

Diplôme national de master

Domaine – Sciences Humaines et Sociales

Mention – Information et Communication

Spécialité – Sciences de l'Information et des Bibliothèques

# **Évolution d'un moteur de recherche fédérée dans un contexte multi- institutionnel : la Bibliothèque Ouverte Montpellier - Languedoc- Roussillon**

**Marion Grand-Démery**

Sous la direction de Chantal Salson, Chef de projets web et systèmes  
d'information Agropolis International, Hanka Hensens, Documentaliste IRD  
Et Thierry Lafouge  
Enseignant en Sciences de l'Information Université Lyon 1



## **Remerciements**

*Mes remerciements vont tout d'abord à Chantal Salson et Hanka Hensens, qui m'ont soutenue et aidée tout au long de mon stage. Ils vont aussi bien sûr à l'ensemble des membres du groupe projet, qui ont suivi avec intérêt chaque étape de mon travail. Je tiens également à remercier le personnel de la Bibliothèque de SupAgro Montpellier pour m'avoir accueillie lors de mes déplacements sur le campus.*

*Je souhaite aussi remercier Jean Cerda, du service informatique d'Agropolis International, pour son aide et sa disponibilité, ainsi que les documentalistes et les chercheurs des différents organismes de recherche, pour leur participation aux enquêtes et aux entretiens.*

*Merci enfin à mon tuteur, Thierry Lafouge, pour l'intérêt qu'il a su porter à ma mission, pour son aide et ses encouragements.*

### **Résumé :**

La Commission Information Scientifique et Technique (CIST) de l'association Agropolis International, qui rassemble les documentalistes des centres de documentation des établissements de recherche et d'enseignement supérieur de la région Languedoc-Roussillon liés aux domaines de l'agriculture, de la biodiversité et de l'environnement, souhaite faire évoluer l'interface de recherche fédérée de la Bibliothèque Ouverte Montpellier – Languedoc-Roussillon en migrant vers un nouvel outil.

Le rapport présente l'environnement dans lequel s'est déroulé le stage, à travers la description du cadre collaboratif du projet d'une part, et la description du contexte documentaire et technique d'autre part. Il décrit également les différentes étapes que j'ai pu mener durant quatre mois pour répondre aux objectifs de la CIST : état des lieux des ressources à interfacier, étude des besoins des utilisateurs, veille sur les moteurs de recherche fédérée et enfin, proposition de solutions.

### *Descripteurs :*

recherche fédérée – étude de marché – Agropolis International – information scientifique et technique – valorisation de l'information – portail de recherche documentaire

### **Abstract :**

The Committee for Scientific and Technical Information of the association Agropolis International, which federate librarians of research and higher education institutes working in the field of agriculture, biodiversity and environment in Languedoc-Roussillon, aims at setting up a new federated search interface to replace the current Bibliothèque Ouverte Montpellier – Languedoc-Roussillon.

The present paper first describes the project background and the technical context. It also gives the different steps and tasks I have carried out during four months in order to answer the CSTI's objectives : taking stock of the data which may be included, needs analysis, suggestion of metasearch tools.

### *Keywords :*

federated search – market study – Agropolis International – scientific and technical information – enhancement of information – information retrieval portal



Cette création est mise à disposition selon le Contrat :

**Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0 France**

disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/2.0/fr/> ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.



# Sommaire

SIGLES ET ABRÉVIATIONS.....	9
INTRODUCTION.....	11
<b>PARTIE 1 – PRÉSENTATION DE LA BIBLIOTHÈQUE OUVERTE MONTPELLIER - LANGUEDOC-ROUSSILLON .....</b>	<b>13</b>
<b>1.1 Un projet collaboratif .....</b>	<b>13</b>
1.1.1 Contexte et objectifs.....	13
1.1.2 Historique.....	13
<b>1.2 La BOMLR aujourd'hui.....</b>	<b>14</b>
1.2.1 Les ressources interfacées.....	14
1.2.2 Le métamoteur AskOnce.....	14
1.2.3 Les inconvénients.....	15
<b>PARTIE 2 – RÉAMÉNAGER LA BOMLR : ANALYSES PRÉALABLES.....</b>	<b>17</b>
<b>2.1 État des lieux des bases de données à interfacier.....</b>	<b>17</b>
2.1.1 Un tour d'horizon des différents établissements membres d'Agropolis.....	17
2.1.2 Des ressources hétérogènes.....	18
<b>2.2 Étude des besoins.....</b>	<b>25</b>
2.2.1 Les publics cibles.....	25
2.2.2 Enquêtes et entretiens.....	25
2.2.2.1 Entretiens avec les documentalistes.....	25
2.2.2.2 Résultats du questionnaire.....	26
2.2.3 De nouveaux objectifs pour la BOMLR.....	27
<b>PARTIE 3 – RÉAMÉNAGER LA BOMLR : ÉTUDE ET PROPOSITION DE SOLUTIONS.....</b>	<b>29</b>
<b>3.1 Les principaux mécanismes de la recherche fédérée.....</b>	<b>29</b>
3.1.1 La technologie des connecteurs.....	29
3.1.2 La technologie du « moissonnage ».....	30
<b>3.2 Panorama des solutions de recherche.....</b>	<b>30</b>
3.2.1 Les moteurs de recherche fédérée.....	31
3.2.2 Les moteurs de recherche intégrée.....	31
3.2.3 Les moteurs hybrides.....	31
3.2.4 Les outils de découverte.....	32
<b>3.3 Solutions libres et propriétaires : quelles différences ?.....</b>	<b>32</b>
<b>3.4 Établissement des critères pour l'évaluation comparative des solutions.....</b>	<b>32</b>
3.4.1 La nécessité d'un outil mixte : un premier critère de sélection .....	32
3.4.2 Les fonctionnalités de recherche et d'affichage des résultats attendues.....	33
<b>3.5 Tableau comparatif des solutions.....</b>	<b>34</b>
<b>3.6 Exploration des solutions .....</b>	<b>39</b>
3.6.1 Les solutions retenues.....	39
3.6.2 Exploration de la sélection : premiers résultats.....	39
3.6.3 Exploration de la sélection : perspectives.....	40
CONCLUSION.....	43
BIBLIOGRAPHIE ET WEBOGRAPHIE.....	45
TABLE DES ANNEXES.....	47
GLOSSAIRE.....	95

**TABLE DES MATIÈRES.....97**



## ***Sigles et abréviations***

ABES : Agence Bibliographique de l'Enseignement Supérieur

AMAP : botAnique et bioinforMatique de l'Architecture des Plantes

BIU : Bibliothèque Inter-Universitaire

BOMLR : Bibliothèque Ouverte Montpellier – Languedoc-Roussillon

BPI : Bibliothèque Publique d'Information

BPMB : Biochimie et Physiologie Moléculaire des Plantes

BRGM : Bureau de Recherche Géologique Minière

CBGP : Centre de Biologie et de Gestion des Populations

CEFE : Centre d'Écologie Fonctionnelle et Évolutive

CEMAGREF : Centre national du Machinisme Agricole, du Génie Rural, des Eaux et des Forêts (devenu IRSTEA)

CIRAD : Centre International de la Recherche Agronomique pour le Développement

CIST : Commission Information Scientifique et Technique

CMS : Content Management System

CRBM : Centre de Recherche de Biochimie Macromoléculaire

HTTP : HyperText Transfer Protocol

ICRA : Centre International pour la Recherche Agricole

INRA : Institut National de la Recherche Agronomique

INSERM : Institut National de la Santé et de la Recherche Médicale

IRSTEA : Institut de Recherche pour l'Ingénierie de l'Agriculture et de l'Environnement

IRC : Institut des Régions Chaudes

IRD : Institut de Recherche pour le Développement

IST : Information Scientifique et Technique

LAMETA : Laboratoire Montpellierain d'Économie Théorique et Appliquée

LBE : Laboratoire de Biotechnologie de l'Environnement

MOISA : Marchés, Organisations, Institutions et Stratégies d'Acteurs

OAI-PMH : Open Archives Initiative Protocol for Metadata Harvesting

OPAC : Online Public Access Catalog

RePEc : Research Papers in Economics

SaaS : Software as a service

SCD : Service Commun de Documentation

SGBD : Système de Gestion de Base de Données

SICD : Service Interétablissements de Coopération Documentaire

SIGB : Système Intégré de Gestion de Bibliothèque

SIST : Scientific and Technical Information System

SRW/U : Search Retrieve Web Service / Search Retrieve via URL

SSII : Société de Services en Ingénierie Informatique

SUDOC : Système Universitaire de Documentation

XML : Extensible Markup Language

# Introduction

---

La Commission Information Scientifique et Technique (CIST) de l'association Agropolis International<sup>1</sup> rassemble les documentalistes des centres de documentation des établissements de recherche et d'enseignement supérieur de la région Languedoc-Roussillon liés aux domaines de l'agriculture, de la biodiversité et de l'environnement. Afin de valoriser l'Information Scientifique et Technique de la région Languedoc-Roussillon, elle s'est lancée dans un ambitieux projet : la Bibliothèque Ouverte Montpellier – Languedoc-Roussillon, un portail de recherche fédérée<sup>2</sup> permettant aux usagers d'accéder en ligne aux bases de données des établissements membres via une interface unique de recherche. La mise en place de cette interface à la fin de l'année 2003 facilite la recherche documentaire des différents publics - chercheurs, étudiants et documentalistes – et augmente la visibilité des différents fonds sollicités, parfois méconnus des publics, en permettant la localisation des données.

Cependant, cet outil est devenu peu à peu obsolète. En effet, le moteur de recherche fédérée utilisé n'a pas bénéficié de mises à jour depuis plusieurs années, et les technologies de recherche, tout comme l'interface, sont désormais dépassées.

Afin de relancer ce service, le groupe de travail sur la BOMLR, constitué d'une partie des membres de la CIST, a décidé de proposer un stage dédié exclusivement à ce projet. Ayant postulé avec succès à cette offre, j'ai donc travaillé pendant quatre mois (du 15 février au 15 juin 2012) sur les perspectives d'évolution de la BOMLR.

Contribuer à un tel projet implique dans un premier temps de prendre connaissance de son histoire, de ses enjeux, de son environnement (partie 1). Dans un second temps, il s'agit d'engager la démarche que suppose la mise en place d'une nouvelle interface : définir les caractéristiques documentaires et techniques des bases de données à interroger, puis établir les attentes des différents publics (partie 2). Ces éléments constituent des étapes essentielles avant d'amorcer la véritable étude des solutions (partie 3), l'objectif majeur du projet étant de proposer un outil évolutif, qui permette de répondre au mieux aux besoins des usagers, actuels et à venir.

---

<sup>1</sup> Se reporter aux plaquettes d'information situées en annexe du document p.50-53

<sup>2</sup> Se reporter au glossaire en p.96



# **Partie 1 – Présentation de la Bibliothèque Ouverte Montpellier - Languedoc-Roussillon**

---

## **1.1 UN PROJET COLLABORATIF**

### **1.1.1 Contexte et objectifs**

L'accès à l'Information Scientifique et Technique (IST) en Languedoc-Roussillon est un enjeu essentiel. Il existe une centaine de services d'information et de documentation spécialisés dans la région. Plus particulièrement, de nombreux établissements membres d'Agropolis International possèdent un centre documentaire ou une bibliothèque IST. Au total, quarante-trois centres de documentation sont réunis dans la Commission Information Scientifique et Technique (CIST), déployée dans le cadre de l'association. Les responsables de ces différents centres ont depuis longtemps fait le souhait de proposer à leurs utilisateurs un accès unique à toutes leurs ressources. Avec le développement d'Internet et des nouvelles technologies, ce projet a pu progressivement prendre forme. Il s'agit en effet de permettre l'interrogation de toutes ces données via une interface unique de recherche accessible en ligne. S'inscrivant dans le sens d'une valorisation de l'information scientifique et technique dans la région, c'est ainsi qu'a pu naître la Bibliothèque Ouverte Montpellier – Languedoc-Roussillon.

### **1.1.2 Historique**

Dès 1999, la commission réalise plusieurs études (état des lieux des ressources à interroger, étude de marché des solutions de recherche fédérée, etc) qui marquent véritablement le début du projet de la Bibliothèque Ouverte Montpellier – Languedoc-Roussillon (BOMLR). Ces études aboutissent à l'élaboration d'un cahier des charges, mettant en évidence les exigences et les contraintes techniques et fonctionnelles du projet. Finalement, la solution retenue pour la mise en place de l'interface unique de recherche de la BOMLR est le logiciel propriétaire AskOnce, alors développé par la société Xerox, et connu désormais sous le nom de ECIS Documentum. Au-delà d'un travail collectif sur le projet mené par différents membres de la CIST, la BOMLR repose alors sur une collaboration entre trois grands partenaires : le Pôle Universitaire Européen de Montpellier, la Bibliothèque Interuniversitaire de Montpellier et Agropolis International. Dès 2002, l'outil est installé sur le site d'Agropolis, hébergé sur un serveur dédié. Pendant plusieurs mois, des développements pour permettre l'interrogation à distance des bases de données des différents centres de documentation sont effectués. À la fin de l'année 2003, le portail de la Bibliothèque Ouverte est ouvert à son public.

## 1.2 LA BOMLR AUJOURD'HUI

### 1.2.1 Les ressources interfacées

Si, lors de son lancement, la BOMLR proposait l'accès à plus de quarante-cinq fonds documentaires via l'interrogation d'une trentaine de catalogues d'ouvrages et de revues, elle ne permet plus désormais l'interrogation simultanée de toutes ces ressources. En effet, l'outil n'a pu être adapté à toutes les évolutions effectuées ces dernières années sur les différentes bases de données des divers centres de documentation (migration vers un nouvel SIGB<sup>3</sup>, modification des protocoles d'échange de données, changement de serveur, etc). Il faut par ailleurs savoir qu'aucune personne n'est actuellement exclusivement mobilisée pour le suivi et la maintenance de la Bibliothèque Ouverte, ce qui limite, de fait, son évolutivité. Les bases qui ne répondent plus aux interrogations à cause de ces divers changements ne sont donc plus interfacées.<sup>4</sup>

On a ainsi, pour la partie « documents » - à distinguer de la partie « revues » -, une recherche qui ne s'effectue plus que sur les bases de données des six établissements suivants<sup>5</sup> :

- Bioversity International
- La Bibliothèque Universitaire de Perpignan
- Le Centre d'Écologie Fonctionnelle et Évolutive (CEFE)
- Le Centre International de la Recherche Agronomique pour le Développement (CIRAD)
- L'École des Mines d'Alès

Pour la partie « revues » - qui est donc le second accès proposé par la BOMLR<sup>6</sup> -, on ne trouve plus que quatre bases de données interrogeables actuellement :

- Le catalogue collectif des revues papier du réseau Agropolis
- La base « revues » de Montpellier SupAgro/ IRC
- Celle du CEFE
- Celle de l'Observatoire Océanologique de Banyuls

### 1.2.2 Le métamoteur AskOnce

L'accès unique à ces ressources hétérogènes est permis par le métamoteur AskOnce mis en place il y a dix ans. Cet outil, écrit en langage Java et commercialisé à l'origine par la société Xerox, appartient désormais à l'entreprise EMC Corporation, spécialisée dans la gestion de l'information en entreprise. Il s'agit d'un outil de recherche de première génération, qui ne génère pas d'index comme dans le cas des moteurs de recherche intégrée<sup>7</sup>, actuellement majoritaires sur le marché. L'interrogation des ressources avec cette solution ne repose donc pas sur le moissonnage des métadonnées<sup>8</sup>

<sup>3</sup> Système Intégré de Gestion de Bibliothèque

<sup>4</sup> Le détail des bases dysfonctionnelles se trouve en annexe à la page 61

<sup>5</sup> Se reporter aux sigles et abréviations p.9 pour l'ensemble des noms des établissements

<sup>6</sup> Se référer aux captures d'écran de la BOMLR actuelle en annexe du document page 56

<sup>7</sup> Se reporter au glossaire en p.96

<sup>8</sup> Se reporter au glossaire en p.96

des documents, mais sur la technologie des connecteurs<sup>9</sup>. Autrement dit, AskOnce traduit la requête de l'utilisateur dans la syntaxe employée par chaque base de données distante puis recueille la liste des résultats correspondant à la requête saisie qu'il affiche de manière harmonisée sur son interface.<sup>10</sup> Le logiciel était à l'époque le seul à répondre au but recherché : interfacier le plus de ressources diverses et distantes possibles, avec le moins de contraintes possibles pour les bases interfacées.

En plus d'une recherche simple, l'outil propose une recherche avancée qui permet à l'utilisateur de sélectionner des critères de recherche (recherche par auteur, collection, date, titre, recherche plein-texte, etc) et d'effectuer sa requête en utilisant des opérateurs booléens. Des *clusters*, sous-ensembles constitués d'éléments récurrents, et communs aux différentes notices des bases interfacées sont proposés à l'affichage des résultats, et permettent d'affiner la recherche.

### **1.2.3 Les inconvénients**

La BOMLR détient aujourd'hui de nombreux points faibles qui sont sans doute à l'origine d'une baisse importante de sa fréquentation. En effet, en étudiant les statistiques de consultation du portail de recherche fédérée réalisées entre 2004 et 2011, on s'aperçoit que le nombre de visites par mois est passé de 1500 à moins de 250 en moyenne.<sup>11</sup>

Outre le nombre désormais limité de ressources interfacées, les fonctionnalités de recherche et, plus encore, les fonctionnalités d'affichage des résultats offertes par la Bibliothèque Ouverte, restent assez restreintes. Effectivement, peu de filtres sont proposés pour l'affinage des réponses et pour la plupart, ceux-là n'apparaissent pas directement sur la page des résultats. Il faut ainsi cliquer sur une icône « loupe » pour pouvoir accéder aux principaux filtres (sept propositions seulement).

De plus, l'outil ne propose aucun outil de veille (alertes ou flux RSS) qui sont pourtant de plus en plus prisés par les publics aujourd'hui, professionnels ou non.<sup>12</sup>

Par ailleurs, à l'heure des interfaces de recherche simples et épurées « à la Google », l'interface de la BOMLR semble désormais dépassée, peu intuitive et peu ergonomique. On notera par exemple l'absence de lien vers la page d'accueil, ce qui oblige l'utilisateur à se servir des flèches « retour » de la barre du navigateur Internet pour revenir aux pages précédentes.

Enfin, on remarque que l'outil est particulièrement lent à afficher les résultats. Ce délai provient très certainement du fait que le moteur repose uniquement sur un système de recherche fédérée, dépendant du temps de réponse des serveurs de chaque base de données interfacées. Mais il est surtout pénalisé par l'absence de mise à jour du logiciel et des passerelles, due au retrait des partenaires universitaires du projet, et à la faiblesse conséquente du budget et du personnel dédiés, quand la solution propriétaire choisie nécessitait des investissements réguliers de mise à jour, développements et maintenance.

Aujourd'hui, la Commission Information Scientifique et Technique d'Agropolis souhaite donc réaménager cette interface qui ne semble plus répondre aux besoins des usagers et n'est plus satisfaisante. La migration vers un nouveau méta-moteur, pour remplacer l'actuel outil AskOnce, a ainsi été envisagée.

<sup>9</sup> Se reporter au glossaire en p.96

<sup>10</sup> Le mécanisme de la recherche fédérée est expliqué plus précisément dans la partie 3 du rapport

<sup>11</sup> Se reporter à l'annexe p.63

<sup>12</sup> Se reporter aux captures d'écran en annexe p.56

Avant d'engager le travail sur des solutions de recherche alternatives, j'ai effectué des analyses préalables. Chacune de ces analyses a d'ailleurs fait l'objet d'une présentation devant les membres du groupe-projet, réunissant les documentalistes de plusieurs organismes de recherche et d'enseignement supérieur de la région. De manière générale, des réunions pour mesurer l'avancement du projet se sont tenues tous les quinze jours.

Dans un premier temps, un état des lieux des ressources à interfacer a permis de mettre en évidence les caractéristiques documentaires et techniques de chaque réservoir de données à intégrer dans la BOMLR. Dans un second temps, une étude des besoins des usagers actuels et potentiels a permis de cerner les attentes des différents publics en terme de recherche d'information aujourd'hui.

## Partie 2 – Réaménager la BOMLR : analyses préalables

---

### 2.1 ÉTAT DES LIEUX DES BASES DE DONNÉES À INTERFACER

#### 2.1.1 Un tour d'horizon des différents établissements membres d'Agropolis

Lors d'une réunion du groupe-projet qui s'est tenue le 17 février 2012, une première sélection des bases à interfacier a été effectuée, à partir d'un panorama recensant les divers centres de documentation des organismes de recherche et des établissements de l'enseignement supérieur du Languedoc-Roussillon et incluant notamment les membres de l'association Agropolis International.

Le groupe-projet a ainsi défini dix bases à interfacier en priorité, correspondant aux ressources de dix établissements proches des thématiques d'Agropolis et situés en région. Si la liste présentée ci-dessous n'inclut pas les bases actuellement interfacées sur le portail, celles-ci seront toutefois bien prises en compte pour l'état des lieux.

Les bases à interfacier en priorité sont donc les ressources des établissements suivants<sup>13</sup> :

- AgroParisTech
- L'AMAP (botAnique et bioinforMatique de l'Architecture des Plantes)
- Le Centre de Biologie et de Gestion des Populations (CBGP)
- Le Centre International pour la Recherche Agricole (ICRA)
- L'Institut National de Recherche Agronomique (INRA)
- L'Institut National de la Santé et de la Recherche Médicale (INSERM)
- L'Institut de Recherche pour l'Ingénierie de l'Agriculture et de l'Environnement (IRSTEA), incluant les ressources de la Maison de la Télédétection (MTD).
- Le Laboratoire de Biotechnologie de l'Environnement (LBE)
- Pierre Bartoli : Centre de Documentation dont le fonds couvre les thématiques liées aux Unités Mixtes de Recherche Innovation, LAMETA et Moisa<sup>14</sup> ainsi qu'au Département Sciences Economiques, Sociales et de Gestion du campus de Montpellier SupAgro.
- Montpellier SupAgro

---

<sup>13</sup> Se référer aux sigles et abréviations en p.9 pour l'ensemble des noms des établissements

<sup>14</sup> Se référer aux sigles et abréviations en p.9

Trois autres bases ont été définies comme des bases de priorité « P2 », à savoir :

- Le BPMB (Biochimie et Physiologie Moléculaire des Plantes)
- Le Bureau de Recherche Géologique Minière (BRGM)
- Le Centre de Recherche de Biochimie Macromoléculaire (CRBM)

### **2.1.2 Des ressources hétérogènes**

Pour effectuer un état des lieux détaillé des différentes bases, j'ai mené divers entretiens entre le 20 février et le 1er mars 2012 auprès des responsables des centres de documentation des organismes. Il s'agit ici de synthétiser les caractéristiques des ressources et des outils dont dispose chacune des organisations, et recueillies lors de ces entretiens.

Ces bases de données sont gérées par des logiciels de gestion de bibliothèque (SIGB, SGBD<sup>15</sup>, logiciels bibliographiques, etc.) variés, et présentent des structures hétérogènes. Elles utilisent également divers protocoles de communication pour l'échange des données.

Le tableau proposé ci-après synthétise les caractéristiques de toutes ces bases de données.

---

<sup>15</sup> Système de Gestion de Base de Données

Centres de documentation	Logiciels	Ressources	Volumétrie	Protocole de communication <sup>16</sup>	Versement dans d'autres bases du type Hal, SUDOC, etc.	Perspectives d'évolution, projets en cours ou à venir / Autres remarques
AgroParisTech	PMB	Catalogues d'ouvrages, de revues, de cartes, etc. + travaux d'étudiants.  Mémoires téléchargeables en format PDF selon autorisations des commanditaires (pour le master Gestion de l'eau notamment).	300 000 notices environ	Webservice OAI-PMH (mis en place depuis mai 2012)	Signalement des revues dans SUDOC mais pas mis à jour  Participe à l'alimentation de Pastel pour les thèses + au projet Star de l'ABES (dépôt dans <i>theses.fr</i> )	Mise en place d'un dépôt des publications dans HAL  Mise en place d'un versement dans AtoZ pour les revues d'ici juillet 2012
AMAP	AmapThèque (développé par l'AMAP sur une base MySQL/php <sup>17</sup> )	Sur Intranet : Catalogues d'ouvrages, d'articles, de rapports, de mémoires, etc. Pas d'accès aux documents en full-text Sur Internet (accès via le site de l'AMAP) : Photothèque	18 000 notices	Aucun protocole de communication mis en place pour l'instant	Alimentation Hal & Prod-INRA pour les publications des chercheurs	Projet de mise en place d'un protocole pour communiquer avec d'autres bases et notamment pour la récupération de données ISBN.
Bioversity International	CDS/Isis  MySQL	Publications Ouvrages Articles de revues → catalogues de notices avec parfois liens vers documents en full-text <u>Musarama</u> (base photo)	13200 références pour la base Musalit, 200 pour la base icono.	Connexion directe au serveur de base de données MySql (port 3306)	Non.	Changement d'outil d'ici début juillet 2012 : Zotero

<sup>16</sup> Se reporter au glossaire en p.96

<sup>17</sup> Se reporter au glossaire en p.96

Centres de documentation	Logiciels	Ressources	Volumétrie	Protocole de communication	Versement dans d'autres bases du type Hal, SUDOC, etc.	Perspectives d'évolution, projets en cours ou à venir / Autres remarques
BIU Montpellier	Aleph/Flora (GED)	Ouvrages, e-books, revues électroniques, images, thèses, mémoires, documents patrimoniaux, annales  Accès aux <i>full-text</i> selon autorisation (accès public ou réservé)	1 956 618 exemplaires localisés sous 907 970 notices bibliographiques	Z 3950	Dépôt dans le SUDOC, TEL (pour les thèses de Montpellier 3), <i>theses.fr</i> (projet Star de l'ABES)	Mise en place d'une base de signalement pour les publications des chercheurs
BU Perpignan	Absys	Catalogues de la BU accessibles en ligne (thèses, mémoires, ouvrages, etc.)  Documents en <i>full-text</i> en accès réservé.	173 000 notices sur l'OPAC	Z 3950	Signalement de tous les exemplaires dans le SUDOC, sauf pour les mémoires et annales d'examens	Adhésion au projet STAR ( <i>theses.fr</i> = plateforme nationale des thèses mise en place par l'ABES) il y a trois ans, toujours en cours.  Changement de système depuis fin mars : nouveau logiciel Absys-net (version web), nouveau catalogue.
CBGP (Campus de Baillarguet)	Fichiers à usage local, mais pas de catalogue en ligne	Réunit plusieurs petits fonds spécialisés liés à la thématique du centre.  Détient surtout des collections de lames (insectes, etc.)  Liste des dernières publications du CBGP en ligne sur le site  Liste des publications 2011	-	-	-	Le CBGP ne peut pas participer au projet de la BO-MLR pour l'instant.
CEFE-CNRS	Reference Manager  EndNote (pour les publications)	Le catalogue en ligne propose un accès aux notices d'ouvrages, des thèses, mémoires, articles scientifiques, catalogue des revues.  Pas de base de publications en ligne	+ de 30 000 références	-	Dépôt des périodiques en série dans le SUDOC	Projet de migration vers le SIGB PMB .

Centres de documentation	Logiciels	Ressources	Volumétrie	Protocole de communication	Versement dans d'autres bases du type Hal, SUDOC, etc.	Perspectives d'évolution, projets en cours ou à venir / Autres remarques
CIHEAM-IAMM	PMB	Catalogues d'ouvrages, de publications, de revues.  Travaux d'étudiants en full-text en accès libre, le reste en accès réservé.	26 000 références dont 11 000 articles de périodiques	Webservice	Revue papier versées dans SUDOC  Revue électronique versées dans CAIRN	
Cirad	Loris (Ever)	Une seule base Agritrop gère les publications, catalogues d'ouvrages, de revues papier et électroniques.  Très peu de documents en <i>full-text</i> : accès via liens vers les PDF, sur sites éditeurs et Hal.	+ de 240 000 références	Z 3950	Participation au SUDOC, versement dans Hal-Cirad et AtoZ	Projet en cours : changement de logiciel documentaire.  Mise en place du protocole OAI courant 2013.
École des Mines d'Alès	Alexandrie-Kentika	Catalogue d'ouvrages, de revues, de publications, de thèses, de travaux d'étudiants.  Accès électronique à des revues papier réservé (il faut faire la demande des codes d'accès au centre de documentation)	+ de 22 000 références	WebService OAI-PMH (portail commun des Écoles des Mines)	Versement de certaines ressources électroniques dans HAL (peu nombreuses car chercheurs réticents)	Il existe un portail commun rassemblant les ressources des Écoles des Mines par moissonnage de toutes les bases de données des écoles (utilisation du protocole OAI-PMH).  Pas de projet de changement d'outil, SIGB actuel plutôt satisfaisant.
ICRA	-	Accès aux rapports de terrains PDF à partir de 2004 via site web.  Le fonds papier contient beaucoup de documents obsolètes	-	-	-	Pas de catalogue en ligne
INRA	PMB (fusion avec le catalogue de SupAgro)	ProdInra est sa principale ressource : elle recense l'ensemble des productions des chercheurs.	125 000 références pour ProdInra (tous instituts confondus = au niveau national)	Webservice (avec un connecteur SOAP)	Quelques thèses de doctorants sont versées dans TEL, sur la base du volontariat.	L'INRA possède une autre base de ressources : Bel-INRA. Cependant ces ressources restent en accès réservé.

Centres de documentation	Logiciels	Ressources	Volumétrie	Protocole de communication	Versement dans d'autres bases du type Hal, SUDOC, etc.	Perspectives d'évolution, projets en cours ou à venir / Autres remarques
INSERM Montpellier	-	<p><u>BIR</u> (répertoires des différentes unités de recherche)</p> <p>Publications des chercheurs</p> <p><u>BiblioInserm</u> (accès strictement réservé aux structures Inserm référencées dans la BIR)</p>	-	-	HAL-Inserm	-
IRD/ MSE	Cadic Intégrale	<p>Base Horizon :</p> <p>Catalogues des acquisitions, articles, productions des chercheurs.</p> <p>Accès aux documents en <i>full-text</i> libre ou réservé (selon autorisations)</p> <p>Également en ligne :</p> <p>Catalogue de revues sur AtoZ (EBSCO) avec accès au texte intégral sur abonnement en Intranet</p> <p>Base iconographique (Indigo)</p> <p>Catalogue des Éditions de l'IRD</p> <p>Bases de cartes (Sphaera)</p>	<p>131 590 notices dans la base</p> <p>55 000 documents numériques (tous niveaux d'accès confondus)</p>	<p>OAI-PMH</p> <p>Z3950 (à venir sur la nouvelle version)</p>	HAL-IRD (dépôt par les auteurs)	<p>Projet de base audiovisuelle</p> <p>Mise en place de la Z39 50 sur la nouvelle version courant 2012</p>
IRSTEA - (Cemagref) /MTD	Cadic Intégrale	<p>Cemadoc :</p> <p>Base dédiée aux publications des chercheurs + travaux étudiants (mémoires, thèses)</p> <p>Bases documentaires : catalogues des ouvrages, rapports, revues (papier), documents cartographiques, etc.</p>	<p>Pour Montpellier :</p> <p>4997 références pour la base doc.</p> <p>2632 pour la base publications (avec entre 100 et 200 réf en + chaque année).</p>	OAI-PMH	<p>Versement dans Hal et dans Tel</p> <p>Recensement des abonnements IRSTEА dans AtoZ</p>	<p>Projet de mise en ligne de photothèque sur Internet, pour l'instant uniquement accessible en Intranet (phase de test)</p>

Centres de documentation	Logiciels	Ressources	Volumétrie	Protocole de communication	Versement dans d'autres bases du type Hal, SUDOC, etc.	Perspectives d'évolution, projets en cours ou à venir / Autres remarques
LBE Narbonne	EndNote (logiciel bibliographique)	Ressources en base interne uniquement.	2000 références	-	Alimentation de ProdInra pour la production scientifique	Projet de mutualisation des ressources avec INRA – Montpellier SupAgro. Migration sur PMB
Montp. SupAgro	PMB	Catalogues d'ouvrages, de revues, travaux d'étudiants, ressources cartographiques  Accès <i>full-text</i> à certains documents (notamment pour les travaux d'étudiants), en accès libre ou réservé selon autorisations.	81 147 notices (ouvrages, articles, périodiques)	Webservice (avec un connecteur SOAP)	Alimentation de <i>theses.fr</i> et du SUDOC pour le signalement des revues (abonnements en cours)	Interrogation externe depuis la base de SupAgro Connecteurs vers CAIRN
Observ. Océanol. Banyuls	Biblio-tech Symphony (Paris IV BUPMC)	Catalogues d'ouvrages, de revues, de rapports, de thèses  Accès à certains documents en <i>full-text</i> (articles) via des liens vers les sites Éditeurs.		Z 3950	Signalement dans le SUDOC et dans le catalogue international des bibliothèques en sciences aquatiques et marines (IAMSLIC)	
Pierre Bartoli	PMB	Littérature grise, ouvrages, mémoires, thèses, périodiques (papier et électroniques)  Accès à certains documents en <i>full-text</i>	Environ 30 000 références.	Webservice	Alimentation de ProdInra et de RePEc (archives ouvertes <sup>18</sup> en sciences économiques)	Suite à l'acquisition du fonds de l'UMR « Innovation », réaménagement du Centre Pierre Bartoli et intégration des ressources PMB en question.
UM3 Bibliothèques d'UFR	Via BIU (Alexandrie n'existe plus)	-	-	-	-	-

<sup>18</sup> Se reporter au glossaire en p.96



L'état des lieux des bases de données synthétisé dans ce tableau permet de recentrer la sélection des ressources à interfacer, en écartant dans un premier temps celles qui ne peuvent être interrogées à distance. En effet, certaines bases ne sont équipées d'aucun protocole de communication, et d'autres ne proposent pas encore d'accès en ligne, comme celles de l'ICRA ou de l'AMAP.

Suite à cet état des lieux, j'ai décidé de mener une analyse des besoins, en relevant les attentes des différents publics cibles au travers d'enquêtes et d'entretiens. Il est effectivement apparu indispensable de faire le point sur les besoins des utilisateurs actuels et potentiels de la BO-MLR, plus de huit ans après sa première mise en ligne.

## 2.2 ÉTUDE DES BESOINS

Avant de présenter les résultats des questionnaires et des entretiens, il est nécessaire de rappeler quels sont les publics visés par la Bibliothèque Ouverte.

### 2.2.1 Les publics cibles

La BOMLR cible plus particulièrement quatre grands groupes d'utilisateurs :

- Les professionnels de l'information, les documentalistes
- Les chercheurs, les enseignants-chercheurs
- Les doctorants
- Les étudiants

### 2.2.2 Enquêtes et entretiens

Au-delà d'une collecte d'informations sur les bases potentielles à interfacer, les entretiens effectués auprès des professionnels de l'information des divers organismes ont permis de mettre en évidence les attentes liées au réaménagement du portail de la BO-MLR. Ces rencontres ont également été l'occasion de recueillir toutes sortes de propositions et suggestions intéressantes. Toutefois, ces entretiens ne concernant que les documentalistes, un questionnaire à destination de tous les publics (étudiants, chercheurs, doctorants et professionnels de l'information) a été mis en ligne et diffusé auprès des différents organismes de recherche et établissements d'enseignement supérieur de la région impliqués dans le projet.<sup>19</sup> Pour ce faire, j'ai utilisé l'outil de création et de traitement d'enquêtes *LimeSurvey*. Les résultats de ce questionnaire, tout comme ceux des entretiens, ont donné lieu à la rédaction d'un compte-rendu diffusé aux membres du groupe-projet.

Sont donc présentés ci-dessous les principaux points abordés par les documentalistes au cours des entretiens puis une brève synthèse des réponses au questionnaire en ligne.

#### 2.2.2.1 Entretiens avec les documentalistes

##### *Valoriser la production scientifique régionale*

De manière générale, les documentalistes attendent d'un tel portail qu'il valorise les productions issues des différents organismes et établissements de recherche de la région. Ainsi, la BOMLR est perçue comme une opportunité pour développer un véritable réservoir collectif des publications produites par les chercheurs des différents instituts de recherche du Languedoc-Roussillon. Plus précisément, on attend de la BO-MLR qu'elle cible en priorité les publications des

<sup>19</sup> La grille d'entretien ainsi que le questionnaire sont présentés en annexe pages 71-73

secteurs de l'agronomie, l'environnement, la biodiversité ainsi que toutes les thématiques connexes, qui restent les champs privilégiés de l'association Agropolis International.

### *Proposer un accès aux documents en texte intégral*

Augmenter la visibilité de la production scientifique locale et régionale dans les secteurs de l'agronomie, de l'environnement, de la biodiversité pourrait se faire dans un premier temps au niveau du signalement. Toutefois, l'accent a également été mis sur la nécessité de proposer un accès aux documents en *full-text*. Dans ce sens, il semble intéressant que le portail interface certains domaines géographiques et/ou thématiques des archives ouvertes telles que HAL ou Prod-INRA ainsi que les collections PDF des différents établissements, à l'instar de celles de Bioversity International. L'interfaçage des archives ouvertes est d'autant plus intéressant que nombreux sont les chercheurs des différents établissements à verser leurs productions dans ce type de réservoirs. L'accès aux documents *full-text* devrait pouvoir être proposé en tant que filtre pour l'affinage des résultats ou comme option dans la recherche avancée. L'utilisateur pourrait alors cibler sa recherche uniquement sur l'ensemble des documents disponibles en ligne en texte intégral.

Un autre type de ressources semble également intéressant à valoriser : il s'agit des travaux des étudiants et des doctorants, regroupant les mémoires et les thèses. Là encore, un interfaçage ciblé de bases communes telles que TEL, pourrait être intéressant. En effet, de nombreux établissements déposent les thèses de leurs étudiants sur ce type de plateforme.

### *Permettre le signalement et la localisation des collections acquises par les bibliothèques*

Par ailleurs, la valorisation des collections d'ouvrages et de périodiques papier et en ligne pourrait se faire via l'interfaçage de catalogues spécialisés de types « AtoZ », pour permettre aux étudiants, en particulier, d'aller chercher les ressources qui les intéressent là où elles sont.

### *S'orienter vers des outils de recherche plus simples et plus intuitifs, axés utilisateurs*

De manière générale, le module de recherche doit être plus simple et plus intuitif. La recherche avancée avec les opérateurs booléens reste utilisée par les professionnels de l'information mais les documentalistes ont remarqué que la majeure partie de leurs usagers privilégiait la recherche simple - « à la Google »- suivie d'un affinage des résultats. Il est donc nécessaire de proposer de nombreux filtres à l'affichage des réponses. Selon les interrogés, les filtres par date et type de ressources semblent indispensables.

La BOMLR doit également s'orienter vers d'autres nouveaux services. Avec le développement du web 2.0, il faut effectivement envisager d'ajouter une dimension participative. Ainsi, le partage de signets sur les réseaux sociaux, la possibilité d'ajouter des notes, des commentaires, etc. semblent être des options à considérer.

L'accès à un compte personnalisé qui permet à l'utilisateur de gérer des paniers, de sauvegarder ses recherches et de se constituer des listes thématiques est également apparu comme un service intéressant à proposer. Ce module d'identification pourrait également permettre à chaque usager de gérer des abonnements à des flux RSS (abonnements à des mots-clés de recherche par exemple).

Actuellement, le portail propose une interface bilingue anglais-français. Dans l'objectif d'un rayonnement international, certains documentalistes ont mis l'accent sur la nécessité de proposer, en plus d'une interface en anglais et en français, une interface en espagnol. Dans cette perspective, une aide à la saisie par traduction automatique des mots-clés a également été suggérée.

## **2.2.2.2 Résultats du questionnaire<sup>20</sup>**

Le questionnaire en ligne, intitulé « Vos besoins en recherche d'information au niveau régional » a permis de recueillir les réponses de 87 personnes, parmi lesquelles 21 professionnels

<sup>20</sup> Une présentation des résultats sous forme de diagrammes se trouve en annexe p.75

de l'information, 20 chercheurs, 15 étudiants, 11 doctorants et 12 enseignants-chercheurs. Les résultats de ce questionnaire ont notamment permis de mettre en évidence les ressources en ligne les plus sollicitées par les personnes interrogées, à savoir :

- Le Web of Science
- Google Scholar
- Pubmed
- Science Direct
- CAIRN
- Scirus
- Les archives ouvertes HAL et ProdiNRA<sup>21</sup>
- Les catalogues de bibliothèques des établissements

De plus, le questionnaire a permis de confirmer que la recherche simple était bien le mode de recherche le plus utilisé par les différents publics. Sur 87 personnes, 60 ont en effet indiqué employer la recherche simple, plutôt que la recherche avancée. L'utilisation des filtres pour l'affinage des résultats – avec l'emploi en priorité des filtres par type de ressources, par année et par mots-clés – est également apparue importante. Les outils de veille (alertes, flux RSS) sont également majoritairement exploités par les utilisateurs.

### **2.2.3 De nouveaux objectifs pour la BOMLR**

L'analyse des besoins a permis d'identifier les objectifs suivants pour le réaménagement du portail de la BOMLR :

- Intégrer de nouvelles ressources : au-delà des catalogues de bibliothèques, proposer une recherche dans des bases de revues en ligne, archives ouvertes ou autres réservoirs de publications scientifiques.
- Développer une interface ergonomique, qui propose à ses utilisateurs des outils de recherche et d'affichage des résultats clairs et intuitifs.
- Offrir de nouvelles fonctionnalités : facettes, alertes/flux RSS, exportation de notices, lien vers le texte intégral depuis la page des résultats, etc.
- Mettre en évidence la localisation des ressources (soit en ligne, soit disponible dans tel ou tel centre de documentation).

Après ces premières analyses, j'ai pu donc véritablement engager le travail sur les outils de recherche documentaire. Le réaménagement de la BOMLR repose effectivement sur la mise en place d'un nouveau méta-moteur, qui doit répondre aux objectifs identifiés par le groupe-projet suite à l'état des lieux et à l'étude des besoins.

La première étape de ce travail a été une étape exploratoire, au cours de laquelle j'ai dû recenser les principaux outils actuellement proposés sur le marché. Ce panorama des solutions a été présenté lors d'une réunion du groupe-projet, à l'issue de laquelle une première sélection a été effectuée.

<sup>21</sup> Se reporter au glossaire en page 96



## Partie 3 – Réaménager la BOMLR : étude et proposition de solutions

---

Avant d'exposer les différentes solutions de recherche envisageables pour le réaménagement de la BOMLR, il m'a semblé nécessaire de mettre en évidence les caractéristiques et les principaux mécanismes de la recherche fédérée, afin de bien cerner le fonctionnement des méta-moteurs. Il n'était effectivement pas évident au départ d'appréhender toutes ces technologies, en perpétuelle évolution.

### 3.1 LES PRINCIPAUX MÉCANISMES DE LA RECHERCHE FÉDÉRÉE

La recherche fédérée est l'interrogation simultanée, à travers une interface unique, de ressources hétérogènes. Elle peut ainsi interroger des catalogues de bibliothèque mais aussi diverses bases de données de thèses, mémoires, articles de périodiques, etc. reposant sur des logiciels de gestion de l'information différents (PMB, Symphony, Alexandrie-Kentika, etc.). Ces ressources sont également hétérogènes du fait de leurs dissemblances sur le plan de leur structuration, leurs formats pouvant être très différents : XML, UNIMARC, etc. Également connue sous le nom de « méta-recherche », la recherche fédérée est donc essentielle pour l'utilisateur qui peut ainsi accéder plus facilement à un tout un ensemble de ressources. Elle repose sur un outil particulier : le méta-moteur.

Un méta-moteur est défini par trois éléments essentiels : l'interface utilisateur, les fonctionnalités de recherche et les fonctionnalités d'affichage des résultats. Il repose aujourd'hui sur deux grands types de technologies, la technologie des « connecteurs » (mécanisme synchrone) et la technologie du « moissonnage » (mécanisme asynchrone, plus récent).

#### 3.1.1 La technologie des connecteurs

Dans le premier cas, le moteur utilise donc des connecteurs – ou *wrappers* en anglais – pour traduire la requête dans la syntaxe employée par chaque base de données distante. Il recueille ensuite la liste des résultats correspondant à la requête saisie par l'utilisateur et l'affiche sur son interface. L'interrogation des ressources se fait ici en temps réel, en mode dit « synchrone ». Le système de connecteurs implique un suivi régulier des passerelles. Il est donc nécessaire d'être en relation avec les services informatiques des différents établissements pour se tenir au courant des évolutions des bases interfacées.

Pour accéder aux bases de données distantes, le moteur doit utiliser plusieurs protocoles de communication, qui varient selon les bases interrogées. De manière générale, on trouve deux grands types de mécanismes d'interrogation pour les échanges de données :

- Les Webservices<sup>22</sup> : basés sur le protocole HTTP, ce mécanisme favorise l'interopérabilité entre diverses applications distantes et systèmes hétérogènes à travers le réseau Internet.

---

<sup>22</sup> Se reporter au glossaire en p.96

- Le protocole Z3950<sup>23</sup> : ce protocole de communication permet l'échange d'informations entre différentes bases de données à travers un réseau informatique. Il est surtout employé par les bibliothèques pour leurs catalogues.

Pour les deux types de mécanismes, la mise en forme des données se fait généralement sous format XML.

### **3.1.2 La technologie du « moissonnage »**

Dans le second cas, le méta-moteur récupère et stocke régulièrement les métadonnées des notices bibliographiques et/ou le texte intégral des documents issus des différentes bases de données. Ce moissonnage est rendu possible grâce au protocole OAI-PMH (Open Archives Initiative's Protocol for Metadata Harvesting)<sup>24</sup>. Ce protocole, issu du mouvement de l'*Open Access* (libre accès)<sup>25</sup> a pour but de faciliter l'accès en ligne aux documents numériques. Il utilise généralement le format de métadonnées «Dublin Core »<sup>26</sup> - une norme de description simple des ressources fondée sur quinze éléments – mais aussi, plus rarement, d'autres formats de données, comme le format XML. À partir de ce moissonnage, le moteur crée un index en local, et c'est cet index qui sera interrogé par l'utilisateur lors de sa recherche. On parle alors plutôt de « recherche intégrée ». Cette technologie permet souvent un délai d'affichage des réponses plus rapide, puisqu'elle dépend d'un seul serveur, celui du fournisseur du service.

Cependant, les résultats proposés à l'utilisateur à l'issue de sa requête sont le reflet d'une collecte de données par le moissonneur à un instant T (moment du dernier passage du moissonneur). Nous avons donc affaire ici à un système asynchrone. Certains outils proposent aujourd'hui les deux technologies : connecteurs et moissonneurs OAI.

Un schéma récapitulatif des différentes technologies liées au fonctionnement d'un méta-moteur se situe en annexe du document (annexe 5 p.55).

## **3.2 PANORAMA DES SOLUTIONS DE RECHERCHE**

Au cours de ma veille sur les différentes solutions, j'ai pu établir un panorama des logiciels, à la fois libres et propriétaires, présents sur le marché. Ce panorama a été présenté lors d'une réunion du groupe-projet qui s'est déroulée le 22 mars 2012.

Les solutions ont pu être classées selon les quatre grandes catégories suivantes, établies progressivement au fil de mes recherches :

- Catégorie 1 : Les moteurs de recherche fédérée
- Catégorie 2 : Les moteurs de recherche intégrée (moissonneurs OAI)
- Catégorie 3 : Les moteurs dits « hybrides » ou « mixtes », qui proposent d'emblée, une recherche fédérée et intégrée.

<sup>23</sup> Se reporter au glossaire en p.96

<sup>24</sup> Se reporter au glossaire en p.96

<sup>25</sup> Se reporter au glossaire en p.96

<sup>26</sup> Se reporter au glossaire en p.96

- Catégorie 4 : Les outils de découverte. Axés interface de recherche, ces outils présentent des fonctionnalités de recherche et d'affichage des résultats particulièrement développés.

### **3.2.1 Les moteurs de recherche fédérée**

Parmi les moteurs de recherche fédérée - solutions dite de « premier niveau », et généralement plus anciennes - j'ai recensé plusieurs solutions libres et propriétaires énumérées ci-dessous :

- **Métalib**, solution propriétaire développée et commercialisée par la société israélienne Ex-Libris.
- **360 Search**, développé et commercialisé par la société Serials Solutions, implantée sur tous les continents.
- **DbWiz**, moteur de recherche fédérée libre développé par l'Université Canadienne Simon Fraser.
- **PazPar2**, moteur de recherche fédérée libre développé par Index Data, un groupe danois, implanté à Copenhague ainsi qu'aux États-Unis.

### **3.2.2 Les moteurs de recherche intégrée**

Parmi les moteurs de recherche intégrée, basés sur l'indexation des métadonnées, j'ai relevé trois solutions propriétaires :

- **EDS (Ebsco Discovery Service)**, solution propriétaire proposée par Ebsco.
- **Primo**, solution propriétaire de la société Ex-Libris
- **Summon**, commercialisée par Serials Solutions.

Ces trois outils peuvent inclure, en plus de la recherche intégrée, la recherche fédérée (technologie des connecteurs), chacun de ces trois éditeurs (Ebsco, Serials Solutions et Ex-Libris) ayant déjà développé son propre moteur de recherche fédérée. On l'a vu par exemple avec Ex-Libris, qui détient la solution Métalib.

Certains de ces outils proposent des fonctionnalités de recherche et d'affichage des résultats particulièrement élaborées, comme Summon, qui peut donc être considéré comme un outil de découverte.

J'ai également trouvé un outil de recherche intégrée libre :

- **Spip-Agora** : historiquement utilisé comme un système de publication de contenu (CMS), il a été complété par des briques logicielles « moteurs », suite au projet SIST (Scientific and Technical Information System) du Cirad. Ce projet a donc permis de développer un moissonneur OAI, créant ainsi une solution de recherche intégrée.

### **3.2.3 Les moteurs hybrides**

Parmi les véritables mécanismes hybrides, qui proposent à la fois la recherche fédérée et intégrée, j'ai recensé les solutions suivantes :

- **Explorit**, de la société Deep Web Technologies (logiciel propriétaire)
- **QES Federator**, de Qwam Content Intelligence (logiciel propriétaire)
- **LibraryFind**, logiciel libre développé par l'Université d'Oregon.

Au départ simple moteur de recherche fédérée, il est devenu, au fil des divers développements de ses utilisateurs, un véritable outil hybride. Il est aujourd'hui adopté par quelques grands établissements français d'information et de documentation comme la Bibliothèque Publique d'Information – BPI – et le Service Inter-établissement de Coopération Documentaire de Grenoble.

### **3.2.4 Les outils de découverte**

Ces outils, lancés plus récemment sur le marché des solutions de recherche, offrent une interface de recherche unique pour l'ensemble des collections d'une organisation. La plupart du temps, ils n'imposent pas de moteurs de recherche particuliers, bien qu'ils incluent de fait un système d'indexation des données. Ce sont des interfaces de recherche avant tout, qui sont souvent intégrées à un portail de bibliothèque déjà en possession d'un moteur de recherche fédérée et/ou intégrée. Ces solutions proposent des fonctionnalités de recherche et d'affichage des résultats particulièrement développées. Plus précisément, ces outils incluent la plupart du temps un système de navigation à facettes qui rend la recherche plus facile et plus intuitive pour l'utilisateur. Ils peuvent également proposer des outils du web 2.0 (partage sur les réseaux sociaux, ajout d'avis ou de commentaires et autres fonctions collaboratives).

J'ai relevé deux outils de découverte libres :

- **Blacklight**, développé par l'Université de Virginie aux États-Unis.
- **Vufind**, développé par la *Villanova University* (Pennsylvanie, États-Unis)

## **3.3 SOLUTIONS LIBRES ET PROPRIÉTAIRES : QUELLES DIFFÉRENCES ?**

Par rapport aux logiciels propriétaires, les solutions libres ont une durée de vie plus longue et présentent une meilleure adaptabilité aux besoins des usagers. En effet, leur code source étant ouvert, il est possible de développer autant de nouvelles fonctionnalités que nécessaires, et de faire évoluer le produit constamment. Par ailleurs, une solution libre est généralement moins coûteuse qu'une solution propriétaire. Si le développement et la mise en place des fonctionnalités du système impliquent la plupart du temps de faire appel à un prestataire extérieur (Société de Services en Ingénierie Informatique, etc.), il n'y a pas de coût de départ, ni d'abonnement annuel à payer.

## **3.4 ÉTABLISSEMENT DES CRITÈRES POUR L'ÉVALUATION COMPARATIVE DES SOLUTIONS**

### **3.4.1 La nécessité d'un outil mixte : un premier critère de sélection**

Étant donnée l'hétérogénéité des ressources et des modes de communication utilisés par les différentes bases de données des membres de la CIST d'Agropolis, le groupe-projet a tout d'abord décidé de ne retenir pour l'étude comparative que neuf des

dix-sept outils présentés, donnant ainsi la priorité aux solutions « hybrides », c'est-à-dire aux outils qui intègrent – ou ont la possibilité d'intégrer – les deux systèmes de recherche : la recherche fédérée et la recherche intégrée. Il est en effet nécessaire que l'outil propose à la fois l'interrogation synchrone (par connecteurs) et l'interrogation asynchrone (par moissonnage OAI-PMH).

Cette première sélection comprend quatre outils libres et cinq outils propriétaires, à savoir :

Solutions libres	Solutions propriétaires
<ul style="list-style-type: none"> <li>• Blacklight</li> <li>• LibraryFind</li> <li>• Spip Agora</li> <li>• VuFind</li> </ul>	<ul style="list-style-type: none"> <li>• EBSCO Discovery Service</li> <li>• Explorit</li> <li>• Primo</li> <li>• QES Federator</li> <li>• Summon</li> </ul>

### **3.4.2 Les fonctionnalités de recherche et d'affichage des résultats attendues**

Au-delà de l'exigence d'un outil « mixte », d'autres critères sont apparus nécessaires à prendre en compte. Ces critères, définis lors de la réunion du 22 mars, ont permis de centrer l'analyse des neuf solutions retenues sur des aspects fonctionnels plus précis. L'outil doit ainsi proposer en priorité les fonctionnalités suivantes :

*Côté « utilisateur »*

- Interrogation des catalogues d'ouvrages/de périodiques et des bases de publications des différents établissements d'enseignement supérieur et organismes de recherche
- Lien vers le texte intégral ou, à défaut, proposition d'un filtre *full-text* à l'affichage des résultats
- Localisation des ressources de la région (filtre par centre de documentation ou par organisme par exemple)
- Distinction entre les ressources en ligne et les ressources à disposition dans tel ou tel centre documentaire
- De manière générale, un affinage développé, comportant un grand nombre de filtres (par type de ressources, auteur, date, thématique, etc.)

*Côté « administrateur »*

En lien avec le caractère évolutif de l'outil - qui reste un critère essentiel - la solution doit comprendre :

- La gestion des connecteurs
- L'indexation des sources externes
- La paramétrabilité de l'interface utilisateur

- Un module statistiques : une telle fonction est effectivement apparue essentielle au groupe-projet, car elle permet de rendre compte concrètement de l'utilisation et de la fréquentation de l'interface. Ce module doit ainsi permettre d'évaluer le nombre de visiteurs et de recherches effectuées sur plusieurs périodes, ou encore de mettre en évidence les ressources les plus interrogées.

D'autres fonctionnalités, si elles ne sont pas essentielles, pourraient être un plus :

- La possibilité d'élargir la recherche vers d'autres sites/moteurs tels que Google Scholar, Web of Science, Science Direct, etc.
- L'authentification des utilisateurs via un *Log In*
- L'aide à la saisie
- L'exportation de notices
- Un système de création d'alertes/flux RSS
- Le partage de notices sur les réseaux sociaux (Twitter, Facebook, etc.)

D'autre part, la solution doit pouvoir bénéficier d'une assez large communauté d'utilisateurs, francophones si possible. S'il s'agit d'une solution propriétaire, un interlocuteur francophone est également apprécié.

Enfin, les solutions libres seront privilégiées. Une solution libre - au code source ouvert - offre la possibilité de développer des fonctionnalités particulières, en accord avec les besoins de ses utilisateurs. Dans la logique de l'*open-source*, les développements effectués par les utilisateurs de la solution sont en principe reversés et partagés avec la communauté des utilisateurs. L'absence de certaines fonctionnalités dans une solution libre ne doit donc pas être considérée comme une véritable limite.

### **3.5 TABLEAU COMPARATIF DES SOLUTIONS**

À partir de ces critères, j'ai pu établir une évaluation comparative des solutions retenues à l'issue de la première sélection. Afin de faciliter la comparaison, cette évaluation a été exposée sous la forme du tableau présenté ci-après.



	Outils de découverte		Moteurs de recherche intégrée				Moteurs de recherche fédérée et intégrée		
	<b>Blacklight</b>	<b>Vufind</b>	<b>Summon</b>	<b>Primo</b>	<b>Ebsco D.S</b>	<b>Spip Agora</b>	<b>LibraryFind</b>	<b>Explorit</b>	<b>QES Federator</b>
Interrogation périodiques	oui	oui	oui	oui	oui	oui	oui	oui	Oui
Interrogation publications	oui	oui	oui	oui	oui	oui	oui	oui	oui
Lien vers texte intégral depuis la page de résultats	non	oui	oui	non mais accès <i>full-text</i> depuis la notice (un « clic » supplémentaire)	oui	non	oui	non	oui
Affinage des résultats	oui	oui	oui	oui	oui	non	oui	oui	non
Filtre par centre de doc/source	oui	oui	oui	oui	possible	non mais sélection des sources interrogeables pour la recherche	oui	oui	oui (pré-sélection pour la recherche)
Filtre par date	oui	oui	oui	oui	oui	non	oui	oui	oui
Filtre « <i>full-text</i> » ou « consultable en ligne »	oui	non	oui	non	non	non	oui	possible	non
Filtre par type de ressource	oui	oui	oui	oui	oui	oui (pré-sélection pour la recherche)	oui	possible	non
Filtre par thématique/ sujet	oui	oui	oui	oui	possible	oui	oui	possible	oui
Filtre par auteur	oui	oui	oui	oui	possible	non	oui	oui	non

	Outils de découverte		Moteurs de recherche intégrée				Moteurs de recherche fédérée et intégrée		
	Blacklight	Vufind	Summon	Primo	Ebsco D.S	Spip Agora	LibraryFind	Explorit	QES Federator
Possibilité d'élargir la recherche vers d'autres sites/moteurs (Google Scholar, Wos, Science Direct, etc.)	?	?	oui (utilisation du résolveur de liens 360 pour Google Scholar – permet d'accéder aux documents en full-text)	oui (rebonds possibles vers Google Scholar, le Web of Science et le SUDOC également)	oui	non	Wos et Science Direct prévus (voir avec le SICD de Grenoble)	Google Scholar	oui
Compte Utilisateurs	oui	oui	oui	oui	oui	non	oui	-	possible (Module complémentaire)
Exportation des notices	oui	oui	oui	oui	oui	oui	oui	oui	oui
Aide à la saisie	non	oui	oui	non, mais corrections automatiques	non	non	non	non	non
Alertes, RSS	oui	oui	oui	oui	oui	oui	oui	oui	via compte personnel uniquement
Partage sur les réseaux sociaux	non	non	à venir	à venir	non	non	oui (voir projet BPI)	non	non
Module Statistiques	à mettre en place	à mettre en place	oui	Oui (générateur de statistiques basé sur le logiciel Open Source BIRT)	oui	non	en cours de développement au SCD de Grenoble (le module devrait permettre d'obtenir des informations sur les recherches, les bases interrogées, l'utilisation des services, les téléchargements, etc.)	oui	oui

	Outils de découverte		Moteurs de recherche intégrée				Moteurs de recherche fédérée et intégrée		
	<b>Blacklight</b>	<b>Vufind</b>	<b>Summon</b>	<b>Primo</b>	<b>Ebsco D.S</b>	<b>Spip Agora</b>	<b>LibraryFind</b>	<b>Explorit</b>	<b>QES Federator</b>
<b>Exemples</b>	Northwest Digital Archive  Stanford University Library	DI-Fusion, portail Université Libre Bruxelles (entrepôts OAI)  Villanova University  Biblioteca Nazionale Braidense	SCD Paris 7  École des Ponts Paris-Tech  University of Calgary	SCD Paris 3 (Portail Virtuose+)  NEBIS, Réseau de Bibliothèques et de centres d'information en Suisse (accès depuis le portail de l'École Polytechnique fédérale de Lausanne)	Bibliothèques HEC  Ankara University	Portail de l'IST du Burkina Faso (outil surtout utilisé pour les sites des pays en voie de développement)	Utah Academic Library Consortium  <u>Projets en cours</u> : - BPI - SCD Grenoble	Science.gov  Mednar.com	Portail des ressources de l'École des Mines de Paris
<b>Classement des solutions (classement effectué par catégorie : O.D, MRI ou MRF/I)</b>	+	++	++	++	+	-	++	+	+

## 3.6 EXPLORATION DES SOLUTIONS

### 3.6.1 Les solutions retenues

Suite à l'étude comparative des neuf solutions, quatre solutions ont été retenues et présentées lors de la réunion de la Commission Information Scientifique et Technique du 3 avril 2012. Il s'agit des solutions libres **Vufind** et **LibraryFind** d'une part, et des solutions propriétaires **Summon** et **Primo** d'autre part.

### 3.6.2 Exploration de la sélection : premiers résultats

Les dernières semaines de mon stage ont consisté à étudier plus en détails ces quatre solutions afin de mesurer - aussi précisément que possible - leurs avantages et leurs inconvénients, et d'émettre des recommandations quant au choix final de la solution.

Tout d'abord, un *mini* cahier des charges, synthétisant les exigences d'Agropolis (aspects documentaires, techniques et fonctionnels) a été envoyé aux différents éditeurs et développeurs des solutions, afin de confirmer la compatibilité de chaque méta-moteur en fonction des besoins et des contraintes de la BOMLR.<sup>27</sup>

Les solutions ont ensuite été observées plus précisément en fonction des critères suivants :

- **Les technologies**

Toutes les solutions peuvent inclure les deux modes de recherche (via les connecteurs ou via le protocole OAI-PMH) puisque, rappelons-le, le caractère « mixte » de la solution a constitué un critère essentiel pour la sélection des outils.

Cependant, il est apparu essentiel de distinguer les solutions qui comprennent d'emblée les deux technologies de celles qui n'en proposent qu'une dans leur *package* initial et nécessitent de développer ou d'intégrer l'autre. Dans ce dernier cas, il faut envisager du temps et un budget supplémentaires.

De ce point de vue, la solution libre **LibraryFind** se démarque des autres. Les deux techniques ont effectivement déjà été développées et utilisées avec succès par ses usagers. Pour **Vufind**, au-delà de l'indexation des données par OAI-PMH et de l'import en format MARC, l'utilisation de la Z3950 est possible, mais celle-ci n'a encore jamais été implémentée : des développements importants seraient donc à faire. Enfin, **Primo** et **Summon** sont avant tout des moteurs de recherche intégrée. Ils proposent donc surtout l'indexation des données par OAI-PMH, qui s'avère par ailleurs être un protocole d'avenir, en plein essor. Dans un second temps, et sur demande, il est possible d'intégrer des connecteurs, et de mettre ainsi en place l'autre mode de recherche.

- **Le langage**

La question du langage de développement ne se pose pas si la solution est gérée par l'éditeur (SaaS<sup>28</sup> *Software as a Service*), mais dans le cas où la solution est hébergée sur le serveur de l'établissement, il est important de posséder des compétences en interne, et plus encore lorsqu'il s'agit de solutions libres. Cela pose un problème pour la solution **LibraryFind**, développée dans le langage « Ruby on Rails », non pratiqué par l'informaticien en poste à Agropolis. Des compétences seront à rechercher localement dans les différents instituts et organismes d'Agropolis et de Montpellier.

<sup>27</sup> Se reporter au document en annexe p. 80

<sup>28</sup> Se reporter au glossaire p.96

- **L'hébergement**

Pour les solutions propriétaires, un problème de dépendance vis-à-vis de la société peut se poser en cas d'hébergement à distance – ou « SaaS ». C'est le cas de la solution **Summon**, proposée uniquement en mode SaaS. Pour l'autre solution propriétaire **Primo** et pour les solutions libres, un hébergement sur place implique un serveur avec une grande capacité de stockage, et une mémoire vive importante, en mesure de supporter le lourd processus d'indexation des métadonnées.

- **Le caractère évolutif de la solution**

Comme nous l'avons déjà évoqué plus haut, les solutions libres présentent une meilleure adaptabilité aux besoins des usagers puisque, grâce à leur code source ouvert, il est possible de développer autant de nouvelles fonctionnalités que nécessaire. Au contraire, l'évolution d'un logiciel propriétaire dépend de la société qui l'a développé.

- **Le coût, à court et plus long terme**

Quelle que soit la solution choisie, il faut considérer, au-delà de la licence, le coût de l'installation, du suivi et de la maintenance de l'outil. S'il y a lieu de personnaliser l'application, il faut prendre en compte le coût d'éventuels développements. Des frais de formation à l'utilisation du logiciel peuvent aussi s'ajouter à ces coûts. Si l'installation d'une solution requiert des compétences qui ne sont pas mobilisées en interne, il faut également envisager de faire appel à une société de services en informatique. On a tendance à dire qu'une solution libre est moins coûteuse qu'une solution propriétaire. Sur le court terme, c'est tout à fait vrai, puisque la licence et l'installation d'un outil libre sont totalement gratuites. Cependant, il ne faut pas négliger les coûts éventuels de développements, qui impliquent bien souvent de faire appel à un prestataire extérieur.

Au-delà des développements, le paramétrage des bases nécessiterait l'embauche d'une personne, tout au moins dans les premiers temps. Sachant qu'une demi-journée de travail est nécessaire pour la configuration d'une base, une dizaine de jours au total devrait suffire pour paramétrer l'ensemble des bases.

La question du prix de la solution restant un critère non négligeable pour le choix de l'outil, j'ai également tenté d'établir une estimation des coûts de développements pour les solutions libres. Pour chaque solution, plusieurs cas étaient possibles; il a fallu estimer les budgets temps et financiers pour chaque cas.

Au final, il s'est avéré que, tant pour **LibraryFind** que pour **Vufind**, la mise en place du protocole OAI-PMH dans les établissements non encore équipés permettrait de limiter les coûts du projet, le développement des connecteurs spécifiques, pour les Webservices en particulier, restant plus coûteux. Une fois ce protocole installé dans tous les établissements, il resterait à paramétrer l'indexation des métadonnées de l'ensemble des ressources. Cependant, cette solution ne peut être envisagée qu'à moyen terme, l'installation du protocole OAI s'effectuant au rythme de chaque établissement.

Des tests ont été effectués avec la solution libre **Vufind**, installée sur le serveur d'Agropolis.

<sup>29</sup> Le moissonnage OAI-PMH a été testé avec succès pour une partie des ressources de l'IRSTEA et de l'IRD, ainsi que pour la totalité des ressources d'AgroParisTech Montpellier (20 000 références environ).

### **3.6.3 Exploration de la sélection : perspectives**

Ayant établi depuis le début de mes recherches sur les méta-moteurs un bon contact avec le SICD de Grenoble - actuellement en phase de production de sa future interface de recherche fédérée « Surgir » via l'outil **LibraryFind** - j'ai organisé un déplacement sur place pour rencontrer

<sup>29</sup> Se reporter aux captures d'écran de l'interface-test Vufind en annexe p. 91

la chef de projet et l'équipe d'informaticiens chargés de la mise en place de l'interface le mercredi 13 juin, accompagnée de Chantal Salson, ma responsable de stage, et de Jean Cerda, du service informatique d'Agropolis.

Parallèlement, afin de poursuivre les tests des solutions, l'installation de la solution LibraryFind sur le serveur d'Agropolis a été programmée.

Ce travail, qui doit conduire au choix final de la solution, a constitué une des étapes les plus difficiles de mon stage. Il n'a effectivement pas été facile d'évaluer précisément les solutions, et ce d'autant plus que, si des tests ont pu être effectués sur les solutions libres, les éditeurs des solutions propriétaires n'ont proposé que des démonstrations.

Cette étape s'est donc clairement plus portée sur l'observation et la comparaison des deux solutions libres, **Vufind** et **LibraryFind**, pour lesquelles le groupe-projet a d'ailleurs à l'heure actuelle largement affirmé sa préférence.



## Conclusion

---

Au cours de ces quatre mois de stage, j'ai pu participer au projet d'évolution de la Bibliothèque Ouverte Montpellier – Languedoc-Roussillon, depuis l'analyse de l'existant jusqu'à l'exploration de nouvelles solutions, en passant par l'étude des besoins et la formalisation des contraintes et des exigences des membres de la CIST d'Agropolis.

Cette expérience m'a permis d'acquérir des connaissances sur de nouvelles technologies et de nouveaux outils documentaires, particulièrement prisés par le monde des bibliothèques qui cherche de plus en plus à faciliter et à élargir l'accès à ses ressources numériques. C'est un travail riche, me semble-t-il, car il implique des connaissances assez précises sur les techniques liées à la recherche documentaire et oblige à réfléchir aux outils de recherche appropriés dans un cadre de projet collaboratif, et plus largement, dans le sens d'une valorisation de l'Information Scientifique et Technique. De plus, la gestion de ce projet a impliqué des liens avec de nombreuses professions, et j'ai particulièrement apprécié les rencontres avec les experts, que ce soit documentalistes IST, informaticiens ou éditeurs de logiciels. La constitution d'une équipe mixte, rassemblant informaticiens et documentalistes, me semble d'ailleurs désormais indispensable pour travailler sur un tel projet, où les problématiques liées à la technique et à l'évolution des technologies tiennent une si grande place.

Cette expérience m'a également permis de mettre en œuvre les connaissances acquises au cours de ma formation à l'ENSSIB, et notamment les méthodes de conduite de projet. J'ai donc rigoureusement suivi les étapes de la gestion de projet, tout en réalisant fréquemment des rapports qui m'ont permis de faire régulièrement le point sur l'état d'avancement du projet et d'en établir les objectifs, remaniés et précisés au fil des réunions.

Malgré tout, ce projet collaboratif a rencontré quelques obstacles, parmi lesquels la difficulté à coordonner emplois du temps et calendriers pour organiser rencontres ou déplacements. Par ailleurs, il semble difficile de trouver un dispositif qui permette de regrouper toutes les ressources souhaitées et ce, principalement pour des raisons techniques. L'hétérogénéité des protocoles de communication et des formats de données semble encore poser problème. De plus, il sera difficile de satisfaire tous les besoins des usagers, et particulièrement en terme d'accès au texte intégral qui, bien qu'encouragé par le mouvement de l'*open access*, suscite encore de nombreuses oppositions, notamment auprès des éditeurs. En outre, cet accès concerne essentiellement les articles et les travaux scientifiques, car l'offre de livres numériques est encore assez restreinte, surtout en France où la question des droits d'auteur appliqués aux contenus numériques reste complexe. Néanmoins, dans la perspective du développement des *e-books*, la Bibliothèque Ouverte Montpellier – Languedoc-Roussillon devra bien sûr envisager l'accès à ces nouvelles ressources.

Il reste donc encore une difficulté de taille : le choix final de la solution, obligatoirement imparfaite. En outre, après cette étape décisive qui clôturera la phase de conception du projet, il faudra défendre le projet auprès des partenaires susceptibles de participer à son financement, et sans lesquels il ne pourra se réaliser.

Ce stage m'a finalement fait prendre conscience qu'un projet complexe de ce type n'a en fait jamais de fin, car il doit sans cesse évoluer pour s'adapter à la transformation constante du monde de l'information et des besoins des usagers.



# ***Bibliographie et Webographie***

## **Bibliographie**

### *Monographies*

Timothy COLE, Muriel FOULONNEAU, *Using the Open Archives Initiative protocol for metadata harvesting*, Linworth, Libraries Unlimited, 2007, 224 p.

Marc MAISONNEUVE, *Logiciels portails pour bibliothèques et centres de documentation : l'offre d'outils de recherche fédérée et de gestion de contenu*, étude réalisée en collaboration avec Cécile Touitou (Tosca Consultants), Paris, ADBS Editions, 2007, 215 p.

Marc MAISONNEUVE, *Le catalogue de la bibliothèque à l'heure du web 2.0 : étude des opacs de nouvelles génération*, Paris, ADBS Editions, 2008, 305 p.

### *Articles de revues*

Medhi ALIPOUR-HAFEZY *et alii*, « Interoperability models in digital libraries: an overview », *Electronic Library*, Vol.28, n°3, 2010, pp. 438-452.

Jerry CASWELL, John WINSTRA, « Improving the search experience : federated search and the library gateway », *Library Hi-Tech*, Vol.28, n°3, 2010, pp.391-401.

Ian GIBSON, Lisa GODDARD et Shannon GORDON, « One box to search them all : implementing federated search at an academic library », *Library Hi-Tech*, Vol. 27, n°1, 2009, pp. 118-133.

Ronda ROWE, « Web-Scale Discovery : a review of Summon, EBSCO Discovery Service and WorldCat Local », *The Charleston Advisor*, Vol.12, n°1, octobre 2010, pp. 5-10.

Tamar SADEH, « Transforming the metasearch concept into a friendly user experience », *Internet Reference Services Quarterly*, Vol.12, n°1-2, 2008, pp. 1-25.

Chantal SALSON *et alii*, « La Bibliothèque Ouverte : pour un accès unique aux ressources de la recherche et de l'enseignement supérieur en Languedoc-Roussillon », *Documentaliste-Sciences de l'Information*, Vol.41, 2004, pp. 14-25.

David STERN, « Harvesting : power and opportunities beyond federated search », *Online*, vol. 33, n°4, 2009, pp. 35-37.

Karen STEVENSON, « Next-Generation Library Catalogues : reviews of Encore, Primo, Summon and Summa », *Serials*, Vol.22, n°1, mars 2009, pp. 68-82.  
[en ligne : <http://uksg.metapress.com/content/c85133g1360mx101/fulltext.pdf> ]

(consultation avril 2012)

### *Rapports et travaux d'études*

Aurélia GIUSTI, *La recherche fédérée des portails patrimoniaux : quelles solutions documentaires ?*, sous la direction de Ghislaine Chartron, mémoire de stage, 2009.  
[en ligne : <http://memsic.ccsd.cnrs.fr/docs/00/52/39/25/PDF/GIUSTI.pdf> ]  
(consultation avril 2012)

Sandrine MOURET, *Une interface sur le site Web d'Agropolis pour interroger en une requête les bases bibliographiques des établissements : étude de marché des solutions possibles*, sous la direction de Hanka Hensens et Marie-Claire Allaya, mémoire de stage, 2000.<sup>30</sup>

### **Webographie**

Penny BAILEY, « Recent developments in federated searching », mars 2008.  
[http://www.baileysolutions.co.uk/images/stories/documents/Federated\\_searching.pdf](http://www.baileysolutions.co.uk/images/stories/documents/Federated_searching.pdf)  
(consultation avril 2012)

Johanne DEGARIE, Guylaine GIRARD, Michèle LEBLANC, Victoria VOLKANOVA, « Les outils de découverte », rapport du groupe de travail du Réseau des Bibliothèques de l'Université de Moncton, mars 2011.  
[http://www.couperin.org/images/stories/documents/outils\\_decouverte\\_rapport\\_final\\_abrege\\_mars2011.pdf](http://www.couperin.org/images/stories/documents/outils_decouverte_rapport_final_abrege_mars2011.pdf)  
(avril 2012)

Claire DUGAST, « Utilisabilité des interfaces de recherche à facettes proposées par les OPAC de nouvelle génération », 1er décembre 2011.  
[http://www.ressi.ch/num12/article\\_074](http://www.ressi.ch/num12/article_074)  
(avril 2012)

Maureen KELLY et Judy LUTHER, « The next generation of discovery », 15 mars 2011.  
[http://www.libraryjournal.com/lj/home/889250-264/the\\_next\\_generation\\_of\\_discovery.html.csp](http://www.libraryjournal.com/lj/home/889250-264/the_next_generation_of_discovery.html.csp)  
(avril 2012)

Marc MAISONNEUVE, « Les OPAC nouvelle génération », 4 mars 2009.  
<http://www.toscaconsultants.fr/articles/opacsdenouvellegenerationADBSGrenoble.pdf>

« OAI-PMH : description du protocole », 2 août 2011  
[http://www.bnf.fr/fr/professionnels/protocoles\\_echange\\_donnees/a.proto\\_oai.html](http://www.bnf.fr/fr/professionnels/protocoles_echange_donnees/a.proto_oai.html)  
(avril 2012)

---

<sup>30</sup>Étude réalisée dans le cadre du projet de la BOMLR, qui a mené à la mise en place du méta-moteur AskOnce pour la BOMLR en 2003.

# Annexes

---

## *Table des annexes*

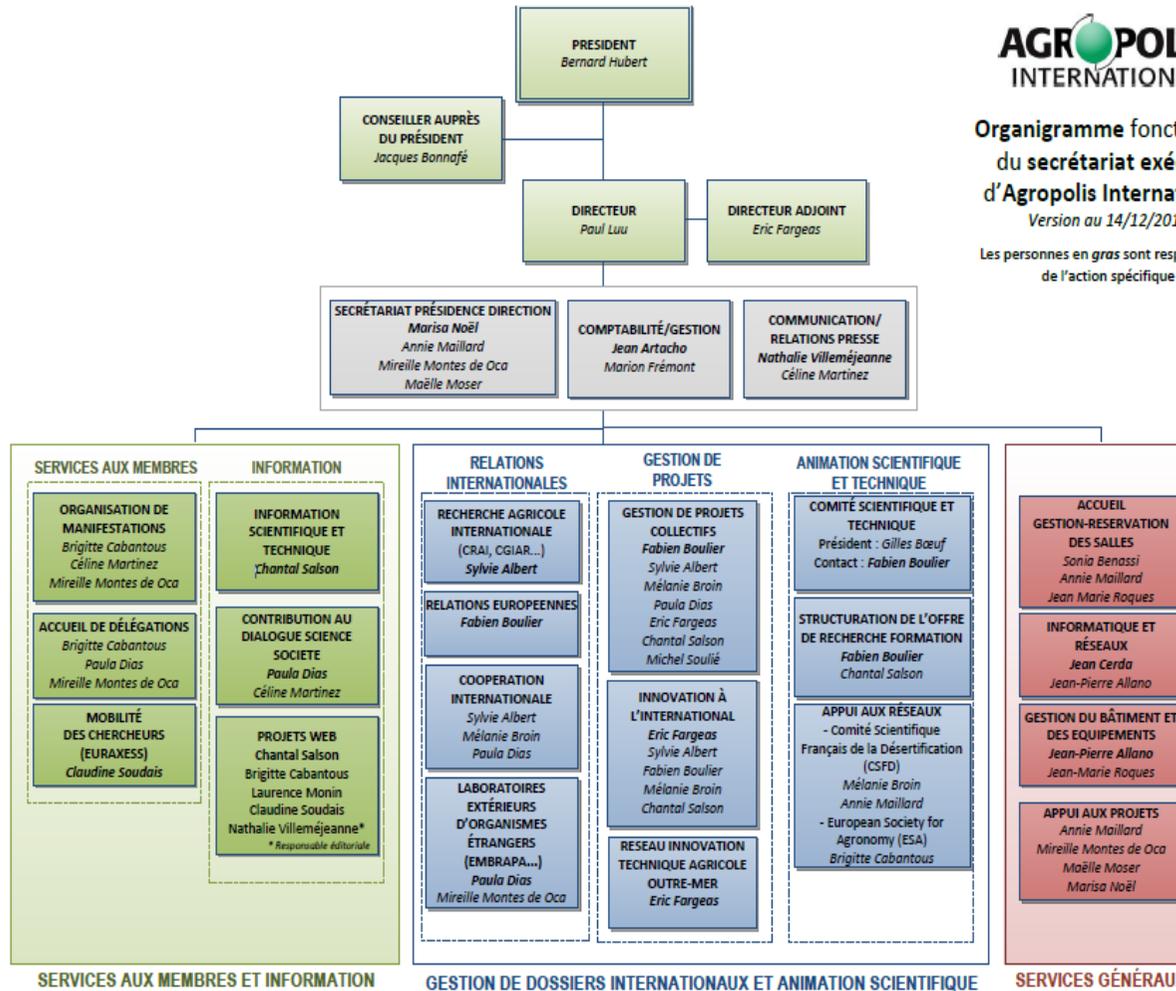
<b>ANNEXE 1 – ORGANIGRAMME (DÉCEMBRE 2011).....</b>	<b>48</b>
<b>ANNEXE 2 – PLAQUETTE AGROPOLIS (MAI 2012).....</b>	<b>49</b>
<b>ANNEXE 3 – PLAQUETTE CIST (MAI 2009).....</b>	<b>51</b>
<b>ANNEXE 4 - PLANNING DE TRAVAIL.....</b>	<b>53</b>
<b>ANNEXE 5 – SCHÉMA DE FONCTIONNEMENT D'UN MÉTA-MOTEUR .....</b>	<b>55</b>
<b>ANNEXE 6 - CAPTURES D'ÉCRAN DE LA BOMLR ACTUELLE.....</b>	<b>56</b>
<b>ANNEXE 7 – BASES ACTUELLEMENT INTERFACÉES PAR LA BOMLR.....</b>	<b>61</b>
<b>ANNEXE 8 - STATISTIQUES DE CONSULTATION DE LA BOMLR ENTRE 2004 ET 2011.....</b>	<b>63</b>
<b>ANNEXE 9 – QUELQUES EXEMPLES DE COMPTES-RENDUS DE RÉUNION DU GROUPE PROJET.....</b>	<b>65</b>
<b>ANNEXE 10 - GRILLE D'ENTRETIEN (ÉTAT DES LIEUX DES BASES DE DONNÉES).....</b>	<b>71</b>
<b>ANNEXE 11 - QUESTIONNAIRE (ANALYSE DES BESOINS).....</b>	<b>73</b>
<b>ANNEXE 12 - RÉSULTATS DU QUESTIONNAIRE .....</b>	<b>75</b>
<b>ANNEXE 13 – MINI CAHIER DES CHARGES.....</b>	<b>80</b>
<b>ANNEXE 15 - CAPTURES D'ÉCRAN DE VUFIND INSTALLÉ SUR LE SERVEUR D'AGROPOLIS (PHASE DE TESTS).....</b>	<b>90</b>

# Annexe 1 – Organigramme (décembre 2011)



Organigramme fonctionnel  
du secrétariat exécutif  
d'Agropolis International  
Version au 14/12/2011

Les personnes en **gras** sont responsables  
de l'action spécifique





- 28 établissements de recherche et d'enseignement supérieur
- 2 300 chercheurs et enseignants dont 400 en poste dans une soixantaine de pays partenaires
- 80 unités de recherche majoritairement inter-institutionnelles et inter-disciplinaires
- 8 campus dont 3 dédiés aux sciences agronomiques, dotés d'infrastructures mutualisées : Écotron, Maisons communes (Technologie, Télé-détection, Sciences de l'Homme, Sciences de l'Eau), serres expérimentales, halles pilotes, incubateurs d'entreprises...
- 3 pôles de compétitivité : Eau, Gestion des Risques et vulnérabilités des territoires, Qualité d'Étirement
- Un réseau de 43 centres documentaires
- 5 Universités et 9 Écoles d'ingénieurs offrant :
  - + 156 parcours de formation (de Bac +2 à Bac +8 : technicien, ingénieur, licence, master, doctorat...) accueillant plus de 3 000 étudiants dont 300 doctorants par an (1/3 provenant de pays étrangers)
  - + 100 modules de formation continue et une capacité d'ingénierie de formation



Au sein de la communauté scientifique de **Montpellier** et de la **région Languedoc-Roussillon**,

- la plus grande concentration mondiale de compétences et d'expertise dans les domaines liés à l'agriculture, à l'alimentation, à la biodiversité et à l'environnement, ouverte sur le développement des régions méditerranéennes et tropicales

- une forte dimension internationale avec le siège du *Consortium* du CGIAR, l'implantation d'institutions internationales et de laboratoires étrangers (Argentine, Australie, Brésil, USA...), l'appui à des réseaux : Société Européenne d'Agronomie, Comité Scientifique Français de la Désertification...

- une fondation pour la recherche agronomique et le développement durable :

Agropolis Fondation



- plusieurs laboratoires d'excellence (LabEx sélectionnés dans le cadre du programme Investissements d'Avenir du ministère chargé de la Recherche)

Un large éventail de **compétences**

- Agronomie, plantes cultivées et systèmes de culture, agroécosystèmes
- Alimentation, nutrition, santé
- Biodiversité et écosystèmes aquatiques
- Biodiversité et écosystèmes terrestres
- Eau, ressources et gestion
- Économie, société et développement durable
- Ecotechnologies
- Interactions hôte-parasites et maladies infectieuses
- Modélisation, information géographique, biostatistiques
- Production et santé animales
- Ressources génétiques et biologie intégrative des plantes
- Vigne et vin

...mobilisées sur des **grands thèmes de recherche liés aux enjeux de sociétés**

- Alimentation durable et sécurité des aliments
- Gestion des ressources, environnement et développement durable
- Innovation, usages, transfert des technologies
- Nutrition, santé
- Productions agricoles et filières de transformation et de commercialisation
- Risques et vulnérabilité des territoires



Une offre de **formation complète et diversifiée**

- Formations initiales diplômantes à visée professionnelle ou de recherche, pour tous niveaux et domaines de compétences
- Formations continues pour les professionnels, diplômants ou non
- Ingénierie des dispositifs de formation



Au sein de la communauté scientifique de Montpellier et de la région Languedoc-Roussillon

Diffusion des savoirs et innovations



Agropolis International, **association fondée et pilotée** par les **institutions de la communauté scientifique régionale**

interface internationale  
coordination  
animation

facilitation  
espace multi-acteurs  
maison commune

Un portail d'accès

- Un accès facilité aux compétences et ressources de la communauté scientifique
- Un accès à des supports d'information : bulletins périodiques, sites web, portails thématiques, synthèses des connaissances scientifiques sur les thèmes et problématiques d'actualité...
- Un appui aux projets collectifs interinstitutionnels à caractère national et international (animation, coordination, gestion)
- Un service d'accueil de délégations
- Une assistance pour les chercheurs internationaux (Centre Euraxess Languedoc-Roussillon)
- Un service pour l'organisation de réunions et de conférences
- Une interface avec les acteurs du développement
- Une plateforme de construction et de coordination de nouveaux partenariats

L'Association Agropolis International au service de la communauté scientifique

# Annexe 3 – Plaquette CIST (mai 2009)

Plus de 40 bibliothèques et centres de documentation

## Centres de documentation de départements et unités de recherche\*

- **AMAP – Botanique et Bioinformatique de l'Architecture des Plantes**  
Montpellier  
Contact : Yannick Brohard (yannick.brohard@ciad.fr)
- **ART-Dev : Acteurs, Ressources et Territoires dans le Développement (ex. MTE)**  
Montpellier  
Contact : Sylvie Hammiel (sylvie.hammiel@univ-montp3.fr)
- **BGPI – Biologie et Génétique des Interactions Plantes-Parasites**  
Montpellier  
Contact : Marie-Carmen Martinez (marie-carmen.martinez@ciad.fr)
- **BPMP – Biochimie et Physiologie Moléculaire des Plantes**  
Montpellier  
Contact : Chantal Baracco (baracco@supagro.inra.fr)
- **CEFE – Centre d'écologie Fonctionnelle et Évolutive – Montpellier**  
Contact : Anne Corgeon (bibcofe@cefe.cnrs.fr)
- **CBAE – Centre de Bio-Archéologie et d'écologie**  
Montpellier  
Contact : Nicole Denis-Touzillier (denis.tou@univ-montp2.fr)
- **CEBTM – Centre de Biologie et d'écologie Tropicale et Méditerranéenne**  
Perpignan  
Contact : Anne Modat (anne.modat@univ-perp.fr)
- **Centre de documentation Pédagogique et Scientifique de Géographie de l'Université Montpellier 3 – Montpellier**  
Contact : Sylvie Deveze (sylvie.deveze@univ-montp3.fr)

\* Les unités de recherche sont constituées d'équipes provenant des établissements suivants : Ciad, CIHEAM-ILAM, Cys, EPHE, Inra, IRD, Montpellier SupAgro, Université de Montpellier 1-2-3, Université de Perpignan via Domitia.

- **Centre Pierre Bartoli - Montpellier** Contact : Caroline Falco (falco@supagro.inra.fr)
- **LAMETA : Laboratoire Montpellierain d'Économie Théorique et Appliquée**
- **MOGA : Marchés, organisations, institutions et stratégies d'acteurs**
- **Département Sciences économiques, sociales et de gestion de Montpellier SupAgro**
- **CILBA - Complexe International de Lutte Biologique Agropolis – Balaruc**  
Contact : Christine Silvy (silvy@supagro.inra.fr)
- **CBGP : Centre de Biologie et de Gestion des Populations**
- **USDA/ARS : United States Department of Agriculture/ Agricultural Research Service/EBCL : European Biological Control Laboratory**
- **CSIRO European Laboratory : Commonwealth Scientific and Industrial Research Organisation**
- **Géosciences Montpellier - Montpellier**  
Contact : Sylvie Reynaud (sylvie.reynaud@univ-montp2.fr)
- **INNOVATION : Innovation et développement dans l'agriculture et l'agro-alimentaire – Montpellier**  
Contact : Catherine Aiquier (caiquier@supagro.inra.fr)
- **ISEM : Institut des sciences de l'Évolution de Montpellier – Montpellier**  
Contact : Christine Bibal (bibal@sem.univ-montp2.fr)
- **LBE – Laboratoire de Biotechnologie de l'Environnement – Narbonne**  
Contact : Veronique Maugeot (veronique.maugeot@supagro.inra.fr)
- **MTD : Maison de la Télédétection**  
Montpellier  
Contact : Sylvie Bill-Sarah (sbil@teledetection.fr)
- **MSE : Maison des Sciences de l'Eau**  
Montpellier  
Contacts : Catherine Ploz / Joëlle O'Sullivan (osullivan@isem.univ-montp2.fr et ploz@isem.univ-montp2.fr)

## Services d'édition

Les services édition des établissements sont également membres de la CIST Agropolis International :

- CNRS Éditions
- Éditions de AgroParisTech/Engref
- Éditions de l'IRD
- Éditions de Montpellier SupAgro
- Éditions du BRGM
- Éditions Quae (Carnegat, Ciad, Ithex, Inra)
- Publications du CIHEAM
- Publications de l'université Montpellier 3 : Presses universitaires de la Méditerranée

Agriculture - Agronomie  
Agroalimentaire Eau Forêt et foresterie Environnement  
Biodiversité Développement durable Sciences de la Vie  
Sciences exactes et de l'ingénieur Alimentation  
Sciences humaines et sociales Sciences de la Terre  
Zones méditerranéennes et tropicales



Montpellier Languedoc-Roussillon  
Campus mondial des sciences vertes

Agropolis International  
Avenue Agropolis  
F-34394 Montpellier Cedex 5

Contact : Chantal Sabon  
sabon@agropolis.fr

## Sites internet

- Agropolis International  
▶ [www.agropolis.fr](http://www.agropolis.fr)
- IST Agropolis  
▶ [www.ist.agropolis.fr](http://www.ist.agropolis.fr)
- Conférence IAALD 2010  
▶ [iaald2010.agropolis.fr](http://iaald2010.agropolis.fr)
- BOMLR  
▶ [www.bomlr.info](http://www.bomlr.info)

## L'information scientifique et technique (IST) : un enjeu pour une formation et une recherche d'excellence



L'IST au sein du pôle scientifique agro-environnement de Montpellier et sa région :

- 250 professionnels de l'information et de la documentation
- plus de 40 centres de documentation et services d'édition
- des fonds documentaires (plus de 1,5 million de documents)
- un réseau et des outils au service des étudiants, enseignants, chercheurs, doctorants de la communauté régionale et internationale



Agriculture  
Alimentation  
Biodiversité  
Environnement



Montpellier Languedoc-Roussillon  
Campus mondial des sciences vertes

[www.ist.agropolis.fr](http://www.ist.agropolis.fr)

Plus de 40 bibliothèques et centres de documentation

## AGROPOLIS INTERNATIONAL

### En chiffres

- 12 établissements de recherche nationaux et internationaux
- 7 écoles supérieures
- 5 universités
- 92 unités de recherche
- 80 formations diplômantes (professionnelles ou recherche)
- 2300 chercheurs et enseignants
- 4000 étudiants et stagiaires
- 1400 techniciens et administratifs

### En bref

- un portail d'accès aux compétences de la communauté de Montpellier et de la région Languedoc-Roussillon dans les domaines liés à l'agriculture, à l'environnement, à la biodiversité, à l'alimentation, à la biodiversité, à l'environnement et aux sociétés rurales, ouvert sur le développement des régions méditerranéennes et tropicales
- une interface au cœur des réseaux de coopération scientifique internationale pour appuyer des projets collectifs
- une plate-forme pilote pour développer des rapprochements inter-institutionnels et de nouveaux modes d'organisation
- un service d'accueil pour les chercheurs, professeurs invités et délégations

## La Commission IST Agropolis

Les responsables des centres de documentation et des services d'édition des membres d'Agropolis International se réunissent au sein de la Commission « Information Scientifique et Technique » créée en 1985.

La commission IST met en œuvre des services et actions communes visant à faciliter la recherche d'information scientifique et technique pour la communauté scientifique régionale et internationale.

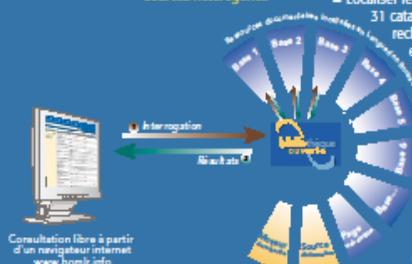
### Actions phares

- Mise en ligne d'un nouveau portail IST en avril 2008 permettant la localisation des ressources et compétences IST en région Languedoc-Roussillon  
► [www.ist.agropolis.fr](http://www.ist.agropolis.fr)
- Mise en place dès 2003 d'une interface d'interrogation unique des catalogues des centres de documentations et bibliothèques des organismes membres d'Agropolis International, la Bibliothèque Ouverte Montpellier-Languedoc-Roussillon  
► [www.bomlr.info](http://www.bomlr.info)
- Organisation de Journées d'étude et d'information destinées aux scientifiques et aux professionnels de l'information
- Mise en place de formations mutualisées à l'Information Scientifique et Technique, destinées aux scientifiques et étudiants
- Création d'une vitrine commune des éditions institutionnelles des membres d'Agropolis International
- Préparation du XIII<sup>ème</sup> congrès international de l'IAALD sur la thématique « IST et développement rural » qui se déroulera à Montpellier en 2010  
► [iaal2010.agropolis.fr](http://iaal2010.agropolis.fr)

**Portail Information Scientifique et Technique (IST)**  
une plate-forme collaborative au service des communautés régionale et internationale



**La BOMLR - Bibliothèque Ouverte Montpellier Languedoc-Roussillon**  
un méta-moteur documentaire pour interroger simultanément des sources hétérogènes



[www.ist.agropolis.fr](http://www.ist.agropolis.fr)

- **Agenda** : Informer sur les formations, réunions, événements
- **BOMLR** : Rechercher et localiser un document ou une revue
- **Bases de données** : Rechercher et localiser une ressource électronique
- **Centres de documentation** : Découvrir et accéder aux centres de documentation et bibliothèques (contacts, composition des fonds, horaires ...)
- **Aide à la recherche d'informations** : Guider l'utilisateur dans sa recherche documentaire
- **Accéder aux catalogues des éditions et aux catalogues des bibliothèques**
- **Des rubriques réservées aux professionnels de l'IST pour partager, archiver des ressources** → Commission IST et Actions en cours

[www.bomlr.info](http://www.bomlr.info)

- Localiser les documents référencés dans 31 catalogues de bibliothèques de recherche et de formation supérieure en Languedoc-Roussillon
- Localiser les revues consultables dans ces bibliothèques
- Consulter les publications scientifiques et travaux d'étudiants (prochainement)
- Interroger simultanément toutes ces ressources documentaires et étendre la recherche à des sources externes (prochainement).

La Commission IST Agropolis

Au sein de la communauté scientifique de Montpellier et de la région Languedoc-Roussillon

## Plus de 40 bibliothèques et centres de documentation

► Des fiches descriptives de ces bibliothèques sont consultables sur le portail IST : [www.ist.agropolis.fr](http://www.ist.agropolis.fr)

### Bibliothèques et centres de documentation des établissements

- **AgroParis Tech/Engref** - Montpellier  
Contact : Claude Maire ([claudie.maire@engref.agroparis-tech.fr](mailto:claudie.maire@engref.agroparis-tech.fr))
- **Bioversity International-France** - Montpellier  
Contact : Claudine Pico ([c.pico@bioversity.org](mailto:c.pico@bioversity.org))
- **CEA Marcoule** - Bagnols-sur-Cèze  
Contact : Jacques Caillaud ([jacques.caillaud@cea.fr](mailto:jacques.caillaud@cea.fr))
- **Cemagref Montpellier** - Montpellier  
Contacts : Carole Clardy, Mylène Tassou ([doc@montpellier.cemagref.fr](mailto:doc@montpellier.cemagref.fr))
- **CIRAD-IAMM** : Institut Agronomique Méditerranéen de Montpellier - Montpellier  
Contact : Marie-Louise Loderic ([mloderic@iamm.fr](mailto:mloderic@iamm.fr))
- **Cirad** : Centre de Coopération Internationale en Recherche Agronomique pour le Développement - Montpellier  
Contact : Marie-Claude Debolin ([ist@cirad.fr](mailto:ist@cirad.fr))
- **EMA** : École des Mines d'Alès - Nîmes et Alès  
Contact : Françoise Armand ([ista-doc@ema.fr](mailto:ista-doc@ema.fr))
- **ENSCM** : École Nationale Supérieure de Chimie de Montpellier Montpellier  
Contact : Veslko Hulka ([veslko.hulka@enscm.fr](mailto:veslko.hulka@enscm.fr))
- **Ifremer** : Institut français de recherche pour l'exploitation de la mer - Station de Palavas-les-Flots  
Contact : Marie-Joséphine Debos ([marie.josephine.debos@ifremer.fr](mailto:marie.josephine.debos@ifremer.fr))
- **Inra** - Centre de Montpellier : Institut National de la Recherche Agronomique Montpellier  
Contact : Dominique Fournier ([fournierd@supagro.inra.fr](mailto:fournierd@supagro.inra.fr))
- **Institut de Botanique** - Service des herbiers Université Montpellier 2 - Montpellier  
Contact : Peter A. Schäfer ([psch@bot.umlv-montp2.fr](mailto:psch@bot.umlv-montp2.fr))
- **IRD** : Institut de Recherche pour le Développement - Montpellier  
Contact : Hanka Horsens ([hanka.horsens@ird.fr](mailto:hanka.horsens@ird.fr))
- **Lycée Frédéric Bazille** - Agropolis - Montpellier  
Contact : Sylvie Bourgeois ([sylviebourgeois@educagri.fr](mailto:sylviebourgeois@educagri.fr))
- **Montpellier SupAgro** - Bibliothèque du Centre international d'études supérieures en sciences agronomiques de Montpellier Montpellier  
Contact : Dominique Fournier ([fournierd@supagro.inra.fr](mailto:fournierd@supagro.inra.fr))
- **Montpellier SupAgro** - Institut des régions chaudes - Montpellier  
Contact : Christiane Rousso ([christianerousso@supagro.inra.fr](mailto:christianerousso@supagro.inra.fr))
- **Florac SupAgro** - Florac  
Contact : Hélène Lazzarini ([helene.lazzarini@educagri.fr](mailto:helene.lazzarini@educagri.fr))
- **Sup de Co Montpellier** - Montpellier  
Contact : Isabelle Bresson ([ibresson@supco-montpellier.fr](mailto:ibresson@supco-montpellier.fr))
- **Universités de Montpellier** : Bibliothèque Interuniversitaire de Montpellier (BIU)  
Contact : Jean-François Foucaud ([biu@ccr.univ-montp1.fr](mailto:biu@ccr.univ-montp1.fr))
- **Bibliothèque Sciences** - Montpellier  
Contact : Philippe Valera ([phvalera@univ-montp2.fr](mailto:phvalera@univ-montp2.fr))
- **Bibliothèque Lettres et Sciences Humaines** - Montpellier  
Contact : Huguetto Briker ([huguetto@univ-montp1.fr](mailto:huguetto@univ-montp1.fr))
- **Bibliothèque Droit et Sciences Économiques** - Montpellier  
Contact : Stéphane Lanoë ([stlano@univ-montp1.fr](mailto:stlano@univ-montp1.fr))
- **Bibliothèque Pharmacie** Montpellier  
Contact : Elisabeth Blanes ([elblanes@univ-montp1.fr](mailto:elblanes@univ-montp1.fr))
- **Université de Nîmes** : bibliothèque Universitaire - Nîmes  
Contact : Catherine Muzson ([camus@univ-nimes.fr](mailto:camus@univ-nimes.fr))
- **Université de Perpignan Via Domitia** : Service commun de documentation - Perpignan  
Contact : Joël Martin ([joel.martin@univ-perp.fr](mailto:joel.martin@univ-perp.fr))

Plus de 40 bibliothèques et centres de documentation

## Annexe 4 - Planning de travail

(S = semaine)

Tâches	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18
A																		
A1																		
A2																		
B																		
C																		
C1																		
C2																		
D																		
D1																		
D2																		
D3																		
E																		
E1																		
E2																		
E3																		
F																		
G																		
H																		

A : Interrogation des bases de données interfacées sur la BOMLR

A1 : Test des catalogues d'ouvrages, rapports, etc. (partie « documents »)

A2 : Test des catalogues de revues (partie « revues »)

B : Panorama des centres de documentation / bases potentielles à interfacier

C : État des lieux des bases de données à intégrer sur la future interface (recueil des caractéristiques documentaires et techniques)

C1 : Recueil des caractéristiques documentaires et techniques des bases de données auprès des documentalistes et des informaticiens des différents établissements

C2 : Synthèse des résultats

D : Analyse des besoins

D1 : Entretiens avec les documentalistes

D2 : Enquêtes en ligne auprès de l'ensemble des publics

D3 : Synthèse des résultats

E : Étude des outils de recherche

E1 : Panorama des solutions

E2 : Étude comparative (après première sélection, et selon une liste de critères établis avec le groupe-projet)

E3 : Synthèse des résultats

### Présentation devant la CIST

F : Formalisation des exigences et contraintes d'Agropolis (rédaction et envoi d'un rapport aux éditeurs et utilisateurs des solutions retenues)

G : Exploration et tests des solutions

H : Rédaction du mémoire de stage



## Annexe 5 – Schéma de fonctionnement d'un méta-moteur

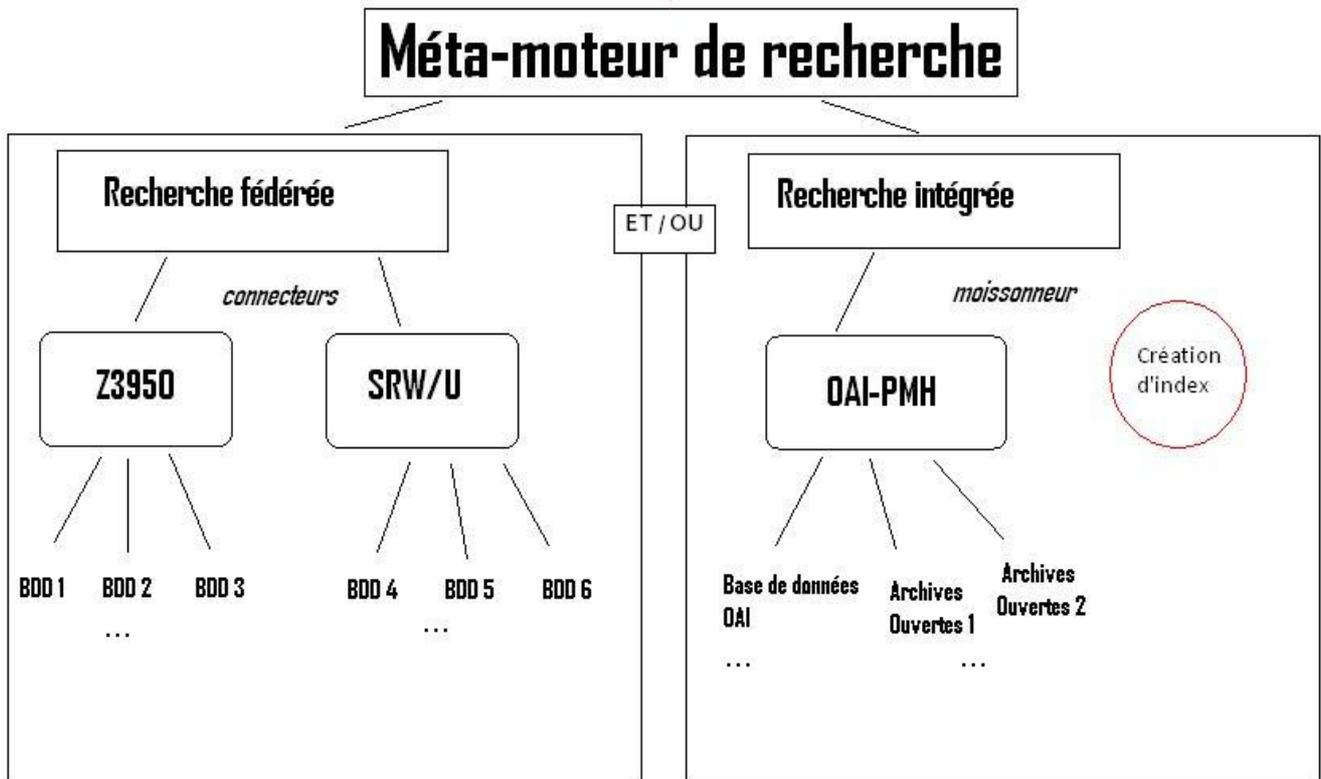


Schéma de fonctionnement d'un méta-moteur de recherche

## Annexe 6 - Captures d'écran de la BOMLR actuelle

Capture 1 : page d'accueil

**bibliothèque ouverte**

# Bibliothèque Ouverte Montpellier Languedoc-Roussillon

Portail de recherche fédérée des **ressources documentaires**  
des établissements de recherche et d'enseignement supérieur en **Languedoc-Roussillon** English

**La Bibliothèque Ouverte Montpellier Languedoc-Roussillon (BOMLR)** permet d'interroger simultanément **1 200 000 notices** issues de **48 unités documentaires**

**Zoom**

**AgroParisTech-Engref Montpellier**

Bibliothèque présente sur le campus de Lavalette  
648 rue Jean François Breton -  
B.P. 7355  
34086 Montpellier Cedex 4

**Annabelle Filatre**  
E-mail | Catalogue  
Fiche descriptive

**Choisissez votre interface pour**

- Localiser **les documents** référencés dans les catalogues
- Localiser **les revues** consultables dans les bibliothèques
- Prochainement:** Interroger simultanément **toutes** ces ressources documentaires

**Choisissez une unité documentaire pour en savoir plus...**

- AgroParisTech-Engref Montpellier • BIU Montpellier Droit-économie-gestion
- BIU Montpellier Lettres-Sciences Humaines • BIU Montpellier Médecine (centre ville)
- BIU Montpellier Médecine Nîmes • BIU Montpellier Pharmacie • BIU Montpellier Sciences
- BIU Montpellier Unité Pédagogique Médicale
- CEFE : Centre d'Ecologie Fonctionnelle et Evolutive (CNRS UMR 5175) • Cemagref
- CIHEAM-IAMM - Institut Agronomique Méditerranéen de Montpellier
- Cirad : Centre de Coopération Internationale en Recherche Agronomique pour le Développement
- DRAC : Direction Régionale des Affaires Culturelles du Languedoc-Roussillon
- EMA : Ecole des Mines d'Alès • INRA - Centre de documentation Pierre Bartoli
- INRA - Centre de Montpellier • IRD - Institut de Recherche pour le Développement
- Montpellier SupAgro - Bibliothèque de la Gaillarde • Montpellier SupAgro - Institut des régions chaudes

## Capture 2 : interface de recherche

The screenshot shows a research interface with a blue header and a yellow content area. The header contains navigation tabs: "Login", "Domaines", "Préférences", and "Rech. avancée". Below these are icons for "Agro-Env", "Sci.Hum.", "Sci.Exactes", "Terre", and "Bio-Santé". A search bar with a "Recherche" button and a "Stop" button is also present. The main content area is divided into several sections: "Interrogation simultanée des : catalogues de références bibliographiques, répartis en domaines scientifiques ou centres d'intérêt" with icons for the same domains; a "FAQ" section with the text "Réponse à vos questions : qui peut se connecter, comment interroger ..."; an "Actu" section with a "Support d'interrogation" link and a "Plaquette à imprimer pour découvrir l'interface" link; a "Votre avis nous intéresse" link; a "Quelques exemples" section with a "Recherche simple" link and a search bar containing "développement durable"; and a "UNE INTERROGATION EN TROIS ÉTAPES" call to action.

Login Domaines Préférences Rech. avancée

Agro-Env Sci.Hum. Sci.Exactes Terre Bio-Santé

Recherche Stop

Etat : Pa:

Interrogation simultanée des :

catalogues de références bibliographiques,  
répartis en domaines scientifiques ou centres d'intérêt

FAQ Réponse à vos questions : qui peut se connecter, comment interroger ...

Actu

Support d'interrogation

Plaquette à imprimer pour découvrir l'interface : Interroger simultanément les ressources documentaires des organismes de recherche et d'enseignement supérieur du Languedoc-Roussillon, c'est facile ...

Démo

Votre avis nous intéresse

cliquer sur l'onglet "Domaines" pour voir un descriptif des bases

UNE INTERROGATION EN TROIS ÉTAPES

Quelques exemples

Recherche simple (bases documentaires + web en option)

Sci. Hum Agro-Env Lettres Sci. Exactes Terre Bio-Santé Revues WWW

développement durable Recherche

Capture 3 : exemple de recherche 1 (test avec le terme « karité »)

The screenshot shows a search interface for 'bibliothèque ouverte'. The search term 'karité' is entered in the search bar. The results are displayed in a table with columns for 'Titre' and 'Contenu'. Two results are shown, both from 'CEFE-CNRS UMR5175'. The interface also includes a sidebar with 'Rubriques par: Thèmes' and a message 'Nombre insuffisant de résultats pour calculer les rubriques'. At the bottom, there is a table showing the number of results for various institutions.

Requête en 'Français'. 'Requête-1' Etat : 2 résultat(s)

Rubriques par: Thèmes	Titre	Contenu	Score
	Aspects du parc à Karites-Neres (Butyrospermum parkii, Parkia biglobosa) d...	DEPOMMIER D ; FERNANDES E ; ICRAF ; / 1985 / CEFE-CNRS UMR5175 [ Aperçu ]	☆☆☆ + =
	Monographie du <b>Karité</b> , Butyrospermum paradoxum (Gaertn.f.) Hepper, espèce a...	BONKOUNGOU EG ; / 1987 / CEFE-CNRS UMR5175 [ Aperçu ]	☆☆☆ + =

Résultat(s) reçus : 2

<b>CEFE-CNRS UMR5175</b>	2 résultat(s)	<b>Cirad-base Agritrop</b>	0 résultat(s)
<b>CIHEAM-IAM.M</b>	0 résultat(s) [ échec ]	<b>IRD-catalogue</b>	0 résultat(s) [ Requête stoppée ]
<b>Musalit - Bioersity International</b>	0 résultat(s)	<b>Sup Agro Montpellier</b>	0 résultat(s)

Nombre insuffisant de résultats pour calculer les rubriques

Capture 4 : exemple de recherche 2 (test avec le terme « agro-environnement »)

The screenshot shows a library search interface with a blue header. The search term 'agro-environnement' is entered in the search bar. The results are displayed in a table with columns for 'Titre' and 'Contenu'. The interface includes navigation buttons like 'Login', 'Domaines', 'Préférences', and 'Rech. avancée'. The 'bibliothèque ouverte' logo is visible in the top right corner. The search results are sorted by 'Requête-1' and show a list of documents with their titles, authors, and scores.

Rubriques par: Thèmes	Titre	Contenu	Score
Tous les résultats Actualiser	La performance environnementale de l'agriculture dans les pays de l'OCDE de...	Organisation de Coopération et de Développement Economiques [OCDE] ; / 2008 / CEFÉ-CNRS UMR5175 [ Aperçu ]	☆☆☆
+ agro alimentaire et environnement, système (11)	Etude agro-environnementale intégrée du marais poitevin	Bouché MB ; AUPÉTIT P ; PHILIPPE MA ; / 1988 / CEFÉ-CNRS UMR5175 [ Aperçu ]	☆☆☆
+ agriculture, protection (8)	Quel marché du travail en agriculture, en agro-alimentaire et en environnem...	métier / Rapiou Marie-Thérèse,Stéphan Nelly / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	La protection de l'environnement dans la stratégie des industriels agro-alim...	Strategie de l'entreprise / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	Plan Bleu. Perspectives du système agro-alimentaire et environnement en Méd...	Région méditerranéenne / Labonne M.,Royer J. / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	Perspectives du système agro-alimentaire et environnement en Méditerranée. ...	Maghreb / Mary Fabienne,Papayannakis Michel / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	Perspectives du système agro-alimentaire et environnement en Méditerranée R...	Pays méditerranéens / Labonne Michel,Royer Jacques / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	Perspectives du système agro-alimentaire et environnement en Méditerranée R...	Pays méditerranéens / Labonne Michel,Royer Jacques / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	Mise en place d'une politique environnement sur le site Nestle France de Ca...	Agro-environnement geed / Prost Aline / Sup Agro Montpellier [ Aperçu ]	☆☆☆
	Guide d'étude des systèmes de		☆☆☆

## Capture 5 : vue détaillée d'une notice



### Vue détaillée

Nom d'utilisateur : guest  
Requête : plein-texte:karité  
Sur les domaines : Agro-Env  
Résultats triés par : ordre d'arrivée

Fermer 

Trier par : titre , auteur , collection , date , sélection , source , score

Export 

Résultats	Sélection
titre	Aspects du parc à Karites-Neres (Butyrospermum parkii, Parkia biglobosa) dans la région de l'Ouham (République Centrafricaine) <input type="checkbox"/>
type_document	OUVRAGE
identifiant	22202
auteur	DEPOMMIER D ; FERNANDES E ; ICRAF ;
descripteur	PARC NATUREL / PARKIA BIGLOBOSA / REPUBLIQUE CENTRAFRICAINE /
cote	53071; rayon Z
descr. géographiques	18 p.
notes	Coll. Baumer
score	077

☆☆☆ source: CEFE-CNRS UMR5175 - date: 1985 -

## ***Annexe 7 – Bases actuellement interfacées par la BOMLR***

### 1. Accès « documents »

Bases interfacées interrogeables actuellement :

1. Bioersity International
2. BU Perpignan
3. CEFE-CNRS
4. Cirad
5. École des Mines d'Alès
6. IRD
7. Maison des Sciences de l'eau (via catalogue IRD)

Bases interfacées dysfonctionnelles – qui ne sont plus accessibles aujourd'hui - :

1. BIU Montpellier (dysfonctionnement momentané)
2. AgroParisTech – ENGREF (développement à venir)
3. CIHEAM-IAM.M (développement à venir)
4. UM3 Bibliothèques d'UFR (fusion avec BIU)
5. Montpellier SupAgro (en cours de développement)

### 2. Accès « revues »

Bases interfacées interrogeables actuellement :

1. Catalogue collectif des revues (papier) du réseau Agropolis (pas mis à jour)
2. Montpellier SupAgro/ Institut des Régions Chaudes (IRC)
3. CEFE-CNRS
4. Observatoire Océanologique de Banyuls

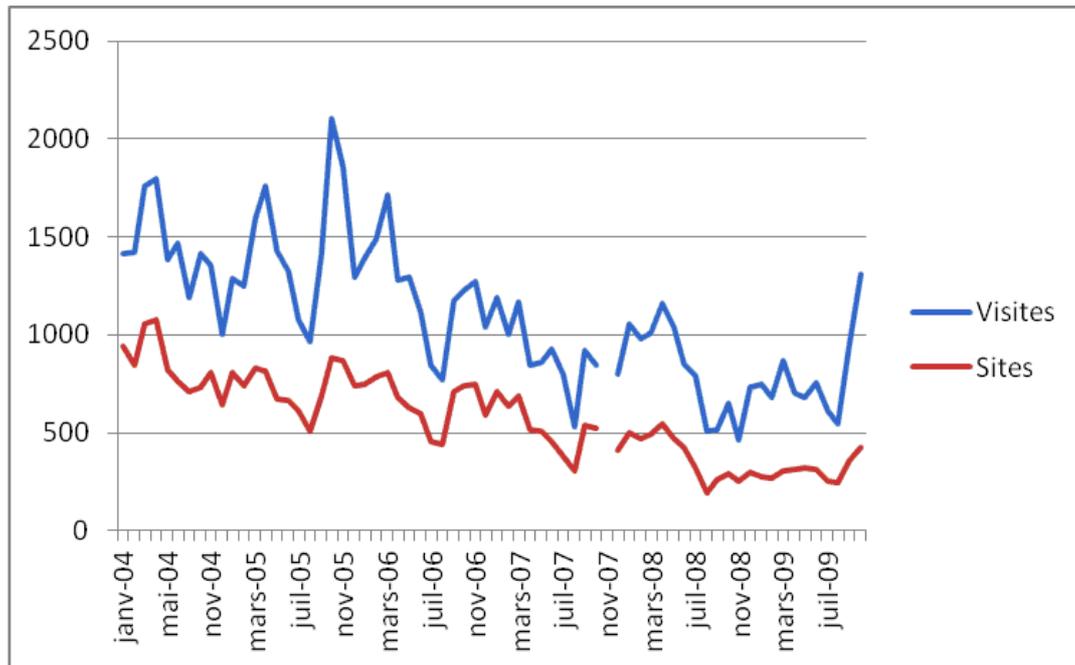
Bases interfacées dysfonctionnelles – qui ne sont plus accessibles aujourd'hui - :

1. Montpellier SupAgro/ INRA
2. Agro ParisTech
3. IRD
4. CRBM
5. CIHEAM.IAM
6. École des Mines d'Alès

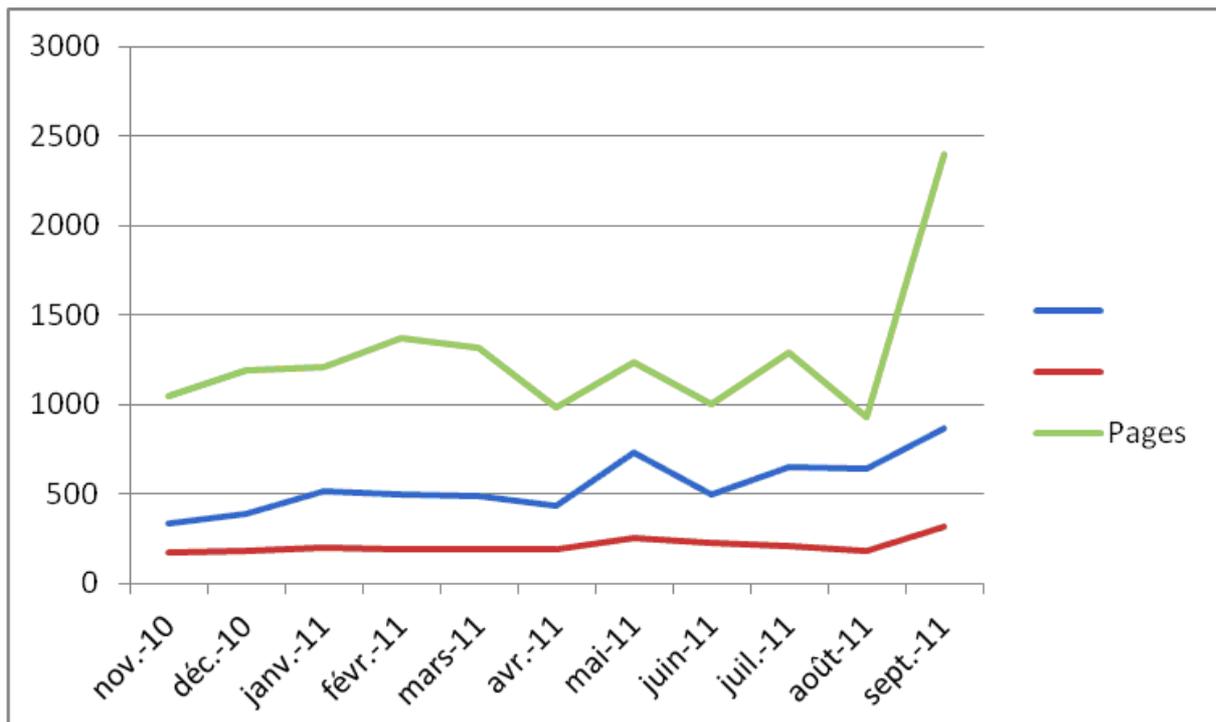
## **Annexe 8 - Statistiques de consultation de la BOMLR entre 2004 et 2011**

Étude réalisée par Chantal Salson, Agropolis International, octobre 2011

### Statistiques 2004 – 2009

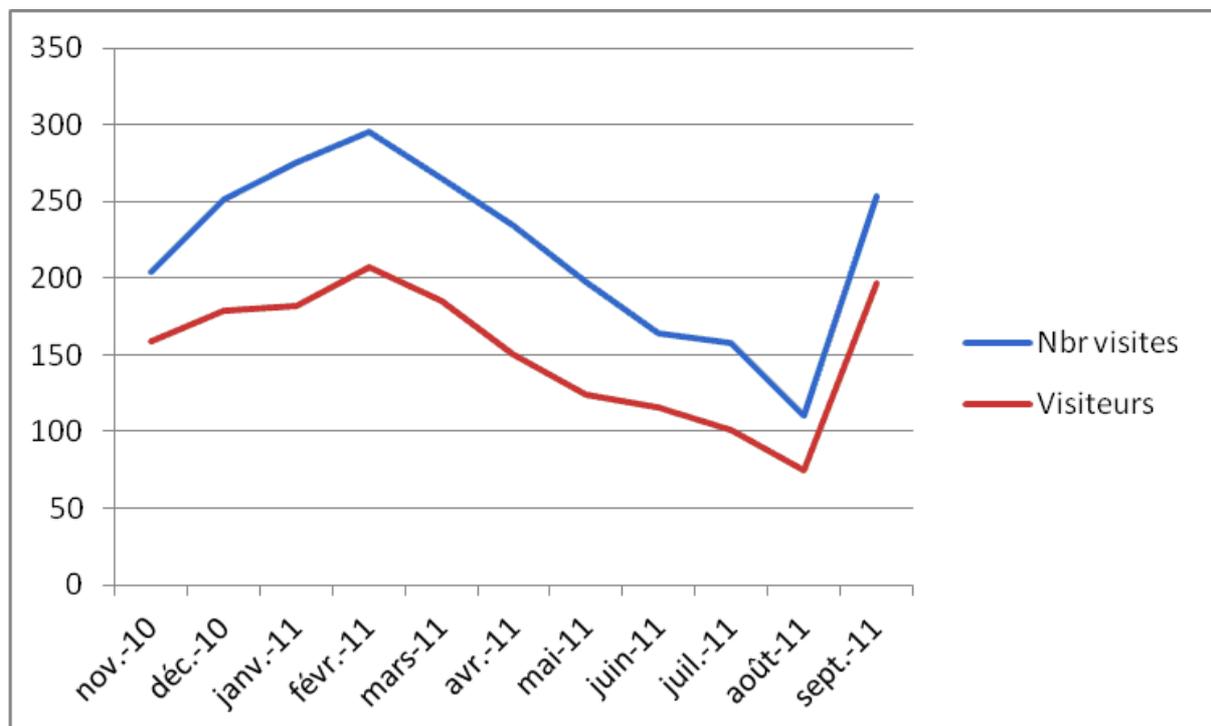


### Statistiques 2010-2011



Pour les années 2010 – 2011, on compte entre 335 et 869 visites par mois.

## Statistiques de consultation de la page d'accueil 2010 – 2011



Concernant la page d'accueil BOMLR, on compte entre 110 et 250 visites par mois.

# ***Annexe 9 – Quelques exemples de comptes-rendus de réunion du groupe projet***

## **Compte-rendu de la réunion BOMLR du 17 février 2012**

Lieu : Agropolis International

Participants : Bernadette Dupeuple, Annabelle Filatre, Marion Grand-Démery, Hanka Hensens, Chantal Salson, Rafaèle Sarrade

Ordre du jour : sélection des bases à interfacier et des organismes à contacter pour les enquêtes et/ou entretiens.

### 1. Rappel des bases actuellement interfacées par la BO-MLR

- La Bioersity International / Musalit
- La BIU Montpellier
- La BU Perpignan
- Le CEFE
- Le CIHEAM-IAM.M
- Le Cirad
- L'École des Mines d'Alès
- L'IRD
- La Maison des Sciences de l'eau (via IRD)
- Montpellier SupAgro
- L'Observatoire Océanologique de Banyuls
- L'UM3 Bibliothèques d'UFR

### 2. Bases à interfacier en priorité

- AgroParisTech
- L'AMAP
- Le BPMP
- Le CBGP

- L'ICRA
- L'INRA
- L'INSERM
- Le LBE
- Pierre Bartoli

### 3. Bases de priorité « P2 »

- Le BPMP
- Le BRGM
- Le CRBM (la base des revues est interfacée par AskOnce mais ne répond plus aux interrogations ; l'accès semble se faire désormais via la plateforme BiblioVie – INIST -)

Durant les quinze premiers jours du stage, il faudra établir un état des lieux des bases des ces différentes organisations en prenant en compte les points suivants :

- Les aspects techniques : outils utilisés, formats, protocoles de communication (incluant la question des perspectives d'évolution de ces outils, logiciels, etc.)
- Les différents types de ressources : publications, catalogues, revues papier et en ligne, documents en texte intégral en ligne,
- Le versement éventuel des différentes ressources dans des bases communes telles que le SUDOC, Hal, etc.
- L'intérêt pour l'interfaçage de ressources externes : bases thématiques en accès libre, etc.

La prochaine réunion se tiendra le vendredi 2 mars. L'heure et le lieu seront communiqués par mail.

Un compte-rendu du travail effectué durant la quinzaine sera envoyé aux participants au moins deux jours avant la réunion.

## Compte-rendu de la réunion BOMLR du 19 avril 2012

Lieu : Agropolis International

Participants : Annabelle Filatre, Marion Grand-Démery, Hanka Hensens, Chantal Salson.

Ordre du jour : exploration et test des solutions de recherche retenues

1 – Présentation de Vufind (libre) installé et testé sur le serveur d'Agropolis.

- Jean Cerda, du service informatique d'Agropolis, a pu installer la solution Libre VuFind sur le serveur d'Agropolis. Quelques tests ont été effectués avec une partie des ressources de l'IRD par moissonnage OAI. Hanka H. se renseigne pour savoir s'il est possible de sélectionner uniquement les ressources de Montpellier et/ou en ligne pour l'interfaçage des données.
- L'importation de données en format MARC a également été testée ; celle-ci fonctionne mais ralentit grandement le serveur (processus d'indexation assez lourd).
- La prochaine étape consiste à tester l'intégration des ressources des catalogues PMB. Annabelle F. doit se renseigner auprès de PMB pour savoir s'il est possible d'installer un serveur OAI pour permettre le moissonnage des données d'AgroParisTech via le protocole OAI-PMH.
- Il faut également voir si l'affichage de la source (localisation) peut se faire dès la page des résultats. Actuellement, la source n'est mise en évidence qu'à travers les filtres pour l'affinage des résultats.
- Pour aller encore plus loin, il est également prévu de tester la technologie des connecteurs. (mettre en place des passerelles Z3950 et/ou webservices)

L'interface est accessible à cette adresse : <http://aigroup.agropolis.fr/vufind/>

2 – Réponse du SICD de Grenoble pour LibraryFind (libre)

Lucie Albaret, du SICD de Grenoble, a évalué la compatibilité de la solution avec l'outil de recherche LibraryFind– en prenant en compte les développements effectués ou à venir dans le cadre de leur projet SURGIR - . Dans l'ensemble, l'outil correspond assez bien à nos attentes. Il répond à toutes nos attentes en matière de fonctionnalités de recherche et d'affichage des résultats (fonctionnalités prioritaires) et propose tous les modes d'interrogation requis. Mme Albaret précise également qu'il est « très simple de programmer l'indexation d'une source externe via une moisson ». Elle ajoute qu'il en est de même pour les connecteurs Z3950 et webservices, grâce à un « formulaire de création des connecteurs très simple d'utilisation ». Elle se renseigne auprès de Atos pour savoir si l'interfaçage de AtoZ est possible, mais *a priori*, les données pourraient être intégrées en format xml. Marion doit également se renseigner auprès de Francine Delmas (IRD) au sujet du catalogue AtoZ.

### 3 – Déplacements à Grenoble et à Paris / rencontre avec les utilisateurs des solutions libres

La question d'un déplacement à Grenoble et à Paris pour rencontrer les utilisateurs et étudier d'un peu plus près le fonctionnement des solutions a été soulevée. Il est question de programmer des rencontres avec ces utilisateurs au cours des prochaines semaines, et notamment avec le SCD de Grenoble, le SCD de Paris 3 et/ou Ecole des Ponts et Chaussées Paris-Tech.

### 4 – Remarques

- Au-delà des fonctionnalités et des technologies permettant l'interrogation à distance, le groupe a également insisté sur l'importance d'avoir des compétences « sur place » pour l'installation et la maintenance des solutions (maîtrise des langages, etc.)
- Le mémoire de stage de Marion doit être déposé à l'ENSSIB au plus tard le 11 juin prochain.

## Compte-rendu de la réunion BOMLR du 15 mai 2012

Participants : Françoise Armand, Annabelle Filatre, Marion Grand-Démery, Hanka Hensens, Chantal Salson, Rafaèle Sarrade.

Ordre du jour : exploration et test des solutions de recherche retenues

1 – Avancement des tests / exploration des différents outils :

- Vufind

*OAI-PMH*

Afin de continuer les tests de la solution libre, il est prévu de tester l'import OAI des ressources de l'École des Mines d'Alès ainsi que celui des ressources PMB d'AgroParisTech.

*Catalogue AtoZ*

Il faudra passer par Francine Delmas (IRD) pour demander à EBSCO de faire si possible un export en format MARC du catalogue AtoZ, l'idée étant aussi de le tester sur l'interface Vufind.

- LibraryFind

En attendant le déplacement au SICD de Grenoble, il s'agit de savoir si des compétences Ruby On Rails (le langage de développement de la solution) sont présentes au sein du réseau Agropolis, ainsi qu'au sein du CINES (Centre Informatique National de l'Enseignement Supérieur).

Une liste des bases de données actuellement interfacées par Grenoble doit également être constituée (bases testées avec succès).

- Summon

Rappel : démonstration de l'outil prévue au début du mois prochain, au cours de la semaine du 4 juin.

- Primo

Ex-Libris est à relancer, pour un retour suite à l'envoi du cctp. À voir aussi pour une démonstration de la solution d'ici la fin du stage de Marion.

2 – Estimation du coût / Chiffrage

Dans les semaines qui viennent, un chiffrage du projet devrait être réalisé, au moins pour les solutions libres, et en fonction des coûts estimés pour les développements, le serveur, l'installation, la maintenance.

### 3 – Valorisation du projet

Dans le cadre d'une première réflexion sur la présentation du projet en vue d'une demande de financement, le groupe-projet a insisté sur le fait de mettre en évidence trois aspects :

- La valorisation des publications et de la production scientifique en région
- La réutilisation du méta-moteur et de ses services par les différents établissements membres
- L'aspect « ressources pédagogiques »

# ***Annexe 10 - Grille d'entretien (état des lieux des bases de données)***

Nom de l'établissement :

## 1. Outils, normes, standards (aspects techniques)

- Quel logiciel utilisé ?
- Quel protocole de communication ? (HTTP Web Service, Z 3950, SRU/SRW, OAI-PMH, etc.)
- Formats et normes de catalogage ?

## 2. Types de ressources accessibles en ligne

- Catalogues d'ouvrages, d'articles, de revues (papier), etc.
- Revues électroniques
- Travaux d'étudiants
- Publications des chercheurs
- Autres : cartes, documentation iconographique, etc.
- Accès libre/réservé aux documents en texte intégral ?

## 3. Volumétrie (nombre de références)

## 4. Alimentation de bases communes du type archives ouvertes ou autres (SUDOC, CAIRN, HAL, AtoZ, theses.fr, etc.)

## 5. Perspectives d'évolution/ projets ? (migration données/ réinformatisation, versement sur bases précédemment citées, etc)



# **Annexe 11 - Questionnaire (Analyse des besoins)**

## **Agropolis International**

Projet de réaménagement de la Bibliothèque Ouverte Montpellier - Languedoc-Roussillon (BOMLR <http://bomlr/home.html>) : un accès unique aux ressources documentaires des établissements et organismes de recherche de la région

### **Vos besoins en recherche d'information au niveau régional :**

#### **Questionnaire**

1. Nom, Prénom (facultatifs) :
  
2. Nom de l'établissement/organisme :
  
3. Statut (enseignant-chercheur, étudiant, doctorant, personnel administratif, professionnel de l'information) :
  
4. Quelles ressources / quels services d'information-documentation sollicitez-vous pour vos recherches ? Précisez le nom des ressources utilisées (ex. : Web of Science, catalogue de la bibliothèque de votre établissement, archives ouvertes HAL, etc.)
  
5. Comment procédez-vous pour effectuer une recherche en ligne, dans une base de données d'articles par exemple ?
  - par recherche simple, puis affinage des résultats
  - par recherche avancée, avec notamment l'utilisation des opérateurs booléens ET-OU-SAUF
  - par navigation (accès par filtres thématiques par exemple)
  
6. Pour l'affinage des réponses, quels sont les types de filtres qui vous semblent les plus pertinents lorsque vous effectuez une recherche ?
  - Filtres par année
  - " par «Peer review status<sup>31</sup>»

---

<sup>31</sup> Se reporter au glossaire p.96

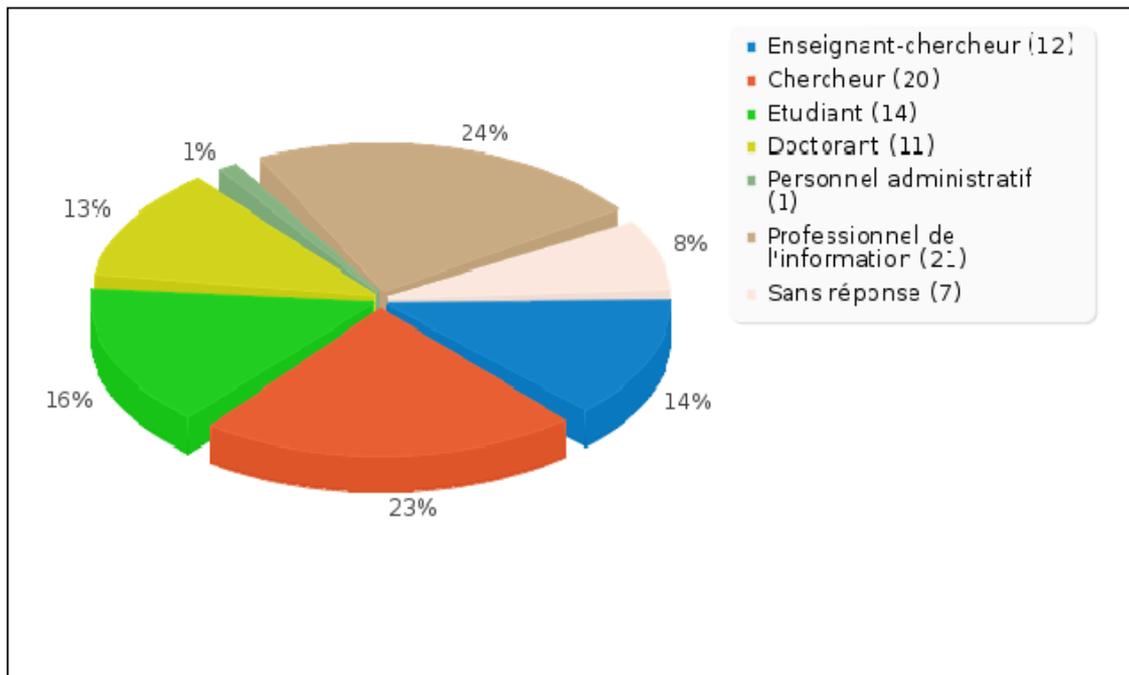
- " par type de ressources (documents full-text ou non, ouvrages, articles, documents cartographiques, etc.)
- " par localisation (nom de l'établissement/organisation détenteur de la ressource)

7. Avez-vous l'habitude d'utiliser des outils de veille (alertes, flux RSS) ?
8. La dimension participative sur les portails documentaires/bases de données en ligne vous semble-t-elle intéressante ? (partage de signets sur les réseaux sociaux, ajouts de commentaires, d'avis, etc.)
9. Connaissez-vous la BO-MLR ? L'utilisez-vous ? (Si non, pourquoi. Si oui, comment l'outil pourrait-il être amélioré ?)

## Annexe 12 - Résultats du questionnaire

On note que les résultats suivants ont été générés par le logiciel d'enquête Lime Survey qui a permis la mise en ligne du questionnaire.

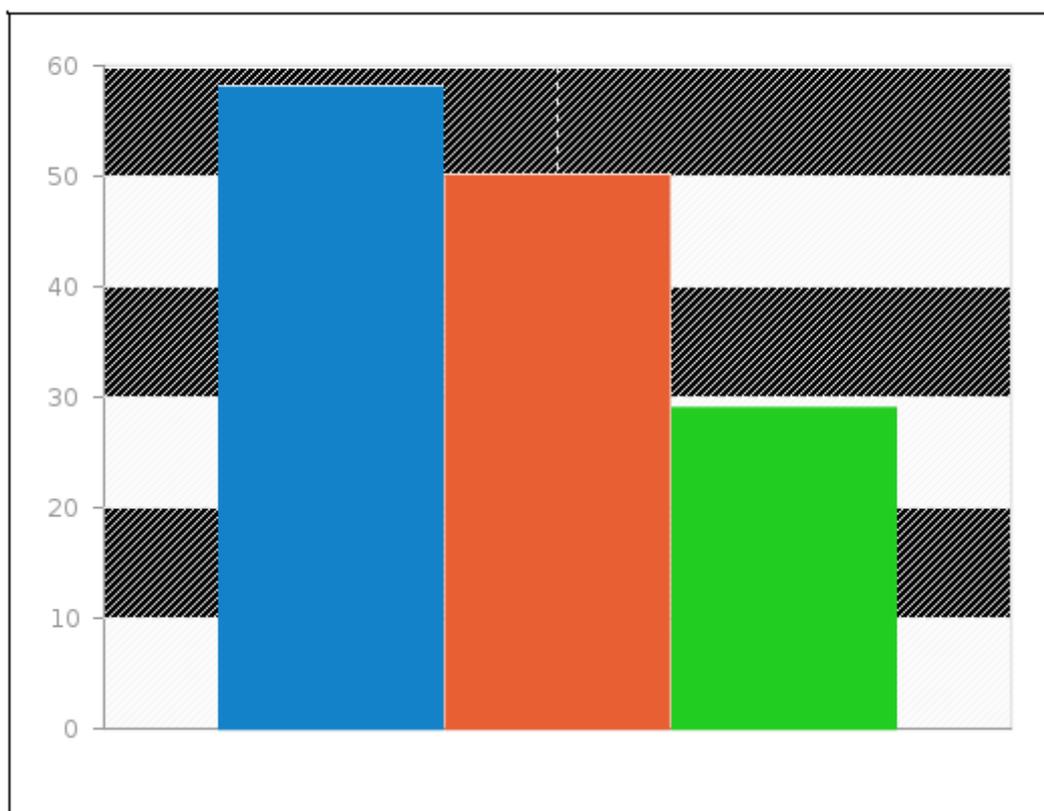
Résultats pour la question 3 (statut) :



Résultats pour la question 5 (mode de recherche utilisé) :

Réponse	Décompte	Pourcentage
<b>Recherche simple</b> (puis affinage des résultats)	58	67.44 %
<b>Recherche avancée</b> (avec utilisation des opérateurs booléens)	50	58.14 %
<b>Recherche par navigation</b> (accès par filtres thématiques par exemple)	29	33.72 %

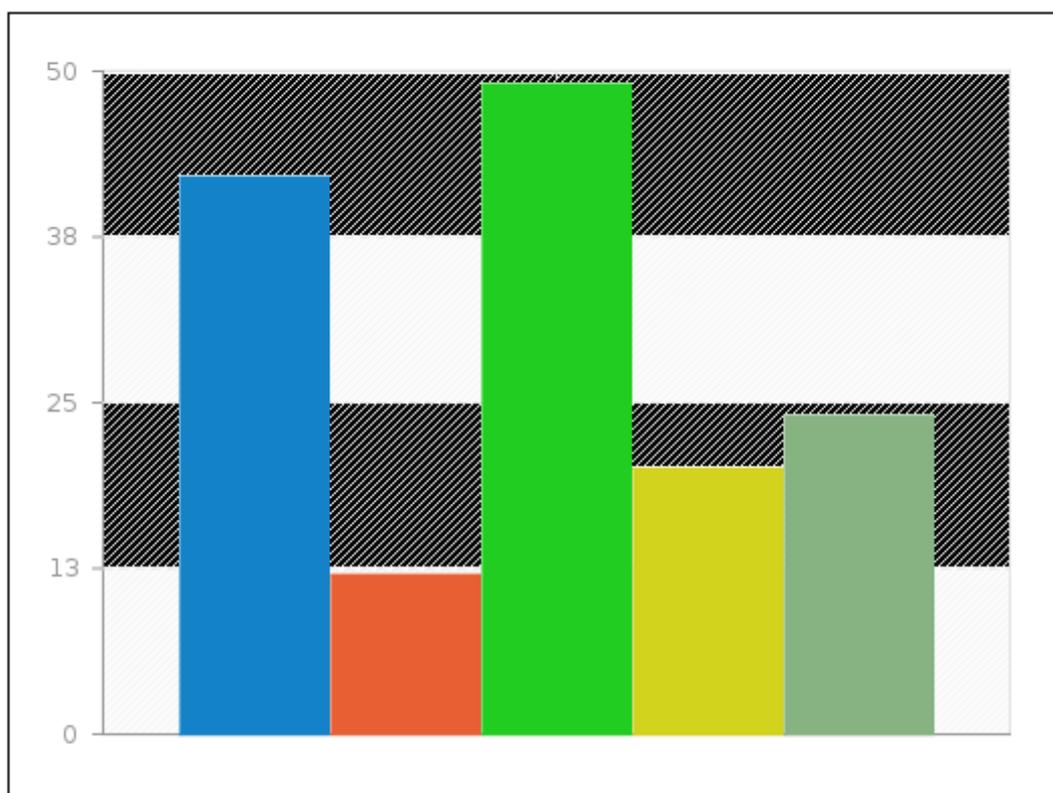
Expression graphique des résultats pour la question 5 :



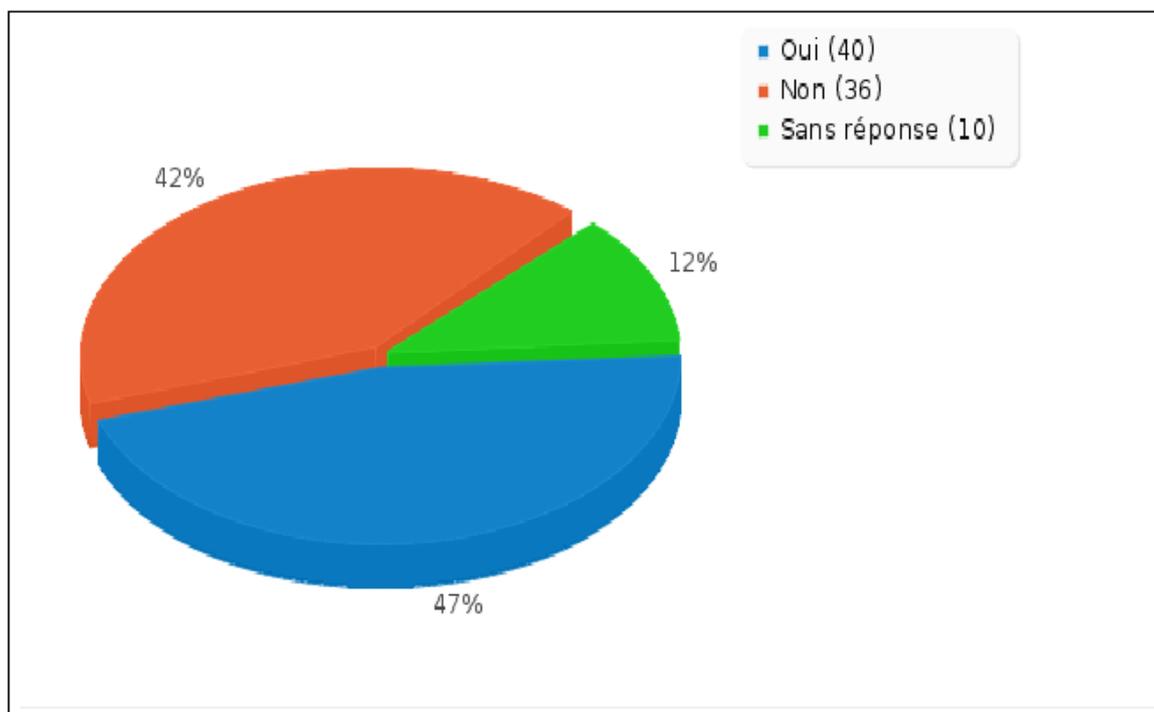
Résultats pour la question 6 (types de filtres utilisés pour l'affinage des réponses) :

Réponse	Décompte	Pourcentage
Filtre <b>par année</b>	42	48.84 %
Filtre <b>par <i>Peer review status</i></b> (évaluation par les pairs)	12	13.95 %
Filtre <b>par type de ressources</b> (documents <i>full-text</i> ou non, ouvrages, articles, etc.)	49	56.98 %
Filtre <b>par localisation</b> (nom de l'établissement détenteur de la ressource)	20	23.26 %
<b>Autre</b>	24	27.91 %

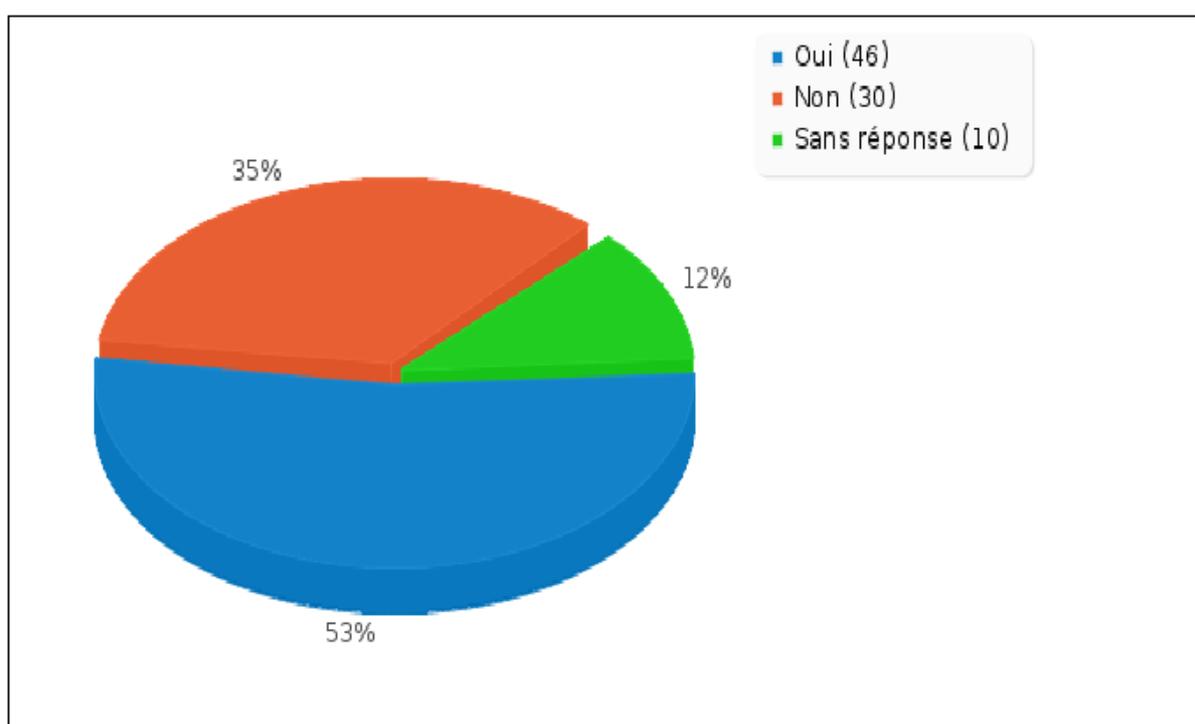
Expression graphique des résultats pour la question 6 :



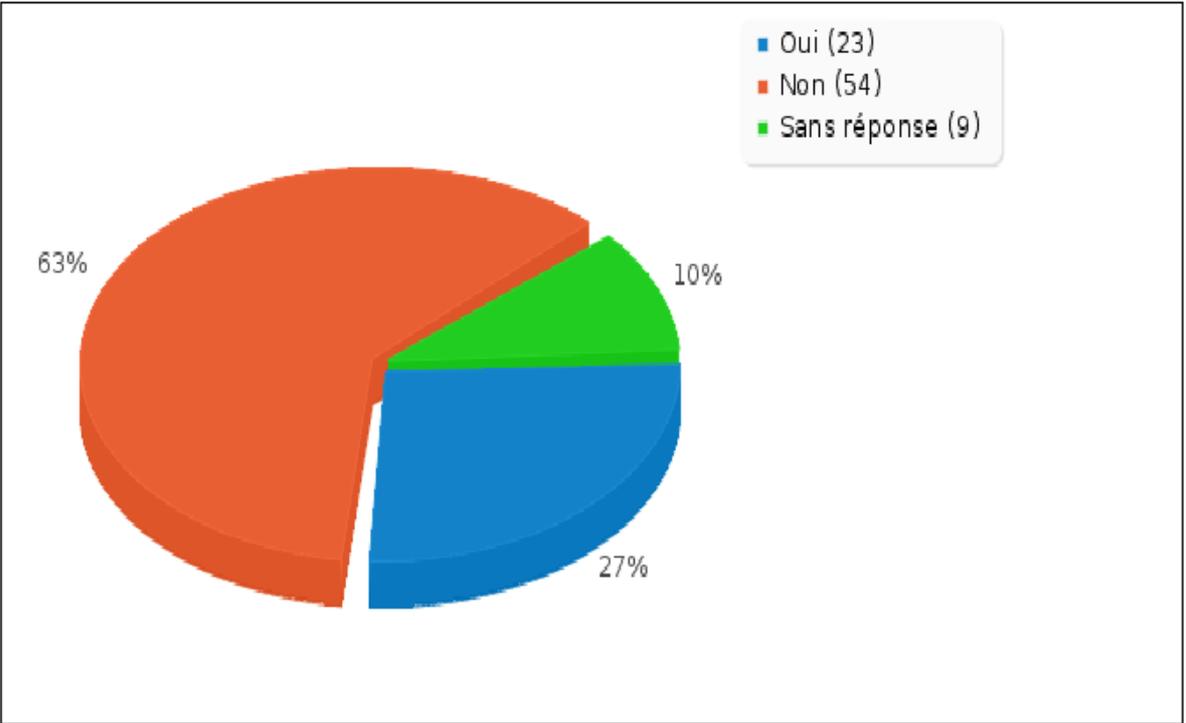
Résultats pour la question 7 (utilisation des outils de veille) :



Résultats pour la question 8 (la dimension participative sur les portails documentaires vous semble-t-elle intéressante ?)



Résultats de la question 9 (connaissez-vous la BOMLR ?) :



## **Annexe 13 – Mini Cahier des charges**



### **Projet de réaménagement du portail de la Bibliothèque Ouvverte Montpellier Languedoc-Roussillon (BO-MLR)**

**Migration vers un nouvel outil de recherche :  
caractéristiques et fonctionnalités attendues**

**avril - mai 2012**

# **Plan**

## **1. DESCRIPTION DES RESSOURCES À INTERFACER**

- 1.1 Types de ressources**
- 1.2 Volumétrie**
- 1.3 Contexte informatique**

## **2. AUTRES RESSOURCES INTERROGEABLES**

- 2.1 Types de ressources**
- 2.2 Volumétrie**

## **3. SYNTHÈSE DES MODES D'INTERROGATION**

- 3.1 Interrogation via des connecteurs synchrones**
- 3.2 Interrogation asynchrone via le protocole OAI-PMH**

## **4. UTILISATEURS**

- 4.1 Typologie des publics cibles**
- 4.2 Nombre potentiel d'utilisateurs simultanés**

## **5. FONCTIONNALITÉS SOUHAITÉES**

- 5.1 Fonctionnalités de recherche**
- 5.2 Fonctionnalités d'affichage des résultats**

## **6. FONCTIONS D'ADMINISTRATION**

## **7. HÉBERGEMENT**

La Commission d'Information Scientifique et Technique d'Agropolis a décidé de mettre en place un nouveau méta-moteur pour l'interface de sa Bibliothèque Ouverte Montpellier – Languedoc-Roussillon (BOMLR) selon les besoins des différents établissements et organismes de recherche impliqués dans le projet.

Il s'agit ici de présenter d'une part les ressources à interfacier sur le portail (aspects documentaires et techniques), d'autre part les fonctionnalités de recherche et d'affichage des résultats souhaitées. Enfin, nous mettrons également en évidence les critères liés aux fonctions d'administration et à l'hébergement de la solution.

## **1. Description des ressources à interfacier en priorité (phase 1)**

### 1.1 Types de ressources

La BOMLR doit pouvoir interfacier les bases de données de plusieurs établissements d'enseignement supérieur et d'organismes de recherche présents en Languedoc-Roussillon.

Les bases de données de ces établissements proposent différents types de ressources qu'il faudra intégrer sur le portail de la BOMLR :

- Catalogues d'ouvrages, de travaux des étudiants et des doctorants (mémoires, thèses) + publications des chercheurs
- Catalogues de revues papier

### 1.2 Volumétrie

<b>Nom de l'établissement</b>	<b>Nombre de références</b>
AgroParisTech Montpellier	20000
Bioversity International	13200
BIU Montpellier	907970
BU Perpignan	173000
CEFE-CNRS	30000
CIHEAM-IAM.M	27000
Cirad	240000

Nom de l'établissement	Nombre de références
Ecole des Mines d'ALès	22000
INRA	125000 pour ProdInra, tous instituts confondus = niveau national. Pour le reste des références au niveau régional, cf. SupAgro.
IRD	131590
IRSTEA	7629 (inclut la base de publications)
Montpellier SupAgro/INRA	81147
Observatoire Océanologique de Banyuls	-

On a donc ici un total d'environ **1 786 536 références**.

### 1.3 Contexte informatique

Nom de l'établissement	Logiciel (SGBD, SIGB, etc)	Protocole de communication
AgroParisTech	PMB	Webservice OAI-PMH (depuis mai 2012)
Bioversity International	CDS Isis → migration vers Zotero d'ici juillet 2012 + MySQL	Connexion directe au serveur de base de données MySQL (port 3306)
BIU Montpellier	Aleph + Flora (GED)	Z3950
BU Perpignan	Absys	Z3950
CEFE-CNRS	Reference Manager (pas de base de publications en ligne)	Connexion directe au serveur de base de données MySQL (après export vers MySQL)
CIHEAM-IAM.M	PMB	Webservice
Cirad	Loris	Z3950  OAI-PMH courant 2013 (projet d'un nouveau système d'information documentaire qui permettra d'inclure ce protocole)
Ecole des Mines d'ALès	Alexandrie-Kentika	Webservice OAI-PMH (portail commun des ressources des 6 écoles)

Nom de l'établissement	Logiciel (SGBD, SIGB, etc)	Protocole de communication
INRA	PMB (fusion avec le catalogue de SupAgro)	Webservice
INRA	ProdInra	OAI-PMH
IRD	Cadic-Intégrale	OAI-PMH Z3950 à venir
IRSTEA	Cadic-Intégrale	OAI-PMH
Montpellier SupAgro	PMB	Webservice (avec un connecteur SOAP)
Observatoire Océanologique de Banyuls	Biblio-Tech + Symphony (via Paris IV Bibliothèque Universitaire Pierre et Marie Curie)	Z 3950

## 2. Autres ressources interrogeables (phase 2)

### 2.1 Types de ressources

- Catalogues de revues (ou d'articles) en ligne : AtoZ, CAIRN
- Autres archives ouvertes pour le dépôt d'articles/ productions scientifiques - en plus de la base ProdInra - : HAL (domaine correspondant à chaque organisme : ex. Hal-Cirad, Hal-Inserm, etc.), RePEc (archives ouvertes en sciences économiques).
- Réservoirs de thèses en format numérique : TEL, Pastel, thèses.fr

### 2.2 Volumétrie

Ressource	Nombre de références
AtoZ	-
CAIRN	25000 références (revues, ouvrages collectifs, magazines, encyclopédies de poche)
Hal-Cirad	1710 documents avec texte intégral

Hal-Inserm	5545 documents avec texte intégral
Hal-AgroParisTech	1267 documents avec texte intégral
RePEc	Plus d'un million de références
TEL	27766 documents avec texte intégral
Pastel (thèses en ligne de ParisTech)	3560 documents avec texte intégral
Thèses.fr (projet de l'ABES)	8000 thèses soutenues depuis 2006, dont 6000 en texte intégral

On a donc ici un total d'environ **1 075 977 références**.

Le protocole de communication utilisé ici est l'OAI-PMH.

### 3. Synthèse des modes d'interrogation

L'outil devra donc inclure deux systèmes : la recherche fédérée et la recherche intégrée.

#### *Recherche fédérée*

##### 3.1 Interrogation via des connecteurs synchrones

- Z3950 (5 bases)
- Webservices (6 bases)
- Connexion directe au serveur de base de données MySQL (exemples de Bioversity International et du CEFÉ)

#### *Recherche intégrée*

##### 3.2 Interrogation asynchrone via le protocole OAI-PMH

Ce mode d'interrogation peut concerner certaines bases de données d'organismes de recherche (ressources de l'IRD ou de l'IRSTEA par exemple) mais il est surtout utilisé pour les autres ressources interrogeables énumérées précédemment, qui incluent principalement les archives ouvertes ou autres réservoirs de documents accessibles en *full-text*.

En plus des connecteurs, le méta-moteur devra ainsi pouvoir moissonner directement les métadonnées des notices bibliographiques et/ou le texte intégral des documents issus des différentes bases. Il créera ainsi un index qui pourra être interrogé par l'utilisateur.

#### Rappel :

- Nombre approximatif de bases de priorité P1 concernées par ce protocole : 4 bases de données, soit 504 210 références.
- Nombre approximatif de références pour les ressources de priorité P2 concernées par ce protocole : 1 075 977 références.

## **4. Utilisateurs**

### 4.1 Typologie des publics

La BOMLR cible plus particulièrement quatre grands groupes d'utilisateurs :

- Les professionnels de l'information, les documentalistes
- Les chercheurs, les enseignants-chercheurs
- Les étudiants et les doctorants

### 4.2 Nombre potentiel d'utilisateurs simultanés

De 2000 à 2500 visiteurs recensés pour les mois où la consultation de la BOMLR a été la plus importante (statistiques 2004/2005), soit environ 80 visiteurs par jour. Il est ainsi difficile d'évaluer le nombre potentiel d'utilisateurs simultanés, d'autant plus que ces taux de fréquentation datent déjà de plusieurs années, mais au vu de ces résultats, on peut penser que le nombre d'utilisateurs simultanés sera relativement faible. Proposer quelques accès seulement devrait être suffisant. Au cas où le nombre de visiteurs dépasse exceptionnellement le quota prévu, un message devra pouvoir s'afficher sur

l'interface pour informer son utilisateur que l'accès est momentanément indisponible et qu'il faut relancer la recherche ultérieurement.

## 5. Fonctionnalités souhaitées

### 5.1 Fonctionnalités de recherche

Fonctionnalités prioritaires :

- Interrogation **unique** sur les catalogues d'ouvrages/de périodiques et les publications scientifiques (connecteurs et moissonneur OAI)
- Modes de recherche : recherche simple et recherche multi-champs
- Historique des recherches (avec la possibilité de combiner des anciennes recherches)

Fonctionnalités secondaires, qui seraient un plus :

- Possibilité d'élargir la recherche vers d'autres sites/moteurs tels que Google Scholar, Web of Science, Science Direct, etc.
- Aide à la saisie (lemmatisation, auto-complétion)

### 5.2 Fonctionnalités d'affichage des résultats

À noter qu'on attend un délai de restitution des résultats assez rapide (deux à quatre secondes maximum), ou tout au moins, un affichage des premiers résultats dans le respect de ce délai, en attendant l'affichage des autres réponses, qui doit également se faire dans un laps de temps raisonnable.

Fonctionnalités prioritaires :

- Lien vers le texte intégral et/ou proposition d'un filtre « full-text » à l'affichage des résultats (→ résolveur de liens)

- Localisation des ressources de la région (filtre par centre de documentation / organisme par exemple)
- Distinction entre ressources en ligne et ressources à disposition dans tel ou tel centre documentaire
- De manière générale, un affinage développé, comportant un grand nombre de filtres (par type de ressources, auteur, date, thématique, etc.)
- Exportation de notices (formats compatibles avec EndNote et Zotero)

Fonctionnalités secondaires, qui seraient un plus :

- Génération d'un nuage de mots-clés
- Système de création d'alertes/flux RSS liés à la requête effectuée
- Partage de notices sur les réseaux sociaux (Twitter, Facebook, etc.)

## 6. Fonctions d'administration

En lien avec le caractère évolutif de l'outil – qui reste un critère essentiel - , la solution doit permettre :

- La gestion des connecteurs
- L'indexation des sources externes
- La paramétrabilité de l'interface utilisateur

Il est aussi primordial que la solution propose un **module de statistiques**, qui permette d'évaluer le nombre de documents (en texte intégral ou non) atteints, le nombre de visiteurs et de recherches effectuées, les sources interrogées (quels établissements sont les plus sollicités ? S'agit-il d'organismes présents en région ? Etc.)

## 7. Hébergement

*A priori*, le système de recherche serait installé sur le site d'Agropolis International.

Pour inclure le moissonnage OAI, un serveur avec une large capacité de stockage, et une mémoire vive importante est envisagé (préconisations : 146 GB – 12 Go).

## Annexe 15 - Captures d'écran de Vufind installé sur le serveur d'Agropolis (phase de tests)

**AGROPOLIS INTERNATIONALE**

Connexion

Lingue: Français

English  
Deutsch  
Español  
Français  
Italiano  
Nederlands  
Português  
Türkçe  
עברית  
Gaeilge  
Cymraeg  
Ελληνικά

agro-alimentaire Tous les champs Rechercher Recherche avancée

Page d'accueil Rechercher: Agro-Alimentaire

Résultat(s) 1 - 20 de 88 pour la requête: 'agro-alimentaire', Temps de recherche: 0.54s Tri Relevance

Recherches alternatives:  
agro-alimentaire » agro-alimentation

1 [Les systèmes agro-alimentaires méditerranéens](#) [Ajouter aux favoris](#)  
par Pérez R.  
[Accéder au texte intégral](#)  
Online

2 [Vision industrielle dans les industries agro-alimentaires](#) [Ajouter aux favoris](#)  
par Guizard, C. Publié 1994  
"... nouveau paysage de la consommation impose aux industriels de l'agro-alimentaire une évolution dans leurs..."  
[Accéder au texte intégral](#)  
application/pdf Online

3 [Vision artificielle dans les industries agro-alimentaires](#) [Ajouter aux favoris](#)  
par Guizard, C. Publié 1992  
[Accéder au texte intégral](#)  
application/pdf Online

Accès au texte intégral

Filters for refining search results: by center of documentation, type of resources, author, thematic, etc

Lien citation

 [Citer ceci](#)  [Envoyer par SMS](#)  [Envoyer ceci](#)  [Ajouter aux favoris](#)

## Atomisation of agricultural sprays: influence of some liquid properties

Auteur principal: [Ben, L.](#)

Autres auteurs: [Vallet, A.](#), [Bonicelli, B.](#)

Format: application/pdf Online

Langue: English

Publié: 2004

Sujets: [PULVERISATEUR](#)  
[PESTICIDE](#)  
[EQUIPEMENT](#)  
[BUSE A FENTE](#)  
[BUSE A TURBULENCE](#)  
[PROTECTION DES CULTURES](#)  
[PROTECTION DE L'ENVIRONNEMENT](#)  
[PULVERISATION](#)  
[OMBROSCOPIE](#)  
[SPRAY AGRICOLE](#)  
[AGENT TENSIO-ACTIF](#)

Accès en ligne: <http://cemadoc.cemaqref.fr/cemoa/PUB00014287>

Tags: Pas de tags, Soyez le premier à ajouter un tag!  [Ajouter un tag](#)



### Exemplaires similaires

[Test method for boom suspension on spray distribution. Part I : Experimental study of pesticide application under a moving boom](#)  
par: Lardoux, Y. Publié: (2007)

[Spray-drift from an axial fan sprayer Full-scale measurement in a wind-tunnel: low-drift v/s standard nozzle performance](#)  
par: Richardson, G.M. Publié: (2005)

[Pesticide atomization modelling: application to a hollow cone nozzle](#)  
par: De Luca, M. Publié: (2007)

[Surfactant influence on droplet size and velocity spectra from hollow cone and air induced sprays](#)  
par: Vallet, A. Publié: (2011)

[Contribution à la modélisation de la pulvérisation agricole : atomisation et transport](#)  
par: Belhadef, A. Publié: (2009)

Suggestion de documents

Indexation par l'utilisateur

[Exemplaires](#) [Description](#) [Commentaires](#) [Affichage Marc](#)

### Internet

<http://cemadoc.cemaqref.fr/cemoa/PUB00014287>

Lien vers le texte intégral

## Gestion du panier

Tous les champs ▼ Rechercher [Recherche avancée](#)

[Page d'accueil](#) > [Votre compte](#) > Liste

### Mes favoris

 [Modifier la liste](#)

 [Détruire la liste](#)

 [Votre compte](#)

Résultat(s) 1 - 2 de 2

Tri Relevance ▼

Choisir une page | avec la sélection: [Courriel](#)  [Supprimer](#)  [Exporter](#)  [Imprimer](#)

1  [Atomisation of agricultural sprays: influence of some liquid](#)  
par: [Ben L.](#)  
application/pdf Online

NO IMAGE AVAILABLE

Export possible vers RefWorks et Endnote

 [Éditer](#)  [Supprimer](#)

2  [Suivi et caractérisation physique de l'aquifère du périmètre irrigué du secteur Ngallenka Amont \(région de Podor\) : mémoire de stage](#)  
par: [Faignant, Maider](#)  
text/pdf Online

NO IMAGE AVAILABLE

 [Éditer](#)  [Supprimer](#)

Choisir une page | avec la sélection: [Courriel](#)  [Supprimer](#)  [Exporter](#)  [Imprimer](#)

- [Favoris](#)
- [Documents empruntés](#)
- [Emprunts et rappels](#)
- [Amendes](#)
- [Profil](#)
- [Recherches enregistrées](#)

### Vos listes

- Mes favoris (2)

NO IMAGE  
AVAILABLE

par [Schlumberger, O.](#) Publié 2003

[Accéder au texte intégral](#)

application/pdf Online

8

NO IMAGE  
AVAILABLE

[Atomisation of agricultural sprays: influence of some liquid properties](#)

♥ [Ajouter aux favoris](#)

par [Ben, L.](#) Publié 2004

[Accéder au texte intégral](#)

application/pdf Online

9

NO IMAGE  
AVAILABLE

[Contribution à la modélisation de l'atomisation d'un jet liquide](#)

♥ [Ajouter aux favoris](#)

par [Vallet, A.](#) Publié 2004

[Accéder au texte intégral](#)

application/pdf Online

10

NO IMAGE  
AVAILABLE

[Evaluation des procédures de gestion volumétrique en Charente. Typologie des exploitations irrigantes et impact économique du mode de gestion de l'eau d'irrigation sur ces...](#)

♥ [Ajouter aux favoris](#)

par [Rousseau, C.](#) Publié 2003

[Accéder au texte intégral](#)

application/pdf Online

11

NO IMAGE  
AVAILABLE

[Etude de la dynamique du peuplement piscicole de la retenue hydroélectrique d'EGUZON \(36\). Rapport d'étape](#)

♥ [Ajouter aux favoris](#)

par [Iz, P.](#) Publié 2002

[Accéder au texte intégral](#)

application/pdf Online

Autres exemples de  
filtres: par genre, période  
/ date, par localisation

#### Genre

[Bibliography](#) (30)

[Congresses](#) (19)

[Biography](#) (17)

[Juvenile literature](#) (16)

[Fiction](#) (15)

[plus ...](#)

#### Période

[20th century](#) (44)

[1945-](#) (16)

[19th century](#) (13)

[1564-1616](#) (8)

[1945-1989](#) (8)

[plus ...](#)

#### Région

[United States](#) (260)

[Great Britain](#) (31)

[Soviet Union](#) (23)

[Europe](#) (16)

[New York \(State\)](#) (15)

[plus ...](#)

#### Année de publication

De:  A:



## Glossaire

**Archive ouverte** : «désigne un réservoir où sont déposées des données issues de la recherche scientifique et de l'enseignement et dont l'accès se veut ouvert c'est-à-dire sans barrière. Cette ouverture est rendue possible par l'utilisation de protocoles communs qui facilitent l'accessibilité de contenus provenant de plusieurs entrepôts maintenus par différents fournisseurs de données.» (Définition *INIST*)

**Connecteur** : le terme, issu de la recherche fédérée, désigne un fichier chargé de traduire l'information du moteur de recherche fédérée vers la base de donnée distante que l'on souhaite interfacer. « Le connecteur récupère la requête selon la syntaxe du moteur de recherche, la traduit dans la syntaxe attendue par le moteur distant, récupère la liste des résultats, et l'affiche selon les règles d'affichage du moteur de recherche fédérée. » (Site Bibliopédia, « Recherche fédérée » - consultation juin 2012)

**Dublin Core** : format descriptif d'une ressource constituant un jeu de métadonnées qui a vocation à permettre l'interopérabilité entre plusieurs systèmes d'information.

**Métadonnées** : «ensemble de données structurées décrivant des ressources physiques ou numériques. Elles sont un maillon essentiel pour le partage de l'information et l'interopérabilité des ressources électroniques. » (définition *INIST*)

**OAI-PMH (protocole)** : «le protocole OAI-PMH est un protocole qui définit les conditions du transfert de métadonnées d'une archive ouverte, produite par un fournisseur de données, vers le serveur d'un fournisseur de services. » (définition *INIST*)

**Open-Access (ou libre accès)** : mise à disposition gratuite d'une ressource sur Internet, permettant aux utilisateurs d'en lire, télécharger, copier, transmettre, imprimer le texte intégral sans autre contrainte que la garantie de son intégrité et de sa paternité (reconnaissance de l'auteur).

**Peer-review (ou évaluation par les pairs)** : désigne, dans les disciplines scientifiques, l'activité collective des chercheurs qui jaugent de façon critique les travaux d'autres chercheurs (d'après Wikipédia).

**PHP** : langage libre de scripts qui permet la production de pages web dynamiques (pages générées à la demande, qui varient en fonction d'information).

**Protocole de communication** : en informatique, un protocole de communication désigne les règles et les conditions spécifiques permettant l'échange de données entre deux machines, ou deux systèmes d'information distincts. L'OAI-PMH est un exemple de protocole de communication.

**Recherche fédérée** : permet en une seule requête l'interrogation à distance et en temps réel de plusieurs sources de données hétérogènes.

**Recherche intégrée** : permet en une seule requête l'interrogation de plusieurs sources de données hétérogènes préalablement indexées. Il s'agit d'un système de recherche asynchrone, rendu possible par l'interrogation d'un index centralisant toutes les métadonnées des différentes ressources, et non par l'interrogation directe des bases de données distantes.

**SaaS** : utilisation d'un logiciel à distance, qui reste hébergé par l'éditeur.

**WebService/SRW** : le webservice ou SRW (Search and Retrieve Webservice) définit un protocole de communication qui permet l'échange de requêtes et de données entre bases distantes via un protocole HTTP.

**Z3950 (protocole)** : « Le protocole Z39.50 spécifie les procédures et les formats pour permettre à un client de requêter une base de données proposée par un serveur, d'identifier les informations correspondant aux critères de la recherche, et de récupérer les informations identifiées. Z39.50 fonctionne en mode synchrone, à la différence d'un moissonneur. Les requêtes émanant du client sont exécutées en direct sur le serveur distant et les résultats sont rapatriés instantanément en retour. » (définition *BNF*)

# Table des matières

SIGLES ET ABRÉVIATIONS.....	9
INTRODUCTION.....	11
<b>PARTIE 1 – PRÉSENTATION DE LA BIBLIOTHÈQUE OUVERTE MONTPELLIER - LANGUEDOC-ROUSSILLON .....</b>	<b>13</b>
<b>1.1 Un projet collaboratif .....</b>	<b>13</b>
1.1.1 <i>Contexte et objectifs.....</i>	13
1.1.2 <i>Historique.....</i>	13
<b>1.2 La BOMLR aujourd'hui.....</b>	<b>14</b>
1.2.1 <i>Les ressources interfacées.....</i>	14
1.2.2 <i>Le métamoteur AskOnce.....</i>	14
1.2.3 <i>Les inconvénients.....</i>	15
<b>PARTIE 2 – RÉAMÉNAGER LA BOMLR : ANALYSES PRÉALABLES.....</b>	<b>17</b>
<b>2.1 État des lieux des bases de données à interfacier.....</b>	<b>17</b>
2.1.1 <i>Un tour d'horizon des différents établissements membres d'Agropolis.....</i>	17
2.1.2 <i>Des ressources hétérogènes.....</i>	18
<b>2.2 Étude des besoins.....</b>	<b>25</b>
2.2.1 <i>Les publics cibles.....</i>	25
2.2.2 <i>Enquêtes et entretiens.....</i>	25
2.2.2.1 <i>Entretiens avec les documentalistes.....</i>	25
2.2.2.2 <i>Résultats du questionnaire.....</i>	26
2.2.3 <i>De nouveaux objectifs pour la BOMLR.....</i>	27
<b>PARTIE 3 – RÉAMÉNAGER LA BOMLR : ÉTUDE ET PROPOSITION DE SOLUTIONS</b>	<b>29</b>
<b>3.1 Les principaux mécanismes de la recherche fédérée.....</b>	<b>29</b>
3.1.1 <i>La technologie des connecteurs.....</i>	29
3.1.2 <i>La technologie du « moissonnage ».....</i>	30
<b>3.2 Panorama des solutions de recherche.....</b>	<b>30</b>
3.2.1 <i>Les moteurs de recherche fédérée.....</i>	31
3.2.2 <i>Les moteurs de recherche intégrée.....</i>	31
3.2.3 <i>Les moteurs hybrides.....</i>	31
3.2.4 <i>Les outils de découverte.....</i>	32
<b>3.3 Solutions libres et propriétaires : quelles différences ?.....</b>	<b>32</b>
<b>3.4 Établissement des critères pour l'évaluation comparative des solutions.....</b>	<b>32</b>
3.4.1 <i>La nécessité d'un outil mixte : un premier critère de sélection .....</i>	32
3.4.2 <i>Les fonctionnalités de recherche et d'affichage des résultats attendues.....</i>	33
<b>3.5 Tableau comparatif des solutions.....</b>	<b>34</b>
<b>3.6 Exploration des solutions .....</b>	<b>39</b>
3.6.1 <i>Les solutions retenues.....</i>	39
3.6.2 <i>Exploration de la sélection : premiers résultats.....</i>	39
3.6.3 <i>Exploration de la sélection : perspectives.....</i>	40
CONCLUSION.....	43
BIBLIOGRAPHIE ET WEBOGRAPHIE.....	45
TABLE DES ANNEXES.....	47
GLOSSAIRE.....	95
TABLE DES MATIÈRES.....	97