



## Projet ORI-OAI

Outil de **R**éférencement et d'**I**ndexation  
Réseau de portails **OAI**

# Moteur d'indexation ORI-OAI-indexing

Formation technique niveau II :  
Configuration avancée (comment naviguer dans les fichiers de configuration)





# Sommaire

**Points d'entrée principaux**

**Le crawler web**

## Points d'entrée principaux

- Voir doc « Navigation dans ... configuration »
- properties/
  - ../configIndexing.xml : configuration générale
  - ../liusConfig.xml : déclaration des MD à indexer
- **+ d'infos sur [ori-oai.org](http://ori-oai.org)**
  - Installation / Les modules / indexing - Documentation / page Installation / Configurations avancées

## Le crawler web

### ■ Idée :

- Permettre d'enrichir la fiche de MD avec une indexation plein texte
- ❶ contenu textuel dans MD "full Text"
- ❷ indexation de cette MD
- ➔ Requête possible sur cette MD depuis le moteur de recherche

- via properties/configIndexing.xml
- **<scheduleCrawler>**
  - Planification du crawler
- **<!-- Crawler web -->**
  - <repository ... > : entrepôts à crawler
  - <doNotCrawl> : entrepôts à ne pas crawler
- **+ d'infos sur ori-oai.org**
  - Installation / Les modules / indexing - Documentation / page Installation / Configurations avancées / Système de crawling

- Copyright © 2009 Consortium ORI-OAI
  - Ce document peut être copié et distribué dans son intégralité, sans modification, retrait ou ajout.
  - Tout usage commercial est interdit.
  - L'utilisation de ce document dans un cadre de formation collective est soumise à l'approbation explicite et préalable de ses auteurs.