

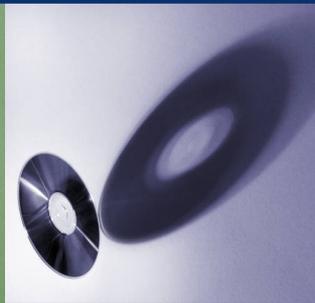


Projet ORI-OAI

Outil de **R**éférencement et d'**I**ndexation

Réseau de portails **OAI**

Rencontres Mondiales du Logiciel Libre 2007
Amiens, 13 juillet 2007





Introduction - contexte

Les fonctions du système

Concepts - principes

Implémentation - architecture

Démonstration

François Jannin
ENSEEIH7 – INP de Toulouse



Introduction



- Croissance de la création numérique institutionnelle (documents scientifiques, pédagogiques, administratifs, ressources documentaires..)
- Mise en place d'une démarche « **Systeme Global d'Information** » à l'échelle d'un établissement
- Organisation des services numériques au moyen de l'**ENT**, s'appuyant sur le SI



- Besoin d'un **référentiel unique** de documents et d'une **indexation professionnelle**
- Demande croissante **d'accessibilité aux ressources numériques** dans différents contextes
- La structuration importante initiée par les projets nationaux (UNT, UNR, SI d'établissement)



- Répondre aux besoins de stockage, de référencement, d'indexation et d'accès aux documents
 - Assurer la valorisation et la gestion pérenne des contenus numériques
 - Développer un SGI en lien avec les ressources numériques d'un référentiel unique de documents
 - Partager les ressources numériques avec d'autres établissements
 - Donner l'accès à distance et selon les droits



- Patrimoine pérenne pouvant se partager
 - Avec les **acteurs** de l'établissement
 - Avec les **partenaires** des réseaux locaux , nationaux et internationaux de l'établissement
 - Avec les **internauts**
 - En respectant les **règles du numérique** et les **standards**
 - Droits d'auteur
 - Interopérabilité des systèmes (OAI / SOAP)
 - Normes d'échange (schémas XML)
 - Vocabulaire commun



Un système de gestion de documents

- Mise en place d'un système de gestion de documents
 - Système de **gestion et de publication** des ressources numériques (gestion des droits, indexation LOM-FR, DC, TEF., indexation classification des UNT)
 - Système **de production** impliquant les acteurs concernés (workflows)
 - Systèmes **de partage** (échange de métadonnées /OAI-PMH)
 - Système **de recherche avancé** (métadonnées, texte intégral), accès thématiques
 - Utilisation de **vocabulaires** provenant de référentiels communs ou spécifiques [issus du SI] à chaque étape du processus de référencement/consultation



Établissements



UNT



Avec le soutien du MENESR





Les fonctions du système



- **Gérer et publier** les documents numériques de l'établissement
 - **Référentiel unique** connecté au SI
 - **Accès thématique** aux ressources
 - Système de **recherche avancée**



- **Partager** les ressources numériques avec d'autres établissements
 - enrichir l'offre numérique
 - augmenter la visibilité des ressources
 - mutualiser les coûts
- **Valoriser** la production scientifique et pédagogique par une indexation de qualité
- **Accéder** aux ressources numériques à **distance** en fonction des **droits d'accès**



- **Ressources d'enseignement**
 - polycopiés numériques, exercices, ressources pédagogiques multimédia (format LOM, LOM-fr)

- **Travaux d'étudiants**
 - rapports de stage
 - synthèses bibliographiques

- **Travaux de recherche**
 - publications, pré-publications, rapports techniques
 - mémoires de Master, mémoires de Thèse (format TEF)

- **Ressources documentaires éditoriales**
 - revues, livres électroniques (format Dublin Core)



Concepts

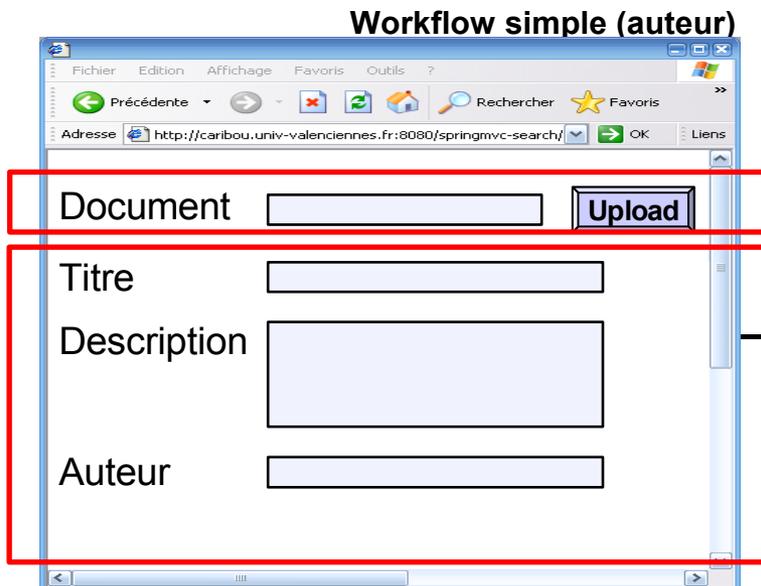


Projet ORI-OAI

principe du référencement



Workflow simple (auteur)



Document Upload

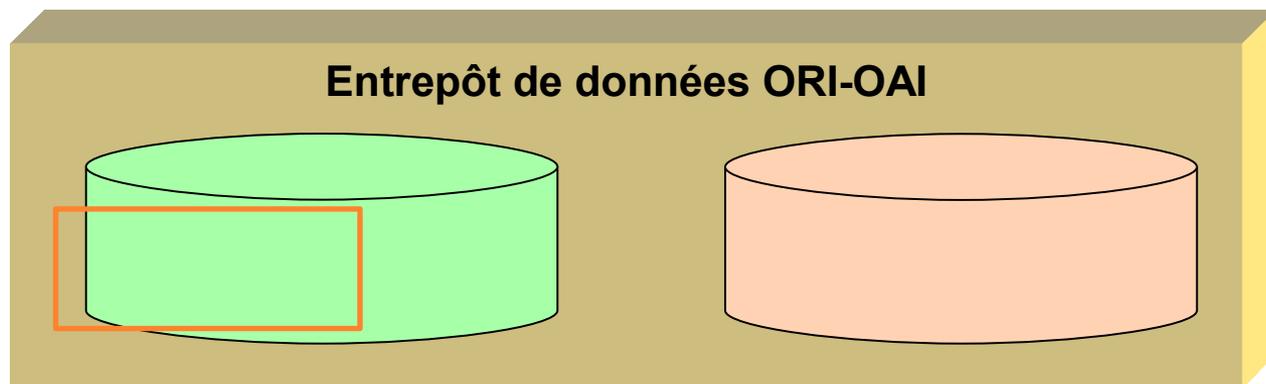
Titre

Description

Auteur

Métadonnées

A publier





Workflow Simple (*étape documentaliste*)



Documentaliste

Workflow simple (documentaliste)

Document

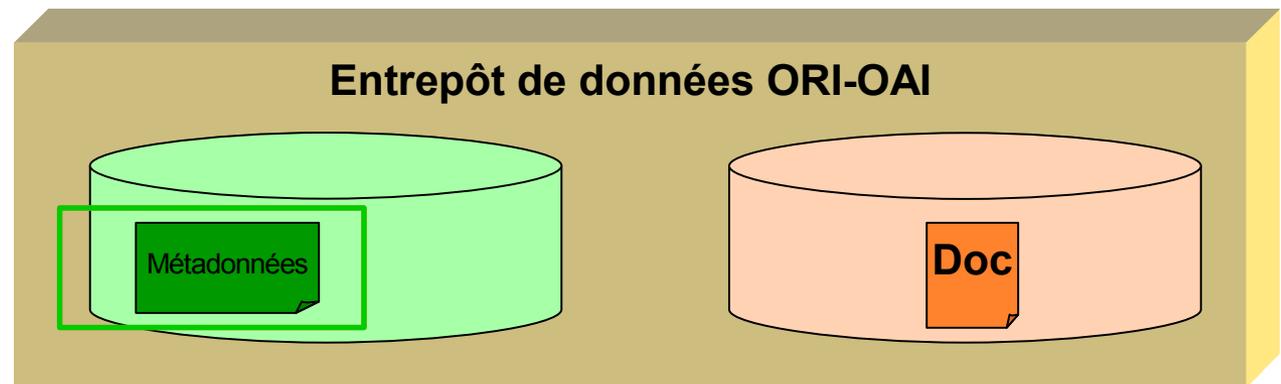
Titre

Description

Auteur

Publier **Rejeter**

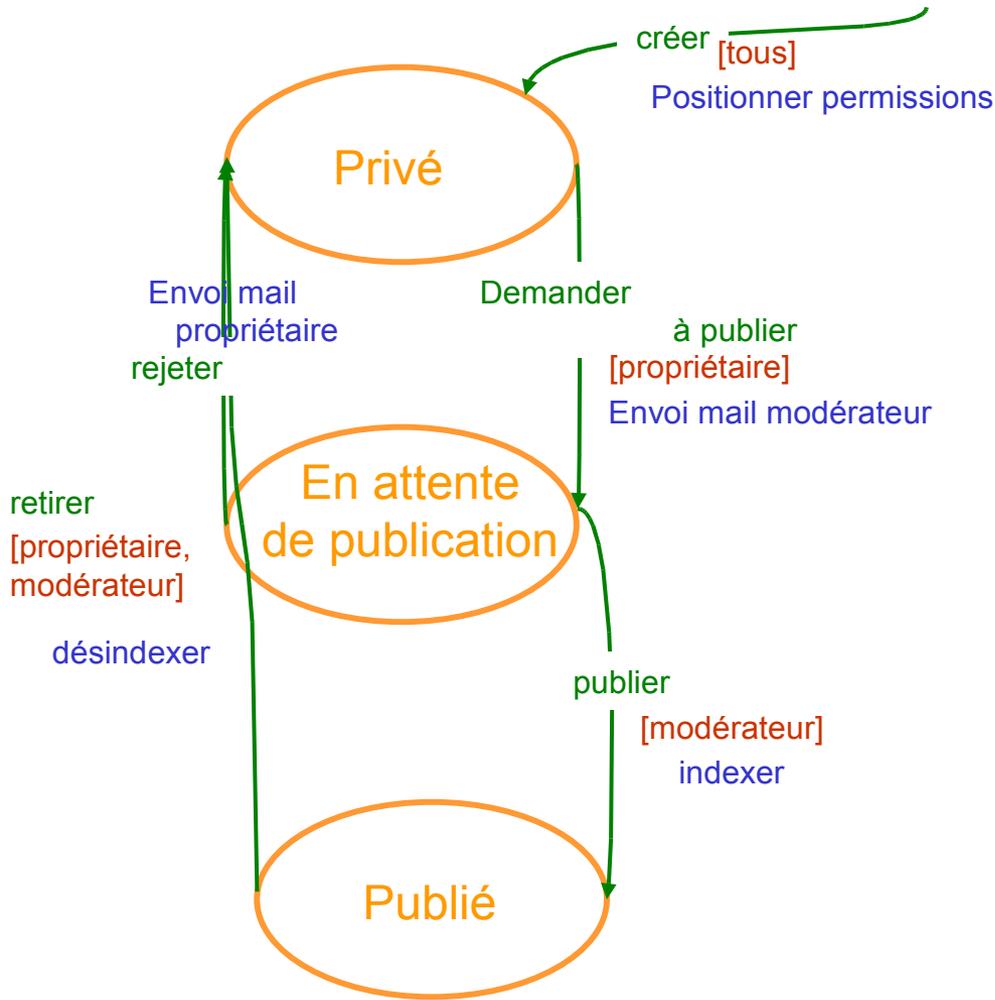
Publié





Principes de ori-oai-workflow

Diagramme Etats / Transitions



Etats
Transitions
Conditions
Actions



- Centralisation et automatisatisation des vocabulaires
 - Répartition en réseau des vocabulaires avec tolérance à la panne (cache intelligent)
 - Utilisés en amont (workflow) et en aval (recherche, OAI)
 - Chaque vocabulaire est partageable par x établissements, UNT/UNR
 - Vocabulaires de référence fermés et **statiques** avec peu de fluctuations (taxonomies, liste de mots-clés fermés)
 - Vocabulaires **dynamiques** générés automatiquement
 - depuis un annuaire LDAP ou Active Directory
 - d'après les valeurs déjà indexées (auteurs, mots-clés libres)
 - d'après des référentiels existants (bases de données...)

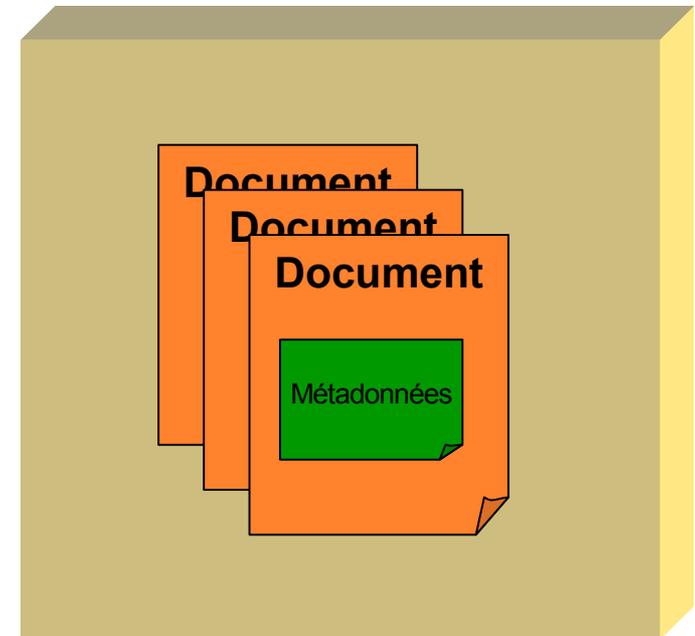


Projet ORI-OAI

principe des échanges

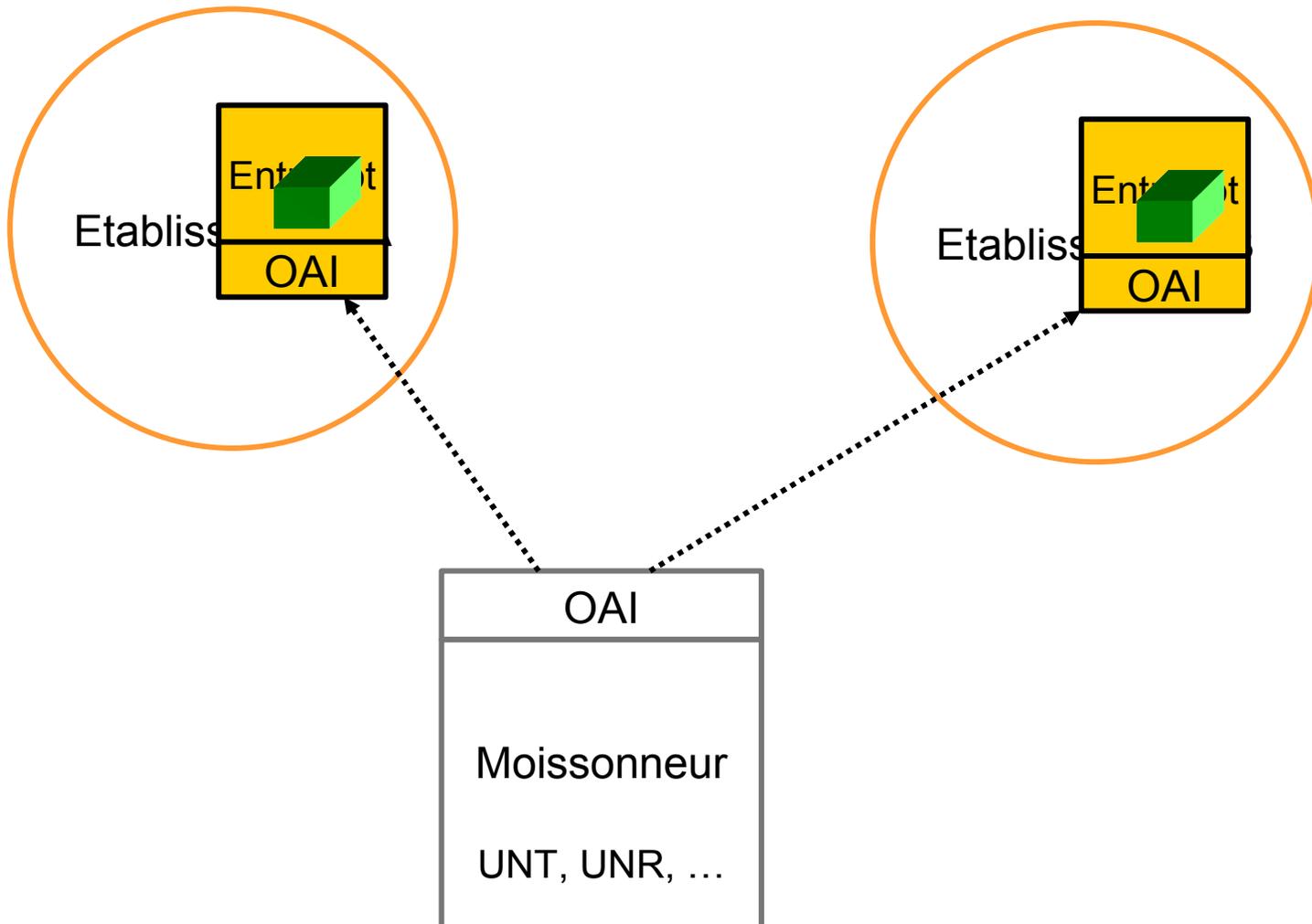
■ Entrepôt ORI-OAI

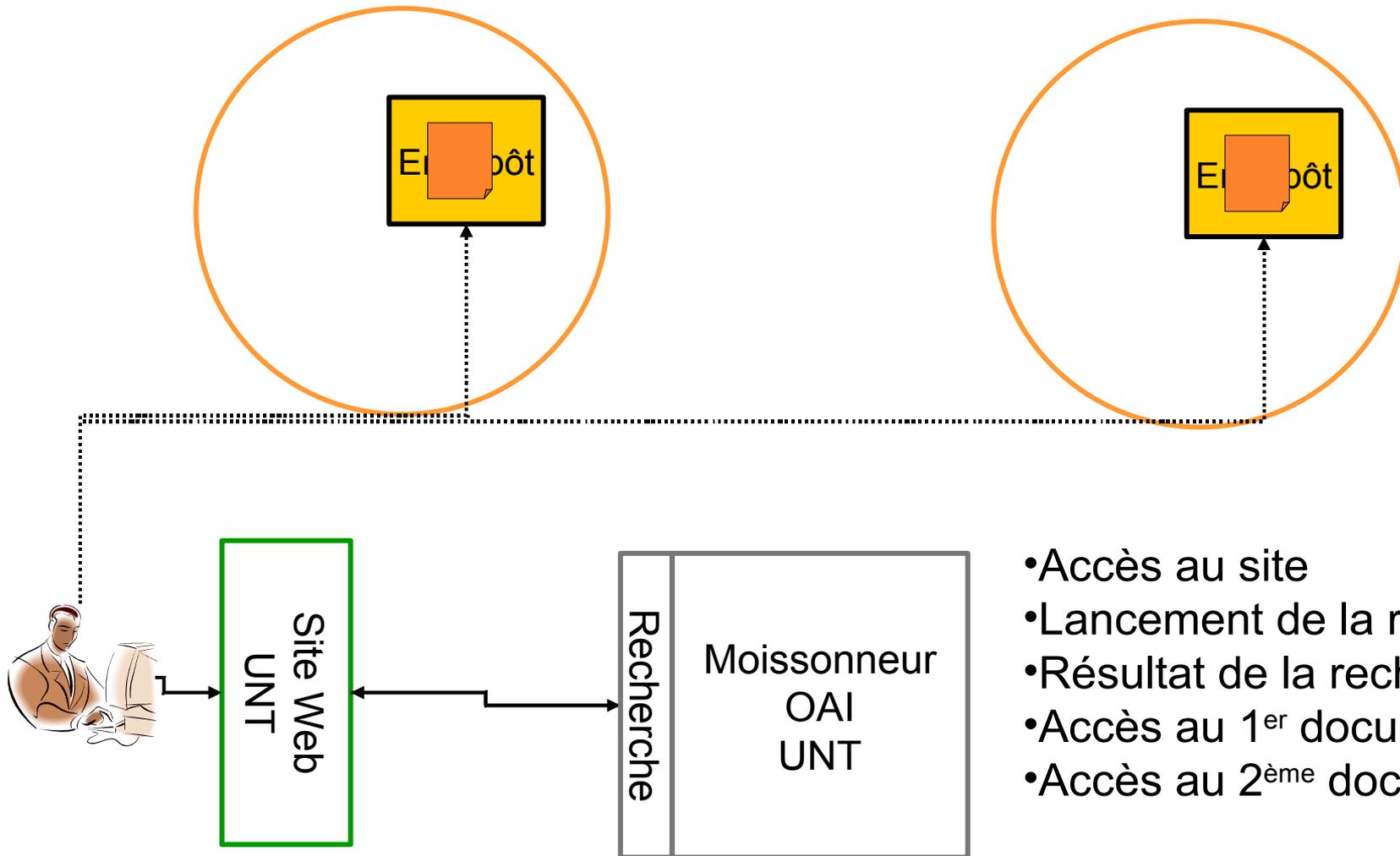
- Stockage des ressources numériques
- Métadonnées sur les documents
 - Dublin Core
 - LOM-fr
 - Etc.





Open Archive Initiative (OAI)

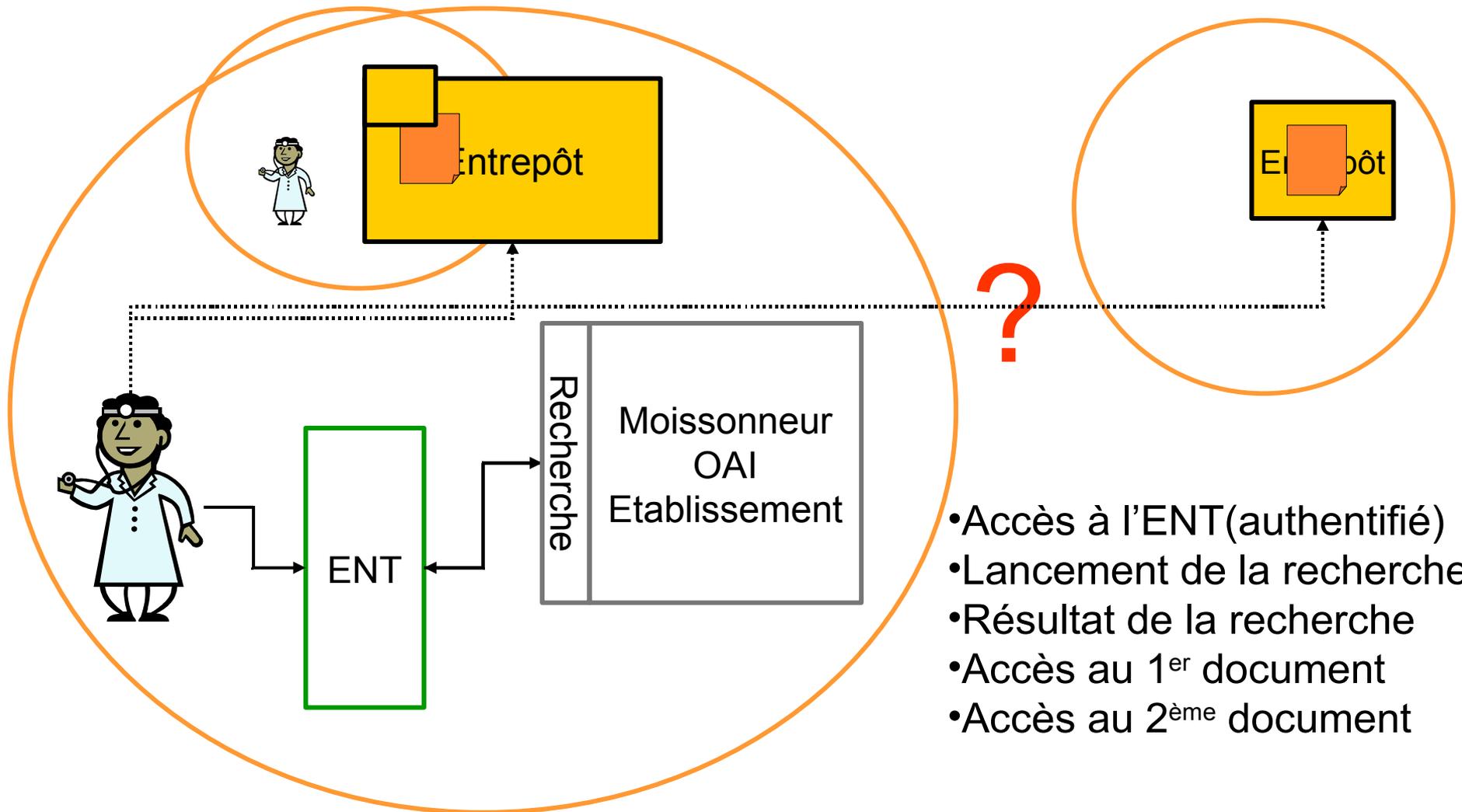




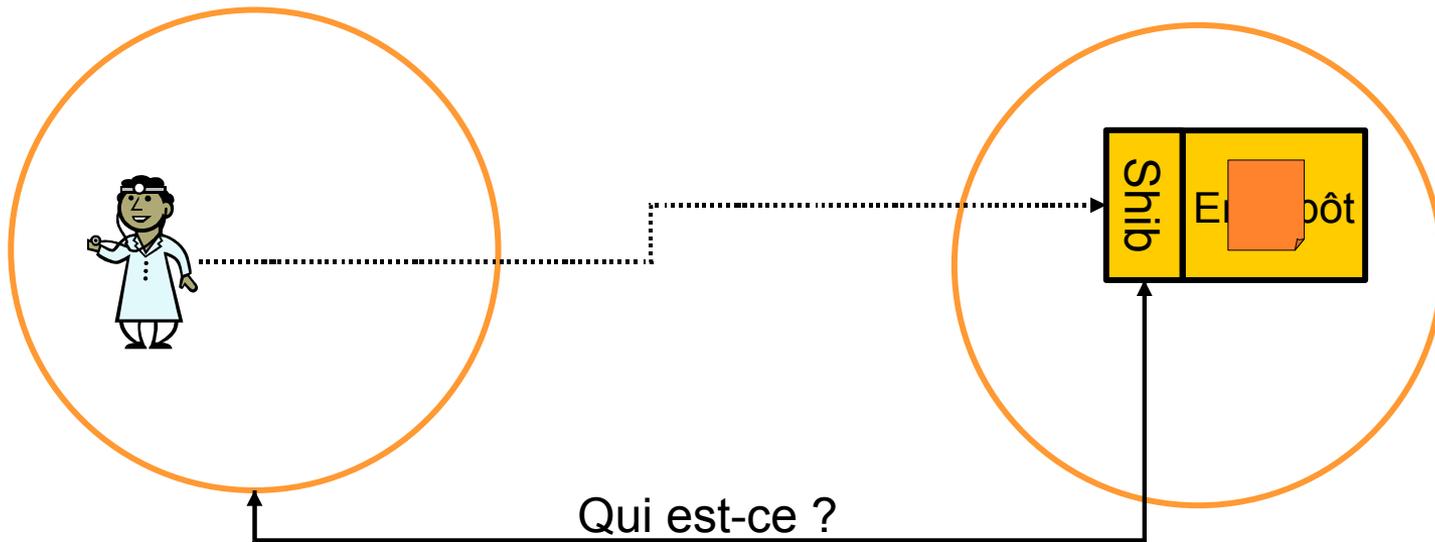
- Accès au site
- Lancement de la recherche
- Résultat de la recherche
- Accès au 1^{er} document
- Accès au 2^{ème} document



OAI pour l'utilisateur d'un établissement

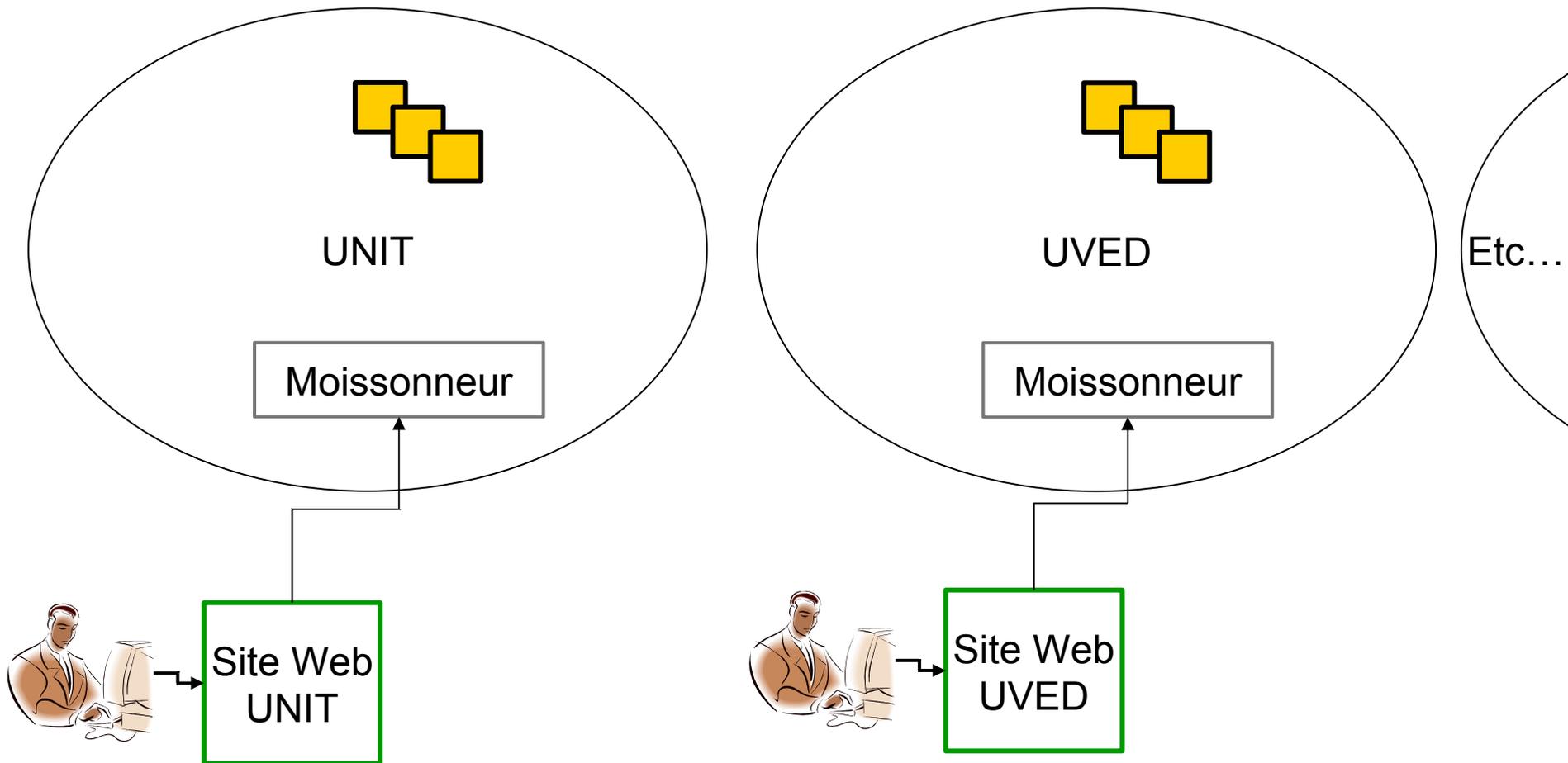


- Accès à l'ENT (authentifié)
- Lancement de la recherche
- Résultat de la recherche
- Accès au 1^{er} document
- Accès au 2^{ème} document



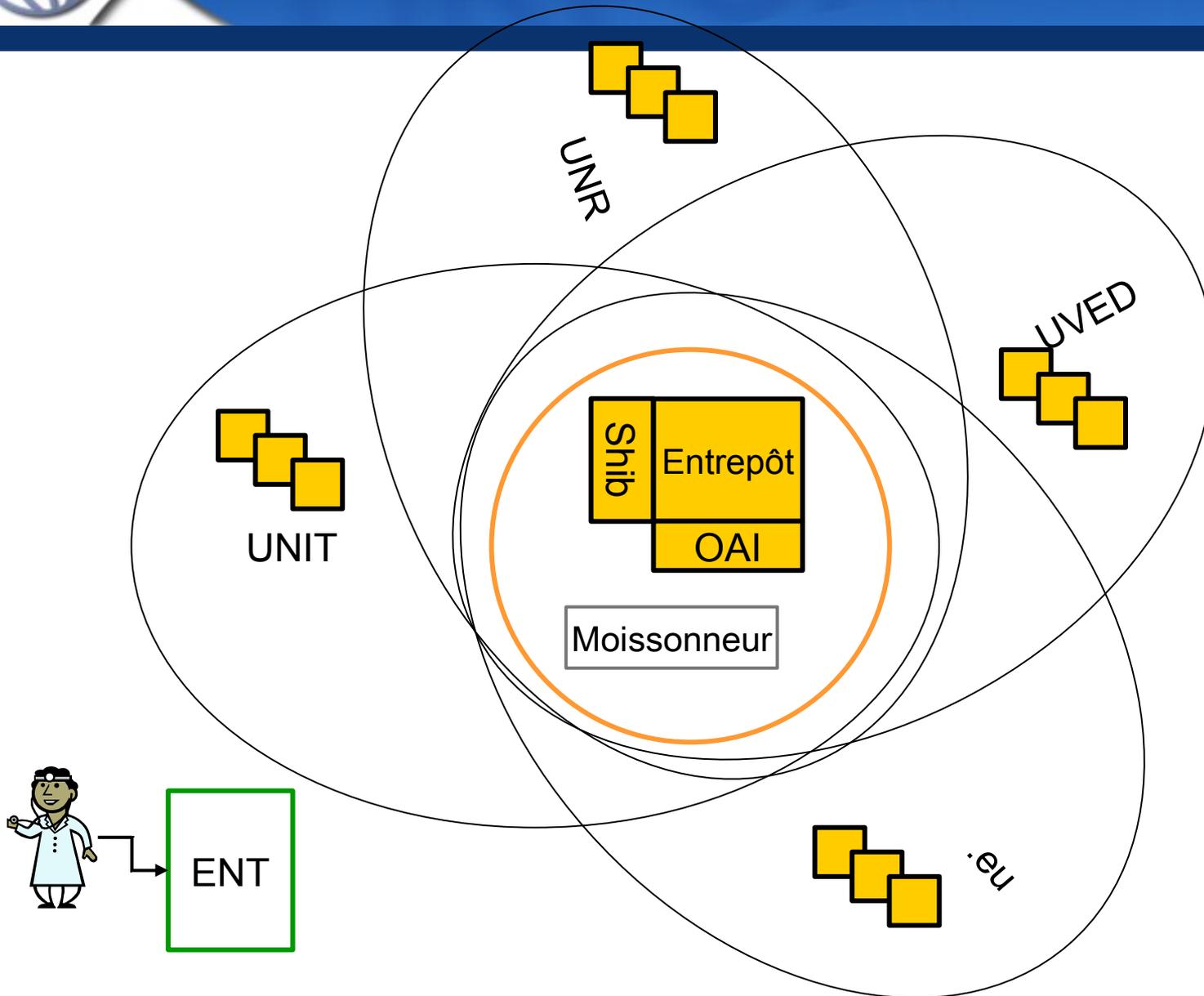
Un étudiant en médecine 2^{ème} année

- Tentative d'accès au document
- Requête à l'établissement « fournisseur d'identités »
- Réponse à l'établissement « fournisseur de services »
- Accès au document





Vue établissement





Implémentation



Projet ORI-OAI

acteurs impliqués



- Équipe de pilotage avec des représentants
 - des établissements
 - des UNT
 - des SCD
- Équipe fonctionnelle
 - 1 coordinatrice
 - 10 membres du groupe « Ressources pédagogiques »
 - 8 membres du groupe « Thèses »
 - 7 membres du groupe « Archives ouvertes »



- Équipe technique
 - 1 coordinateur
 - 1 conseiller technique
 - 5 développeurs
 - 2 Université de Rennes 1
 - 1 INP Toulouse - ENSEEIHT
 - 2 Université de Valenciennes



Projet ORI-OAI

développement



- ORI-OAI est composé de 7 modules
 - Composants interopérables avec des outils existants
 - Utilisation de standards comme OAI, XML-RPC, SOAP, WEBDAV
 - Différentes versions des composants
 - Portlet à intégrer dans un moteur type ENT(JSR 168)
 - Standalone: application indépendante
 - Plusieurs distributions prévues
 - Les modules peuvent être associés les uns aux autres suivant le cas d'utilisation (UNT, UNR, établissement, ...)



- Java J2EE
- Spring : Inversion de Contrôle, injection de dépendance
- Hibernate : Liaison facilitée aux bases relationnelles
- XFire : Web services SOAP simplifiés et efficaces
- AOP/AspectJ : Gestion transversale non-intrusive des transactions et des permissions
- Lucene et LIUS : indexation XML, plein-texte et de formats binaires (PDF, Word, RTF...) par Apache et l'université Laval (Québec)
- OAICat : implémentation Java populaire de OAI-PMH par OCLC
- OSWorkflow : Framework pour workflow très souple et paramétrable
- Orbeon / OPS : Puissant framework de formulaires Web basé sur Xforms
- I18N : localisation des IHM

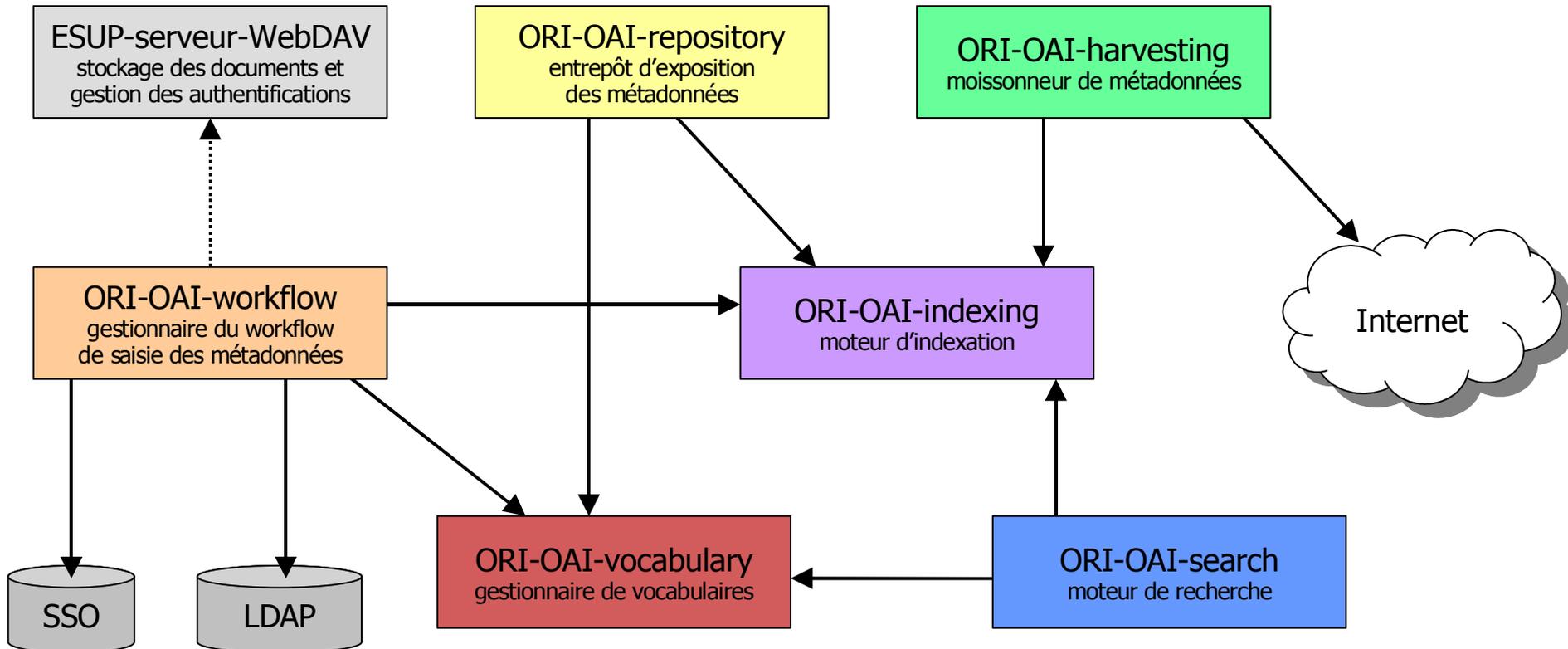


Projet ORI-OAI

architecture

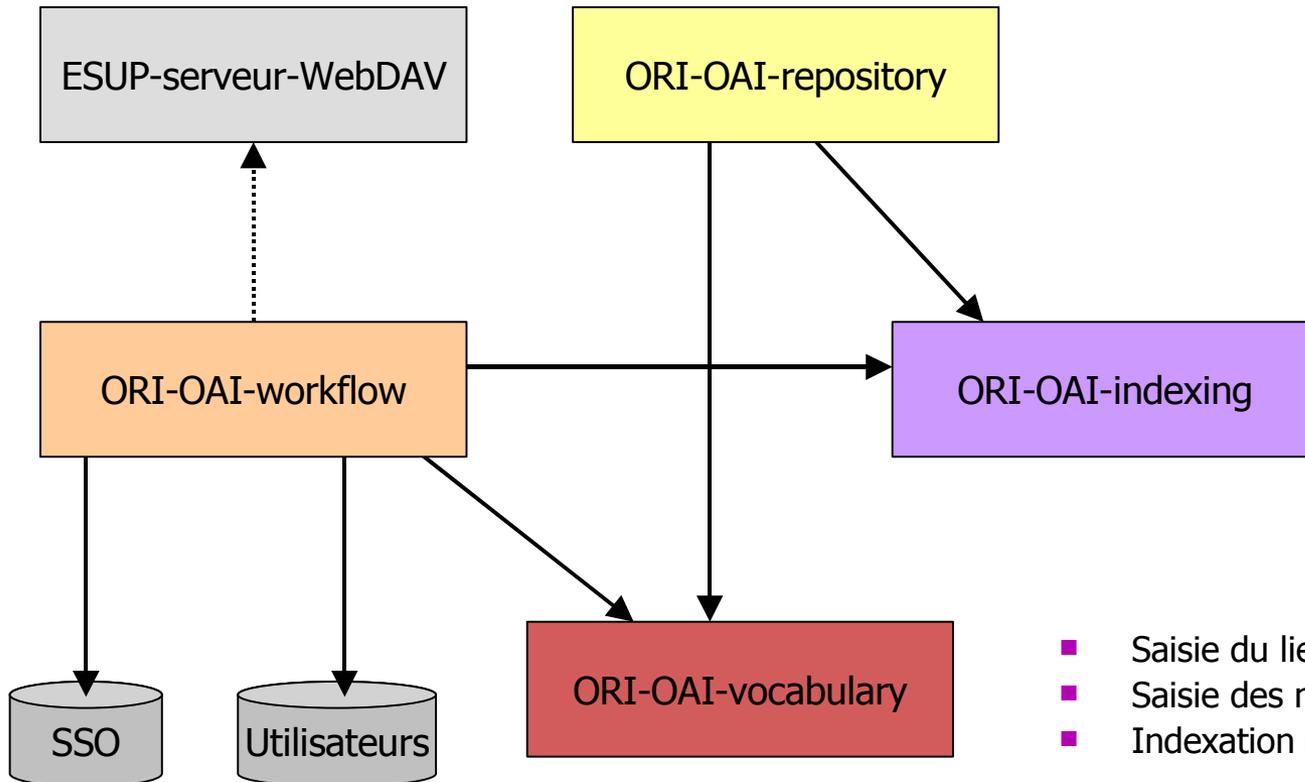


Architecture globale





Gestion des documents locaux

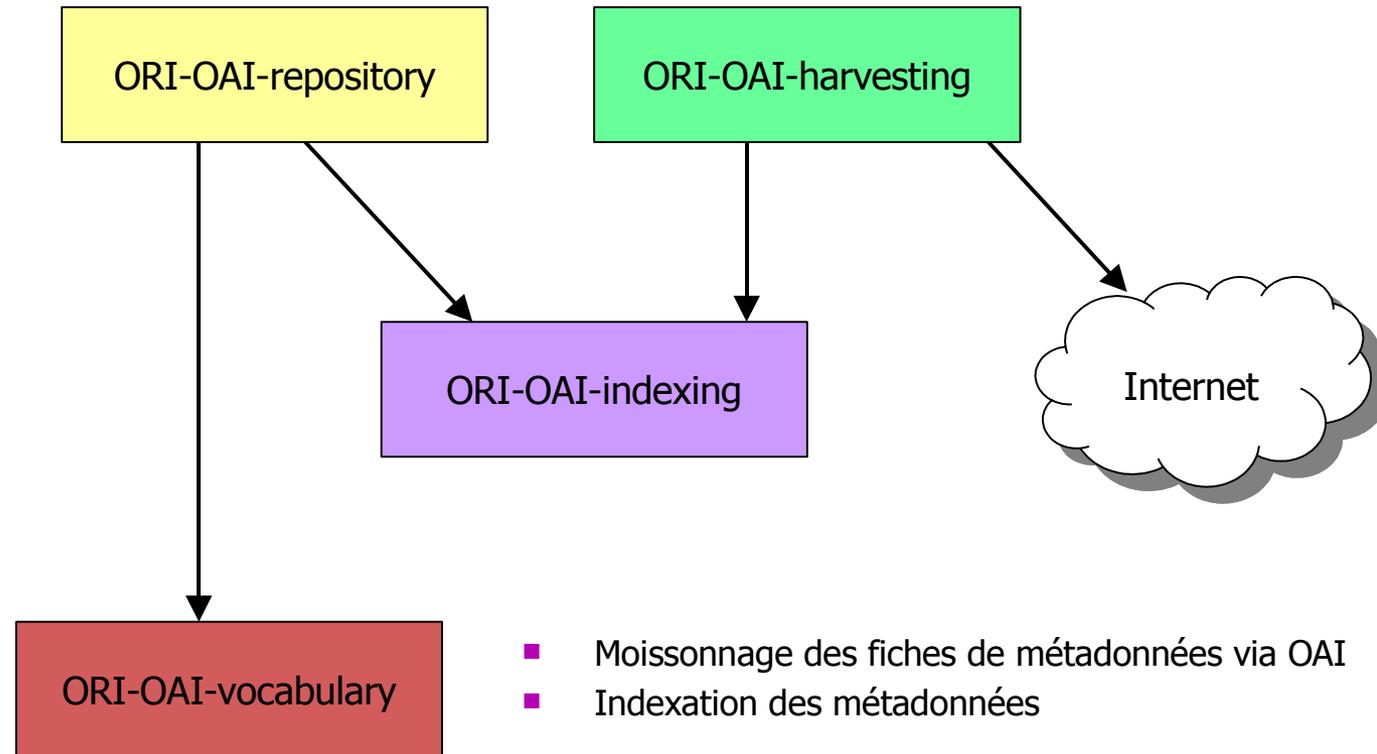


- Saisie du lien vers le(s) document(s)
- Saisie des métadonnées via vocabulaires
- Indexation des métadonnées

- Exposition OAI (optionnelle)
 - Exposition des fiches par requêtes sur l'index
 - Génération de sets OAI via le vocabulaire



Moissonnage par OAI

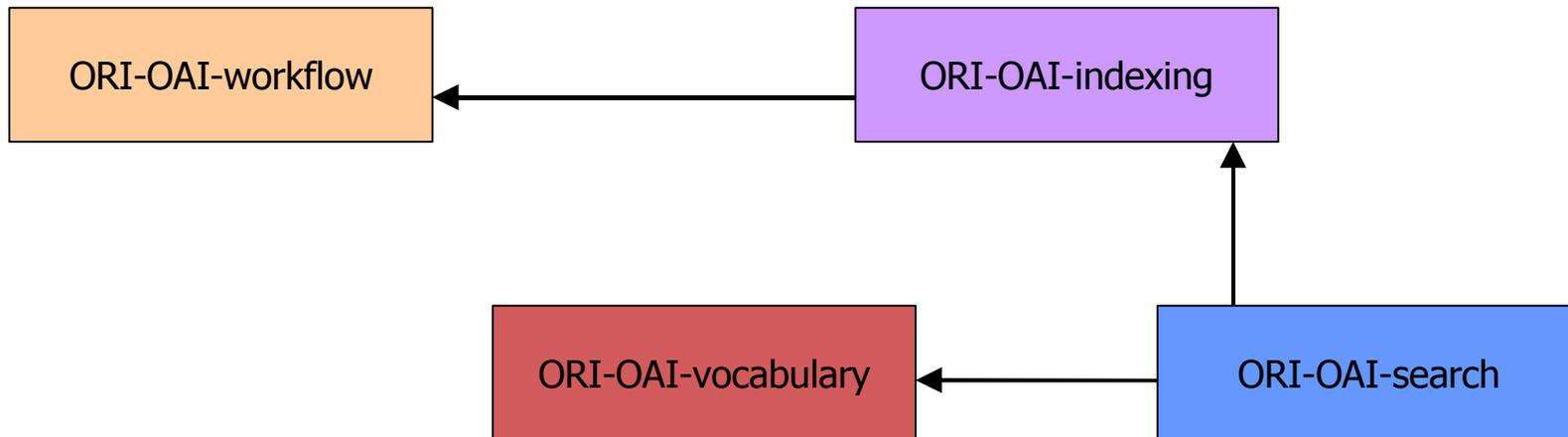


- Moissonnage des fiches de métadonnées via OAI
- Indexation des métadonnées
- Exposition OAI (optionnelle)
 - Exposition des fiches par requêtes sur l'index
 - Génération de sets OAI via le vocabulaire



Recherche de documents locaux

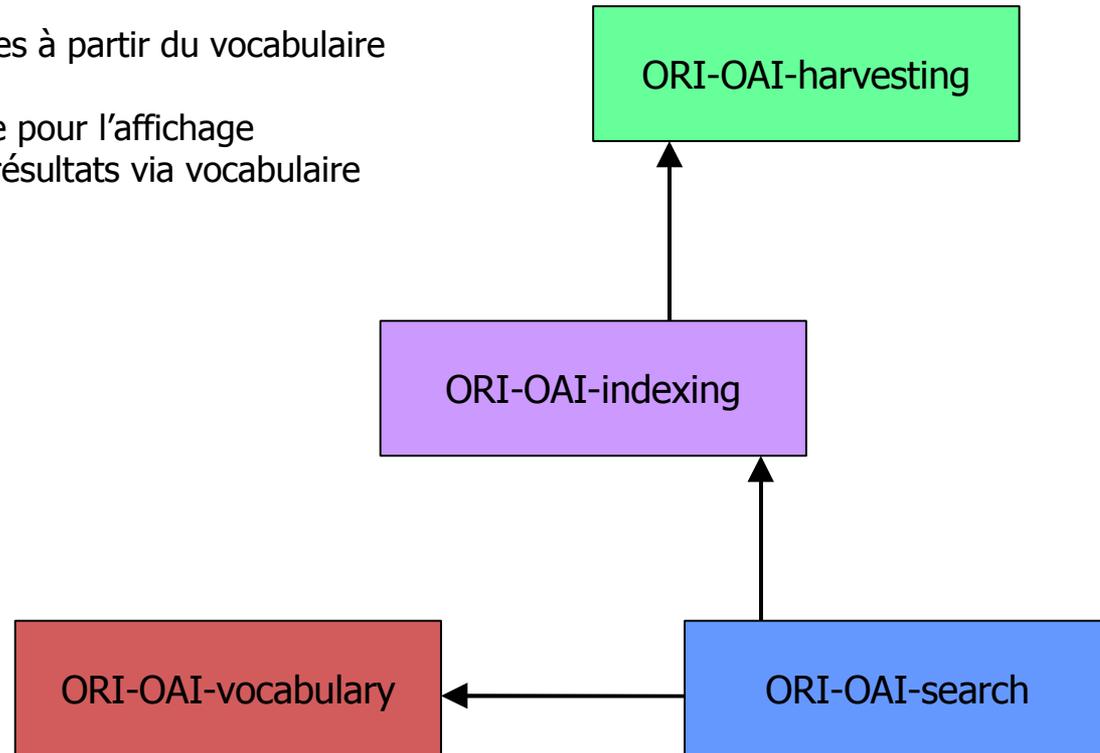
- Génération des interfaces à partir du vocabulaire
- Recherche dans l'index
- Récupération de la fiche pour l'affichage
- Traduction de certains résultats via vocabulaire





Recherche de documents distants

- Génération des interfaces à partir du vocabulaire
- Recherche dans l'index
- Récupération de la fiche pour l'affichage
- Traduction de certains résultats via vocabulaire





Projet ORI-OAI

calendrier



- Version BETA
 - Distribuée fin juin 2007 à 7 établissements
 - Tests juillet et août
 - Retours fin août

- Version 1.0
 - Finalisation septembre 2007
 - Après retours des BETA-testeurs



- Site du projet

- <http://www.ori-oai.org>

- Sources

- http://sourcesup.cru.fr/softwaremap/trove_list.php?form_cat=318

- Contacts

- Liste de diffusion publique **ori-oai-utilisateurs**
<http://listes.univ-rennes1.fr/wws/subscribe/ori-oai-utilisateurs>



Démonstrations



- **Dépôt d'un document LOM, workflow :**
<http://cas.enseeiht.fr/ori-oai-workflow>
- **Recherche thématique & avancée**
<http://cas.enseeiht.fr/ori-oai-search>
- **Recherche relookée**
<http://cas.enseeiht.fr/moodle-web>
- **Moissonneur**
moissonneur ENSEEIHT



Merci !

francois.jannin@enseeiht.fr

Ce travail est mis à disposition sous une licence Creative Commons

Vous êtes libres

De reproduire, distribuer et communiquer cette création au public

De modifier cette création



Cette création est mise à disposition selon le Contrat Paternité-NonCommercial-ShareAlike 2.5
disponible en ligne <http://creativecommons.org/licenses/by-nc-sa/2.5/>

■ Remarque : Les transparents présentés ici ont été réalisés par :

- Rosa-Maria Gomez (INSA de Lyon)
- Yohan Colmant (Université de Valenciennes)
- Raymond Bourges (Université de Rennes 1)
- Monique Joly (INSA de Lyon)
- Alain Mayeur (Université de Valenciennes)
- François Jannin (INP Toulouse – ENSEEIHT)