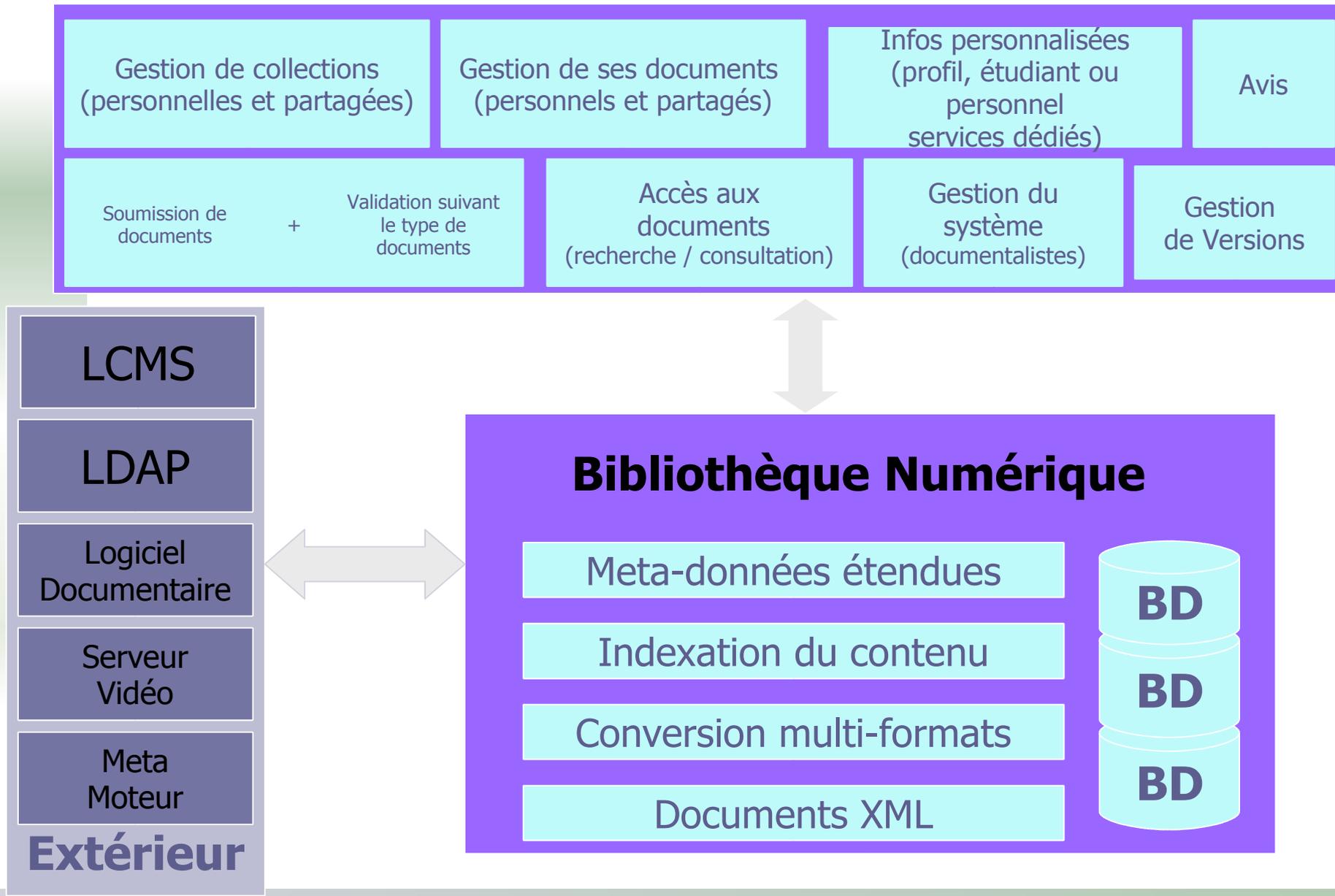
Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
Positionnement  
Architecture technique

# CASTORE

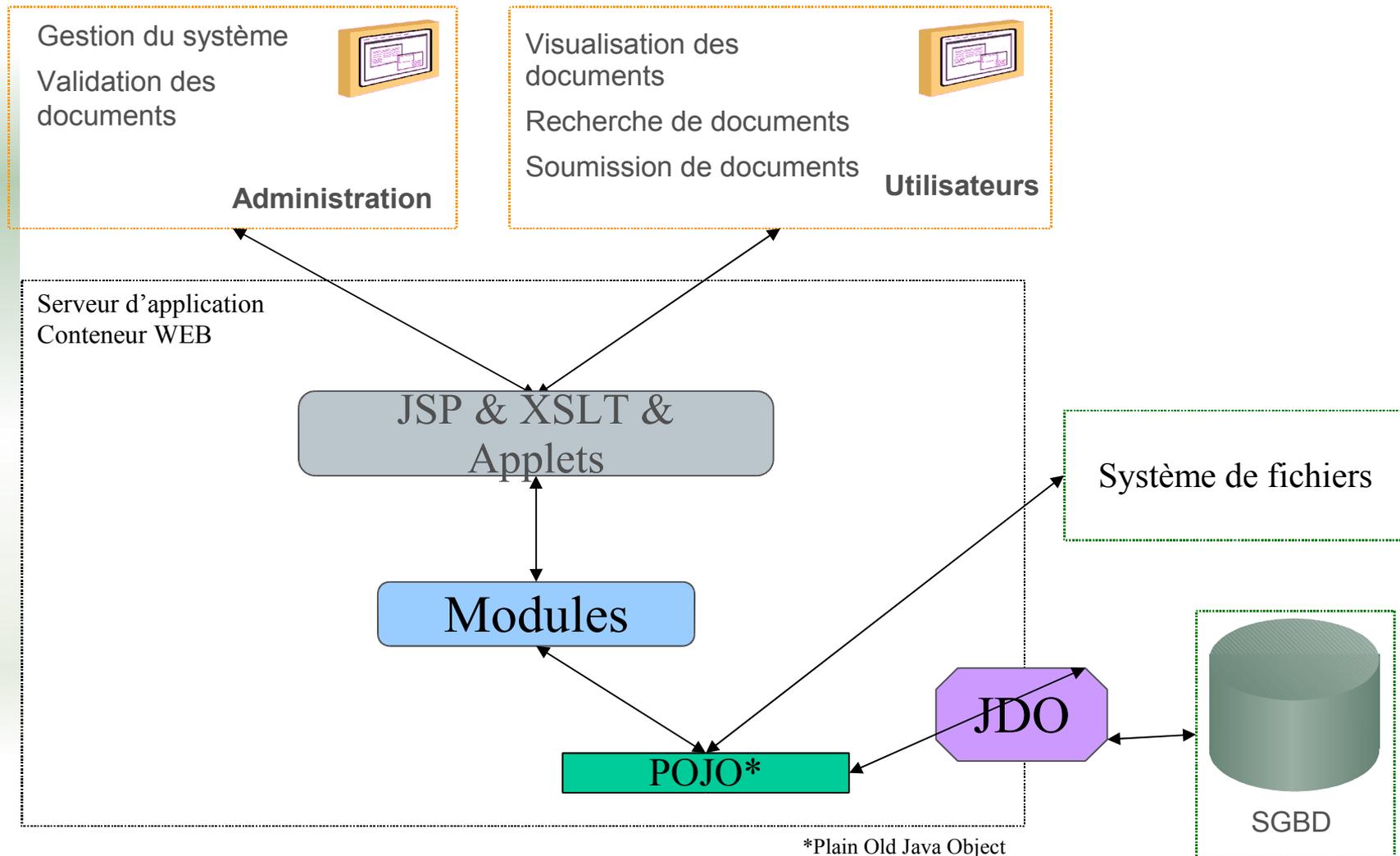
---

Bases du projet  
La plateforme - démonstration  
Castore aujourd'hui  
Positionnement  
**Architecture technique**

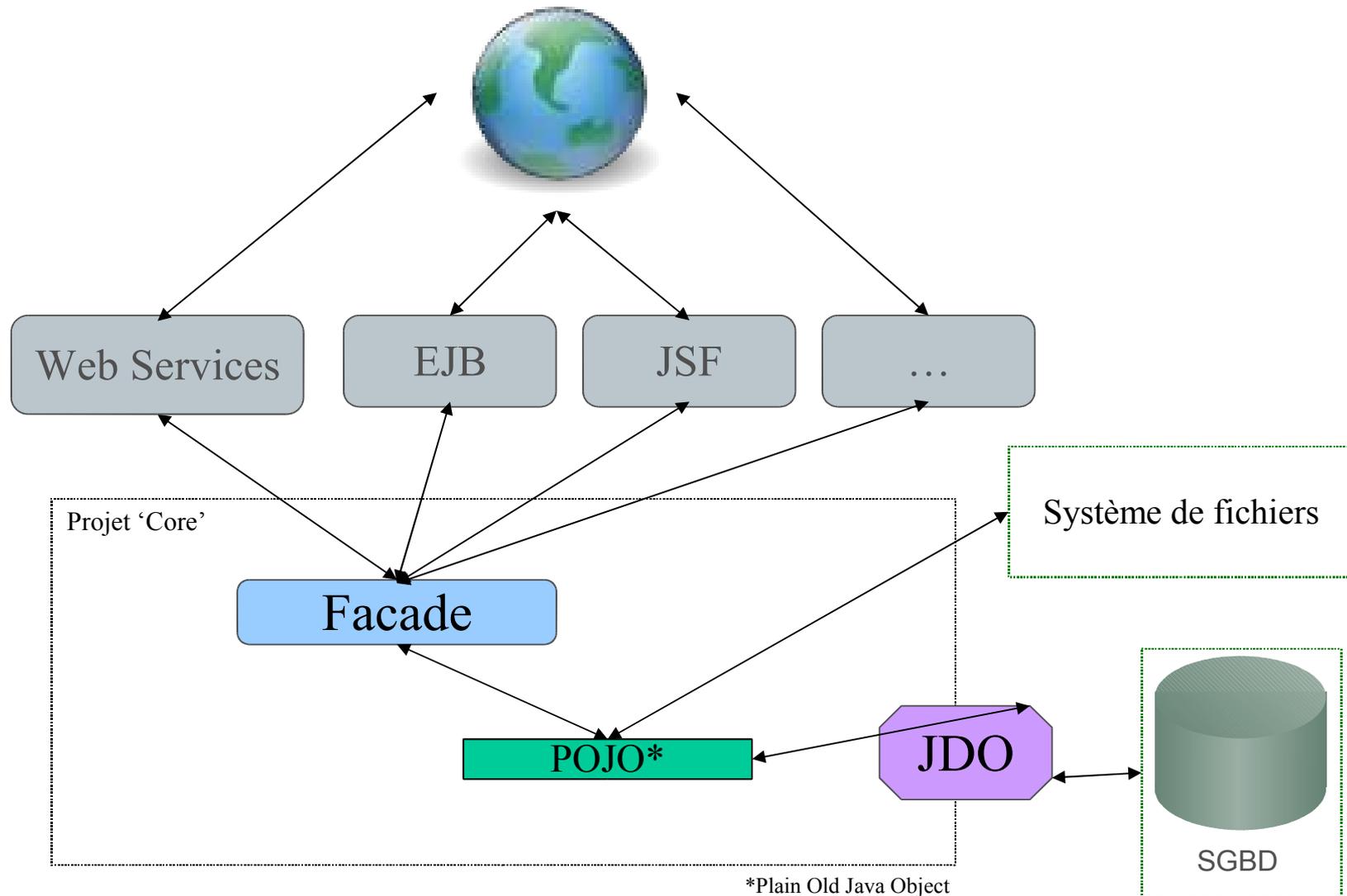
# Schéma conceptuel



# Architecture technique Actuelle



# Adaptation de la plateforme



# Outils

- JDK 1.4 ou plus, AspectJ
- Struts (projet Apache)
- Tomcat
- Kodo ou JPOX
- Oracle ou n'importe quelle BD
- SAXON/FOP
- Lucene
- Jcharts
- OAICat
- Ant

- JUnit
- CVS Adullact
- Maven :
  - Rapport de style
  - Rapport de couverture de code
  - Rapport de mesure de complexité (Metrics)
  - Gestion des compilations

# Propriétés et avantages de la plate-forme

- Architecture logicielle ouverte à base de composants Open Source :
  - Maîtrise en local
  - Libre modification/adaptation aux besoins des utilisateurs
- Utilisation de XML comme format de conservation des documents

# Propriétés et avantages de la plate-forme

- Intégration avec des outils déjà existants :
  - annuaire LDAP
  - plate-forme e-learning
  - logiciel documentaire
  - etc.
- Ouverture du système vers l'extérieur

# CASTORE

---

<http://www.castore.org>

<http://demo.castore.org>

<http://forge.castore.org>

<http://blog.castore.org>