

Résumés – descripteurs

Réorganisation de l'annuaire et du site web des « bases de données gratuites sur Internet ».

Résumé :

Ce rapport présente les étapes de la réorganisation d'un système documentaire offrant la possibilité de rechercher puis d'accéder en ligne à plus d'un millier de bases de données disponibles gratuitement sur Internet. Une étude de l'existant a permis d'établir un bilan général, évaluant le fonds documentaire et le traitement à effectuer. Des améliorations ont pu être alors proposées, du choix d'une classification à la restructuration de la base de données et des interfaces de recherche et de saisie, totalement refondues.

Descripteurs :

système d'information, base de données en ligne, classification, Internet.

Reorganization of the directory and of the web site of the « bases de données gratuites sur Internet ».

Summary :

This report presents the stages of the reorganization of a documentary system offering the possibility to search and reach online over a thousand databases, free of charge and available from the Internet. Initially, a study of the system allowed to draw up general estimate of the references and data processing. Improvements were then suggested from the choice of a classification to the reorganisation of the database and of the research and input interfaces, totally revised.

Keywords :

information system, online database, indexing, Internet.

Remerciements

Je tiens à remercier

Jean-Pierre Lardy pour la confiance et l'autonomie qu'il m'a
immédiatement accordées ;

Hugo Etiévant dont l'expérience de la conception de site web a été
riche d'enseignements et dont les talents d'infographiste ont mis en
valeur mon travail ;

Adel Larouci et Marc Bertin, stagiaires à l'URFIST, pour leurs conseils
judicieux ;

Et toutes les personnes qui ont eu la gentillesse de tester DADI aux
différentes phases de sa conception.

Sommaire

RÉSUMÉS – DESCRIPTEURS	2
REMERCIEMENTS.....	3
INTRODUCTION	6
PRÉSENTATION DE LA STRUCTURE D'ACCUEIL	7
1. Trois axes de travail	7
1.1. La formation de formateurs	7
1.2. La production de supports de cours	8
1.3. La veille informationnelle	8
2. L'URFIST de Lyon	8
2.1. L'équipe.....	8
2.2. Son site web.....	9
PARTIE 1 : « LES BASES DE DONNEES GRATUITES SUR INTERNET » :	
ANALYSE DE L'EXISTANT	10
1. Les bases de données gratuites disponibles sur Internet.	10
1.1. De l'idée à la réalisation.....	10
1.2. Analyse de l'existant	11
1.2.1. L'interface visiteur et les options de recherche	11
1.2.1.1. La page d'accueil	11
1.2.1.2. Les options de recherche.....	13
1.3. Structure de la base de données et interface administrateur	15
1.3.1. Structure de la base de données	15
1.3.2. L'interface administrateur	17
1.4. Bilan du système documentaire	17
PARTIE 2 : MISE À JOUR DU RÉPERTOIRE ET RÉORGANISATION DU	
SITE : DES BESOINS ET DES SOLUTIONS.....	20
1. Analyse des besoins	20
1.1. Changement de la structure de la base de données	20
1.2. Vérification des liens	20
1.3. Recherche de nouvelles bases de données.....	21
1.4. Veille informationnelle	21
1.5. Indexation intellectuelle	21
1.6. Présentation des notices	21
1.7. Recherche en ligne.....	22
1.8. Nouvelles interfaces web.....	22
1.9. Un site bilingue : français-anglais	22
1.10. Sigle et charte graphique	22
1.11. Promotion	23
2. Mise à jour de l'annuaire	23
2.1. Vérification des liens	23
2.2. Recherche de nouvelles bases de données.....	23

2.3. Mise en place d'une veille avec Copernic	24
2.4. Nouvelle indexation.....	24
3. La nouvelle administration de l'annuaire	26
3.1. Création d'une nouvelle notice	26
3.2. Une gestion simplifiée	30
4. Le nouveau site	33
4.1. Un nouveau nom	33
4.2. Schéma des liens.....	34
4.3. Le plan du site	35
4.3.1. Le plan du site.....	35
4.4. Trois options de recherche.....	38
4.4.1. La recherche avancée	39
4.4.2. La navigation dans les catégories de la classification Dewey .	41
4.4.3. Une notice interactive	43
PARTIE 3 : NOUVELLE ARCHITECTURE INFORMATIQUE ET PÉRENNITÉ DU PROJET.....	44
1. Architecture informatique	44
1.1. Le site visiteur	44
1.1.1. L'environnement technique	44
1.1.2. Charte graphique et maquette	45
1.1.2.1. <i>Charte graphique</i>	45
1.1.2.2. <i>Maquette</i>	47
1.2. Structure du site	47
1.2.1. Partie visiteur.....	47
1.2.2. Partie administration	48
1.2.2.1. <i>Description de la base de données</i>	48
1.2.2.2. <i>Scripts généraux</i>	49
1.2.2.3. <i>Scripts des pages</i>	50
1.2.2.4. <i>Scripts des objets</i>	54
2. Difficultés rencontrées	56
2.1. La question légale.....	56
2.2. Les contraintes informatiques	56
3. Promotion et avenir	57
3.1. Promotion	57
3.2. Avenir	58
CONCLUSION	59
BIBLIOGRAPHIE.....	60
TABLE DES ANNEXES.....	61

Introduction

Ce stage a été réalisé à l'Université Claude Bernard de Lyon, à l'URFIST (Unité Régionale de Formation et de Promotion pour l'Information Scientifique et Technique), durant une période de 4 mois, du 3 juin au 27 septembre 2002.

Le stage a eu pour objet la réorganisation de l'annuaire recensant et gérant les bases de données gratuites disponibles sur Internet dans le but de le rendre plus performant, c'est-à-dire de basculer une base de données gérée sous Access vers MySQL et d'y apporter des améliorations. Il s'est agi également de mettre à jour un fonds documentaire constitué de bases de données et de l'organiser de façon à retrouver rapidement et à coup sûr l'information recherchée en mettant en place un nouveau système d'indexation et de recherche.

Une partie du stage a été consacrée à l'analyse. Il a fallu tout d'abord dresser le bilan du fonctionnement et de l'administration de la base de données afin de pouvoir proposer ensuite des améliorations. Ces dernières passaient par la création de nouvelles interfaces visiteur et administrateur aussi fonctionnelles qu'efficaces et donc par une collaboration étroite avec l'informaticien chargé du développement de la nouvelle application.

Il sera fait, dans une première partie, une analyse détaillée de l'ancien système. Une analyse des besoins sera menée dans une seconde partie ainsi que la description de la mise en œuvre du nouveau système. La troisième partie sera consacrée d'une part à la structure informatique de la nouvelle application et d'autre part aux difficultés rencontrées au cours de ces quatre mois de stage avant d'envisager une suite à donner au projet pour améliorer encore le système.

Présentation de la structure d'accueil

Les **URFIST** sont des Unités Régionales de Formation et de Promotion pour l'Information Scientifique et Technique. Elles sont une réponse apportée par le Ministère de l'Education Nationale, de la Recherche et de la Technologie à l'enjeu majeur qu'est devenue la gestion du flux de l'information spécialisée face à l'expansion considérable du marché de l'information et la multiplication de ses différents supports. La nécessité de former les futurs acteurs et utilisateurs de cette industrie de l'information s'imposait.

Dans cette optique, un réseau de sept URFIST a été installé dans les pôles universitaires stratégiques : Bordeaux, Lyon, Nice, Paris, Rennes, Strasbourg et Toulouse. Ces unités de formation sont en liaison avec l'ensemble des partenaires nationaux de la production et de la diffusion de l'information spécialisée et assurent ainsi la couverture de l'ensemble des régions.

1. Trois axes de travail

1.1. La formation de formateurs

Les URFIST ont pour principale mission la formation et la sensibilisation des personnels de bibliothèque, des enseignants-chercheurs des universités et des étudiants des 2^e et 3^e cycle, à l'usage de l'information électronique, quel qu'en soit le support ou bien le mode d'accès (CD-Rom, Internet ...), dans un grand nombre de disciplines.

Une partie de leurs actions s'adresse plus particulièrement aux bibliothécaires et enseignants susceptibles de former des étudiants à la maîtrise des technologies de l'information.

Les URFIST développent par ailleurs des actions de formation qui permettent de poser les conditions d'une réflexion autour des problématiques émergentes : Intelligence collective, distribuée, déploiement des dispositifs hypertextes et

hypermédias, représentation cartographique de l'information, questionnements et enjeux liés au développement des nouvelles Mémoires numériques.

1.2. La production de supports de cours

Ces unités mettent à la disposition des utilisateurs et des formateurs les documents pédagogiques et les publications qu'elles produisent.

1.3. La veille informationnelle

Elles contribuent enfin, par la mise en place de stages thématiques, à améliorer la connaissance et l'exploitation des gisements d'information spécialisée, ainsi que la pratique de la veille informationnelle.

2. L'URFIST de Lyon

2.1. L'équipe

L'équipe de l'URFIST de Lyon¹ est composée de trois personnes :

- **Jean-Pierre LARDY** : Maître de conférences à l'Université de Lyon 1 et co-responsable de l'URFIST de Lyon. Spécialiste de l'information en sciences exactes et bio-médicales.
- **Lise HERZHAFT** : Conservateur des bibliothèques et co-responsable de l'URFIST de Lyon. Spécialiste de l'information en sciences sociale, économique, humaine et juridique.
- **Denise ETIEVANT** : Secrétaire chargée de l'inscription et de l'accueil des stagiaires.

L'Unité de formation met à la disposition de ses visiteurs et stagiaires une salle de formation équipée en matériels et logiciels.

¹Université Claude Bernard - LYON I
Bibliothèque Universitaire de la DOUA, 2ème étage
 43 Bd du 11 Novembre 1918
 69622 VILLEURBANNE

2.2. Son site web

L'URFIST de Lyon possède un site web² qui rend compte de son actualité, propose le programme de ses formations et offre la possibilité de s'y inscrire en ligne. Le site comporte également des liens vers un certain nombre de ressources : cours virtuels et outils de recherche permettant à chacun de trouver l'information adéquate sur Internet voire, et c'est l'un des buts, de s'auto-former à la recherche pertinente d'informations en ligne. Enfin, le site donne accès gratuitement à un **annuaire de bases de données** que l'on peut trouver et exploiter de façon gratuite sur Internet.

The screenshot displays a grid of resource cards on a yellow background. Each card contains text, icons, and hyperlinks. The cards are organized into five columns:

- Column 1:** 'STAGES' (Stages), 'Formation des usagers : FORMIST', 'serveur francophone pour l'autoformation à l'IST', and 'La Lettre de l'URFIST trimestrielle'.
- Column 2:** 'DADI' (Bases de données gratuites DADI), 'Bases de données et serveurs commerciaux', 'Bases de données d'images et de photographies', and 'Outils pour le web invisible'.
- Column 3:** 'Le détective de l'Internet' (Cours interactif), 'L'information sur Internet' (Le labyrinthe virtuel), 'Les outils de recherche sur l'Internet' (Maîtrisez les nouveaux outils), and 'Tout pour les sciences exactes' (constantes, nomenclatures ...).
- Column 4:** 'Nos supports de cours et autres lectures', 'Cours de méthodologie juridique à l'intention des étudiants de 1ère année', 'Formateurs', and 'Ressources pédagogiques pour les formations de formateurs à la recherche documentaire'.
- Column 5:** 'RISI' (Recherche d'information sur l'Internet), 'RIME' (Ressources sur Internet en Management et en Economie), and 'The Spire Project' (Spire, Guide de ressources australien en anglais).

Cette dernière option est l'œuvre de Jean-Pierre Lardy qui l'a créée et intitulée « **Les bases de données gratuites sur Internet**³ » en 1996 et en est depuis l'administrateur unique.

² <http://urfist.univ-lyon1.fr/>

³ <http://urfist.univ-lyon1.fr/gratuits/index.htm>

Partie 1 : « LES BASES DE DONNEES GRATUITES SUR INTERNET » : ANALYSE DE L'EXISTANT

1. Les bases de données gratuites disponibles sur Internet.

1.1. De l'idée à la réalisation

L'idée de créer cet annuaire est venue à Jean-Pierre Lardy d'un constat : au milieu des années 80, le développement d'Internet changeait petit à petit la donne pour la diffusion des bases de données bibliographiques et factuelles. En effet, des bases de données qui étaient auparavant accessibles de façon externe (et payante) ou interne aux universités, laboratoires de recherche et même offices de brevets devenaient consultables en accès gratuit par Internet qui, s'il ne constituait pas alors un média nouveau, a engendré une explosion des échanges d'informations et une généralisation de son utilisation.

Cependant si ces bases de données étaient disponibles sur le web, elles constituaient, et constituent toujours d'ailleurs, un type de ressource mal signalé par les moteurs de recherche car l'information y est diffusée sous forme de pages dynamiques pas toujours indexées par les robots. On en parle sous le nom de **web invisible**. Faire une recherche sur le Net en utilisant uniquement les moteurs de recherche et annuaires généralistes équivaut en effet à ignorer des milliers de sources validées et, en l'occurrence, des producteurs de base de données. C'est donc dans un souci de mise à disposition de l'information, qui est l'une des missions de l'URFIST, que l'auteur a commencé un travail de fourmi qui consiste toujours à recenser les bases de données disponibles gratuitement sur Internet et à les rassembler dans un annuaire accessible à tous gratuitement via le site web de l'Unité de formation.

1.2. Analyse de l'existant

1.2.1. L'interface visiteur et les options de recherche

1.2.1.1. La page d'accueil

Il n'a jamais été réalisé de charte graphique puisque l'interface a été créée et modifiée petit à petit par l'éditeur de la base de données. La présentation générale ainsi que les fonctionnalités et les couleurs du site ont donc évolué au gré des modifications apportées par son créateur depuis la première version mise en ligne en 1996.

La page d'accueil se compose de deux parties distinctes : le menu et la présentation du site.

Le menu : il décrit les fonctionnalités de l'outil et présente le classement des bases de données en un sous-menu de 39 catégories.

les bases de Données gratuites Disponibles sur Internet

[à propos de jpl...](#) | [proposer une base de données](#) | [feedback](#) | [liens](#)

► Interrogez la [base de données](#)

ou choisissez une catégorie :

[Agriculture \(23\)](#) - [Arts \(18\)](#) - [Astronomie \(7\)](#) - [Brevets \(18\)](#) - [Chimie \(19\)](#) - [Cinéma \(14\)](#) - [Dictionnaires \(4\)](#) - [Economie \(25\)](#) - [Environnement \(10\)](#) - [Experts \(4\)](#) - [Génétique \(25\)](#) - [Histoire \(26\)](#) - [Images et photographies, Galeries \(52\)](#) - [Images et photographies, Outils de recherche \(6\)](#) - [Informatique \(16\)](#) - [Linguistique \(5\)](#) - [Littérature \(16\)](#) - [Loisirs \(4\)](#) - [Marques \(2\)](#) - [Mathématiques \(11\)](#) - [Médecine \(70\)](#) - [Multidisciplinaire \(46\)](#) - [Musique \(23\)](#) - [Normes \(5\)](#) - [Recherche communautaire \(3\)](#) - [Répertoire-Portail \(16\)](#) - [Sciences de la vie \(69\)](#) - [Sciences de l'éducation \(8\)](#) - [Sciences exactes \(31\)](#) - [Sciences humaines \(13\)](#) - [Sciences juridiques \(12\)](#) - [Sciences politiques \(15\)](#) - [Sciences sociales \(5\)](#) - [Sport \(3\)](#) - [Technologie \(20\)](#) - [Thesaurus et Classifications \(51\)](#) - [Thèses \(11\)](#) - [Vétérinaire \(6\)](#)

- ❑ **à propos de jpl...** : ce lien renvoie à une page web⁴ présentant les travaux réalisés et les ouvrages écrits par Jean-Pierre Lardy.
- ❑ **proposer une base de données** : renvoie à un formulaire qui propose au visiteur de remplir des champs précisant la nature, l'URL, le producteur et le thème de la base de données. Un espace est réservé à l'identification du visiteur (nom et adresse électronique). Les propositions sont envoyées directement dans la boîte aux lettres électronique de l'administrateur de la base.

⁴ <http://urfist.univ-lyon1.fr/jpl.html>

Proposer une base de données gratuite...

- Indiquez les informations suivantes sur la base de données :

Nom de la base

URL

Producteur

- Sélectionnez une des catégories suivantes :
- Indiquez vos coordonnées suivantes :

Nom

Courrier électronique

- feedback** : cette option permet les commentaires et le signalement de liens morts. Un espace est ici encore réservé à l'identification du visiteur avec en plus du nom et de l'adresse électronique, des champs prévus pour recueillir le numéro de téléphone et de télécopie.

Quel type de commentaire voulez-vous envoyer ?

Problème Suggestion Éloge

Quel est l'objet de votre commentaire ?

Autre :

Tapez vos commentaires dans l'espace ci-dessous :

Indiquez vos coordonnées :

Nom

Adresse de messagerie

Téléphone

Télécopie

Veuillez me contacter dès que possible concernant ces commentaires.

- **liens** : cette page propose des liens vers une sélection des ressources (présentées sous la forme d'une bibliographie) et des outils de recherche permettant de franchir les limites du web visible et ainsi d'améliorer la pertinence de ses recherches sur la toile.
- **Page d'accueil Urfist** : ce lien renvoie au menu général du site.

La présentation : Il s'agit d'un texte rédigé par Jean-Pierre Lardy présentant le site et le contexte de sa création ainsi que son contenu (le type de bases de données répertoriées et les limites de la recherche d'information sur Internet). Cette introduction comporte de nombreux liens vers des sites de recherche d'information sur le web invisible.

1.2.1.2. *Les options de recherche*

On en dénombre deux : la sélection par catégories et la recherche d'une chaîne de caractères dans tous les champs à la fois.

La sélection par catégories : il s'agit pour le visiteur de sélectionner l'une des 39 catégories choisies par l'administrateur pour indexer les bases de données présentées dans le site. Ces catégories n'obéissent pas à une règle classificatoire et ont été rajoutées au fur à mesure des besoins. Elles sont donc très générales (agriculture, sciences humaines..).

[Agriculture](#) (23) - [Arts](#) (18) - [Astronomie](#) (7) - [Brevets](#) (18) - [Chimie](#) (19) - [Cinéma](#) (14) - [Dictionnaires](#) (4) - [Economie](#) (25) - [Environnement](#) (10) - [Experts](#) (4) - [Génétique](#) (23) - [Histoire](#) (26) - [Images et photographies, Galeries](#) (52) - [Images et photographies, Outils de recherche](#) (6) - [Informatique](#) (16) - [Linguistique](#) (5) - [Littérature](#) (16) - [Loisirs](#) (4) - [Marques](#) (2) - [Mathématiques](#) (11) - [Médecine](#) (70) - [Multidisciplinaire](#) (46) - [Musique](#) (23) - [Normes](#) (5) - [Recherche communautaire](#) (3) - [Répertoire-Portail](#) (16) - [Sciences de la vie](#) (69) - [Sciences de l'éducation](#) (5) - [Sciences exactes](#) (31) - [Sciences humaines](#) (13) - [Sciences juridiques](#) (12) - [Sciences politiques](#) (15) - [Sciences sociales](#) (5) - [Sport](#) (3) - [Technologie](#) (20) - [Thesaurus et Classifications](#) (51) - [Thèses](#) (11) - [Vétérinaire](#) (6)

En cliquant sur une catégorie, le système renvoie le visiteur à une page unique qui liste les bases de données correspondant au thème sélectionné. A chaque base correspond une description plus ou moins complète selon des critères qui seront passés en revue plus avant lors de la description de la table « bases » d'Access (nom, description, producteur...). Seul le champ URL n'apparaît pas puisque le lien vers la base de données se fait en cliquant le lien hypertexte contenu dans son nom.

Astronomie

1. [ARIBIB](#)
Universitat heidelberg
- Anglais
ARI Bibliographical Database for Astronomical References
2. [ARNIE](#)
LEDAS, the Leicester Database and Archive Service
- Anglais
Interrogation de bases en astronomie
3. [Astronomy and Astrophysics Abstract Service - ADS](#)
NASA
452 451 références - Anglais

La navigation est assurée dans cette page par des boutons « précédent » (renvoi à la catégorie précédente), « suivant » et « retour aux catégories » présents en bas de celle-ci.

Il s'agit d'un mode de recherche très général mais qui a le mérite d'être exhaustif et de renseigner rapidement sur un thème global.

Pour les besoins plus précis, on préférera l'interrogation de la base par chaîne de caractères.

La recherche d'une chaîne de caractères : Elle constitue une option récente puisqu'elle a été implantée au mois de janvier 2002. Ce mode de recherche consiste en un balayage de tous les mots contenus dans tous les champs de la notice d'une base de données, grâce à la fonction GREP⁵, afin de trouver la chaîne de caractère demandée.

Si cette fonctionnalité constitue un progrès indéniable, elle génère également beaucoup de bruit, comme l'illustrera plus tard un exemple concret.

⁵ Fonction utilisée sous Unix pour balayer une chaîne de caractère, avec troncatures à droite et à gauche, à la recherche d'un motif, en l'occurrence le mot-clé désigné par l'utilisateur.

1.3. Structure de la base de données et interface administrateur

1.3.1. Structure de la base de données

Il s'agit d'une **base de données relationnelle gérée sous Access**. Ce choix a été fait en raison de la relative simplicité d'utilisation et de gestion du logiciel et des performances qu'il offrait à l'époque pour gérer un type de données dont le volume était alors assez restreint. Nous examinerons par la suite les limites de ce choix.

Un logiciel GDIBd transforme les fichiers Access en pages html statiques classées par catégories. La base de données Access permet la saisie de fiche renseignant les références (c'est à dire les bases de données répertoriées).

Il existe deux tables : l'une pour les catégories et l'autre pour les bases de données.

La table des catégories comporte deux champs : **Num_cat** (de 1 à 39) qui fait office de clé primaire et **catégorie** (champ texte limité à 50 caractères) qui recense les catégories employées dans la base pour classer les bases de données répertoriées. Ce sont ces catégories que l'utilisateur retrouve dans le menu du site.

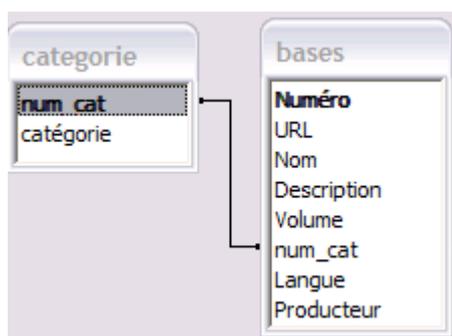
num_cat	catégorie
1	Agriculture
2	Brevets
3	Cinéma
4	Economie
5	Génétique
6	Histoire
7	Informatique
8	Littérature
9	Marques
10	Mathématiques
11	Médecine
12	Multidisciplinaire
13	Musique
14	Normes
15	Recherche communautaire
16	Sciences de la vie
17	Sciences exactes
18	Sciences humaines
19	Sciences politiques
20	Technologie
21	Thesaurus et Classifications
22	Thèses
24	Linguistique
25	Sciences de l'éducation
26	Sciences juridiques
27	Chimie
28	Images et photographies. Galeries
29	Experts
30	Sport
31	Répertoire-Portail
32	Arts
33	Sciences sociales
34	Images et photographies. Outils d
35	Dictionnaires
36	Loisirs

La table des bases comporte huit champs :

- ❑ **Numéro** : il s'agit d'un numéro automatiquement attribué par le système à une référence et qui fait office de clé primaire
- ❑ **URL** : champ texte limité à 250 caractères
- ❑ **Nom** : champ texte limité à 255 caractères
- ❑ **Description** : champ texte limité à 250 caractères
- ❑ **Volume** : champ texte limité à 50 caractères (nombre de références bibliographiques contenues dans les bases de données référencées)
- ❑ **Catégorie** : Num_cat de type numérique permettant la jonction des deux tables lors de requêtes SQL (ici : SELECT DISTINCTROW [categorie].[num_cat], [categorie].[catégorie] FROM [categorie];)
- ❑ **Langue** : champ texte limité à 50 caractères. Il fait appel à une liste de valeurs comportant des langues pré-sélectionnées :
"Anglais";"Français";"Italien";"Allemand". L'anglais est la langue qui s'affiche par défaut
- ❑ **Producteur** : champ texte limité à 125 caractères

Numéro	URL	Nom	Description	Volume	Catégorie	Langue	Producteur
942	http://www.b2btextile.v	@nnaire des Professio	Produit, savoir-faire	plus de 10 000	Economie	Français	Application On Line, s.a.r.l
518	http://www.fontshop.fr/	200 000 images libres d			Images et photi	Français	FontSHOP
888	http://www.abbott.com	Abbott Laboratories vets			Vétérinaire	Anglais	Abbott Laboratories
854	http://www.chemie.fu-t	Abbreviations of Chemic			Chimie	Anglais	Université de Berlin RFA
799	http://www.accessible	Accessible Archives	Primary Source Mat	6 bases de don	Histoire	Anglais	Accessible Archives, Inc.
666	http://www.accpuf.org/	ACCPUF	Droit constitutionnel		Sciences juridi	Français	Association des Cours Constitutionnell
558	http://www.agora21.orj	Activités de recherche fi			Economie	Français	Association des Pratiques du Développ
357	http://boree.cnusc.fr/3	ACUBASE	Références d'article		Médecine	Français	
826	http://www.cerc.usgs.g	Acute Toxicity Databasi	Tests de toxicité ag		Sciences de la	Anglais	Columbia Environmental Research Cent
684	http://scriptorium.lib.d	Ad*Access Project	Publicités parues ds	plus de 7 000 p	Images et photi	Anglais	Duke University's Durham North Carolin
742	http://cca.qc.ca/adhen	Adhémair	Propriété, bâti et poj		Arts	Français	Centre Canadien d'Architecture
355	http://www.fda.gov/sea	Agence US du médicam			Médecine	Anglais	U.S. Food and Drug Administration
919	http://www.africultures	Agenda des événements	tous les événements		Multidisciplinair	Français	Africultures
891	http://www.agric.org/	AgNIC	AgNIC is a guide to		Agriculture	Anglais	AgNIC
577	http://www.dainet.de/d	AgriBib	Permet l'interrogatio		Agriculture	Allemand	ZADI - Zentralstelle für Agrardokumenta
250	http://www.nal.usda.g	Agricola	Tous les aspects de		Agriculture	Anglais	National Agricultural Library
251	http://www.fao.org/agn	AGRIS	Sciences et technol	2,7 millions de	Agriculture	Anglais	FAO - Organisation des Nations Unies
1015	http://www.inra.fr/agn	Agritox	Propriétés physique		Agriculture	Anglais	INRA - Centre de Versailles-Grignon
358	http://www.med.univ-er	Aide au Diagnostic Méd	Descriptions de patf	plus de 10 000	Médecine	Français	Laboratoire d'Informatique Médicale de I

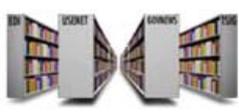
Les liaisons entre les tables s'opèrent de la manière suivante :



1.3.2. L' interface administrateur

Afin de faciliter l'alimentation de la base de données, un formulaire plus convivial que la table « bases » a été créé, reprenant les champs des intitulés dans une interface plus parlante :

Catégorie	Agriculture	Les bases de données gratuites sur internet	
Nom	AgriBib		
Description	Permet l'interrogation de plusieurs bases de données allemandes en agriculture		
URL	http://www.dainet.de/dain/service/agribase/index.htm		
Langue	Allemand	Producteur	ZADI - Zentralstelle für Agrardokumentation und -information
Volume			



Aucun champ n'est obligatoire. Seuls les champs « catégorie » et « langue » sont contrôlés par une liste de choix ; les autres peuvent être remplis par l'administrateur à sa guise. Aucune règle de catalogage n'a été établie (format de date, présentation des noms de bases, casse...).

1.4. Bilan du système documentaire

Le système présente deux handicaps majeurs :

Le premier concerne la **gestion technique** de la base de données sous Access. Il s'agit d'un logiciel tout à fait performant à un niveau individuel et pour une utilisation restreinte mais qui ne convenait plus dans le cas présent pour plusieurs raisons. D'une part, les limites de l'exportation de données ont commencé à se faire sentir lors de l'utilisation en réseau de bases de données ; le développement d'Internet et de systèmes de gestion beaucoup plus puissants ont rendu ce logiciel inadéquat. Les « bases de données gratuites sur Internet » n'échappent donc pas à cette règle. D'autre part, Access est adapté à de petites quantités de données mais dès que leur volume s'accroît, la gestion des tables et des liens entre elles devient périlleuse et les listes par catégories s'allongeant, le temps de recherche d'une référence s'allonge également.

La deuxième remarque vise la **gestion documentaire** de la base. En l'absence de classification et de règles de catalogage, les notices telles qu'elles apparaissent à l'utilisateur manquent d'homogénéité (voir l'exemple des bases en astronomie ci-dessous). De plus la description des bases de données présentées était souvent succincte et issue d'une description fournie par le producteur de la base lui-même ou par le moteur de recherche l'indexant. Dans tous les cas, ce commentaire se révélait trop souvent sommaire et peu représentatif du produit documentaire, en particulier lorsqu'il s'agissait de comparer plusieurs bases sur un même thème. Par exemple, dans la catégorie astronomie, les trois premiers résultats qui apparaissent sont :

[ARIBIB](#)

Universitat heidelberg

- Anglais

ARI Bibliographical Database for Astronomical References

[ARNIE](#)

LEDAS, the Leicester Database and Archive Service

- Anglais

Interrogation de bases en astronomie

[Astronomy and Astrophysics Abstract Service - ADS](#)

NASA

452 451 références – Anglais

Il semble donc difficile de faire son choix sans être spécialiste.

Enfin, les possibilités de requêtes et d'exploitation de celles-ci se sont avérées trop peu nombreuses pour suivre la nécessaire évolution de la base et l'accroissement perpétuel du nombre de références (716 au début de mon stage).

Concrètement, le champ « description » ne peut comporter plus de 256 caractères pour permettre une gestion pertinente des mots-clés, c'est à dire l'inverse de la situation actuelle où tous les mots de tous les champs sont considérés par le système comme des mots-clés alors qu'il s'agit en réalité de chaînes de caractères vides de sens pour le système. Ce fut d'ailleurs l'argument documentaire plus que technique qui a conduit à la décision de modifier le système.

La principale difficulté de la gestion de la base de données vient du fait que Jean-Pierre Lardy est le seul à l'administrer et qu'il ne s'agit pas de sa fonction principale. Ainsi, lorsque de nouvelles références lui sont communiquées ou qu'il en trouve lui même, il a tout d'abord le souci de les porter à la connaissance des visiteurs du site et l'organisation de ces dernières pâtit donc du manque de temps accordé à leur référencement.

Il est donc apparu nécessaire après cinq années d'existence d'effectuer une mise à jour de l'annuaire et de le faire évoluer en fonction d'exigences techniques et documentaires que je vais décrire.

Partie 2 : Mise à jour du répertoire et réorganisation du site : des besoins et des solutions.

1. Analyse des besoins

La rapidité et la simplicité de l'alimentation de l'annuaire ainsi que l'efficacité des recherches étaient des éléments fondamentaux dans la réflexion sur les besoins et les solutions à apporter.

1.1. Changement de la structure de la base de données

Un choix informatique devait être fait afin de pallier les défauts de la base de données gérée sous Access : la consultation se faisait par choix de catégories ; elle devait être améliorée car la limite évidente de ce système est une liste trop longue à consulter de manière séquentielle. De plus, la base de données Access n'était pas consultable directement en ligne.

La question devait donc être rapidement abordée avec l'informaticien chargé de la programmation de la nouvelle application.

1.2. Vérification des liens

Toutes les bases de données recensées dans l'annuaire des bases de données gratuites sur Internet possèdent une URL qui est susceptible de changer avec le temps ou même de disparaître. En effet, le site mettant sa base de données à disposition des internautes peut changer d'hébergeur ou d'adresse. Il arrive dans ce cas que le changement soit signalé à l'ancienne adresse et une redirection est alors effectuée, la plupart du temps de façon automatique. Il existe pour cela des outils permettant de détecter automatiquement les **liens morts** mais ils ne sont malheureusement pas efficaces pour tous les types d'erreurs.

C'est le cas par exemple des liens erronés qui peuvent aller de la simple « erreur type 404 » (adresse inconnue) à des situations cocasses lorsque l'adresse a été rachetée et que l'on y découvre un site traitant d'un tout autre domaine. L'URL conduit bien à une page valide donc le système ne détecte pas d'erreur.

Il est donc nécessaire de vérifier périodiquement « à la main » la validité des liens.

1.3. Recherche de nouvelles bases de données

Lorsque j'ai pris mes fonctions le 3 juin, l'annuaire comptait 716 bases de données. L'un des objectifs du stage devait être la recherche de nouvelles bases afin d'agrandir l'annuaire. Pour cela, plusieurs moyens devaient être testés.

Parallèlement aux moteurs de recherche et annuaires généralistes il existe sur le Net un certain nombre de guides, d'annuaires sélectifs, de sites fédérateurs, de portails thématiques et d'agents de veille, qui franchissent les limites du web visible et interrogent ou identifient les sites du web invisible. Les bases de données n'étant pas toujours indexées par les moteurs de recherche pour les raisons exposées plus haut, cet aspect de la recherche documentaire devait être primordial durant le stage.

1.4. Veille informationnelle

Il m'a également été demandé de tester le logiciel de veille Copernic et de décider s'il se révélait approprié à la recherche de nouvelles références par des équations générant les meilleurs résultats.

1.5. Indexation intellectuelle

L'annuaire ne bénéficie pas, nous l'avons vu, d'un système de classification plus élaboré qu'une liste de thèmes ou catégories, ajoutés au fur et à mesure des besoins. Il a donc été décidé qu'il fallait trouver et mettre en place un système de classement pertinent pour référencer correctement les bases.

1.6. Présentation des notices

Il est apparu évident qu'une fois les bases de données convenablement indexées, il faudrait les présenter au visiteur sous la forme de notices le plus détaillées possible et standardisées pour plus de clarté.

1.7. Recherche en ligne

Le système de recherche actuel, tel qu'il a été décrit, ne permet pas une exploitation pertinente des mots-clés. L'une de mes tâches devait donc consister à trouver un système de gestion de mots-clés adéquat et à les choisir pour décrire les nouvelles comme les anciennes base de données.

Sans description exacte de ce que devrait être la nouvelle interface de recherche, il a été précisé que l'essentiel était de permettre d'améliorer au maximum les requêtes par le croisement de critères.

1.8. Nouvelles interfaces web

Il a été décidé que le site web, accessible librement au public, devait être refondu pour donner au nouveau système une interface de recherche adaptée et fonctionnelle mais aussi que l'interface administrateur devrait être modifiée afin de bénéficier d'une gestion aussi standardisée que possible dans le but de maintenir l'annuaire à jour facilement (contrôle à la saisie grâce à des listes prédéfinies).

1.9. Un site bilingue : français-anglais

Une grande nouveauté du site serait aussi une localisation⁶ en anglais, qui n'existait pas jusqu'alors. En effet, une majorité de bases de données sont de production anglo-saxonne et une majorité d'internautes à la recherche d'informations en ligne sont anglophones. C'est un aspect qu'il convenait effectivement de ne pas ignorer car il semblait constituer une promesse d'accroissement du nombre des visiteurs du site.

1.10. Sigle et charte graphique

En ce qui concerne l'habillage du site, il fallait aussi trouver un nom plus facilement identifiable que « les bases de données gratuites sur Internet » et définir une charte graphique.

⁶ C'est-à-dire une version du site et du système de recherche en anglais.

1.11. Promotion

En dernier lieu, une promotion devait être organisée pour informer le public de la mise en ligne de la nouvelle version du site et de ses nouvelles fonctions. Il fallait dresser une liste des différents publics à toucher et trouver des moyens matériels de communication.

2. Mise à jour de l'annuaire

2.1. Vérification des liens

Ce fut la première tâche réalisée. Elle a consisté en une vérification de la validité des liens des 716 bases référencées dans l'annuaire des « Bases de données gratuites sur Internet ».

J'ai donc constitué un fichier avec les adresses actuelles et leurs adresses corrigées lorsque cela s'imposait.

CIB officielle

OMPI - Organisation Mondiale de la Propriété Intellectuelle
- Français

En vigueur depuis le 1/1/1995

Adresse actuelle : <http://classifications.wipo.int/fre/main.htm>

Adresse correcte : http://www.wipo.int/classifications/fulltext/new_ipc/index.htm

Certaines bases de données n'étaient plus accessibles et ont été signalées comme telles pour être supprimées ultérieurement de l'annuaire.

2.2. Recherche de nouvelles bases de données

Outre le « ménage » des références de l'annuaire, une extension du choix a été faite en recherchant de nouvelles bases de données. Le but était de renseigner des domaines jusque là pauvres en références (en particulier les sciences humaines).

Les moyens ont été multiples : interrogation de moteurs de recherche, métamoteurs de recherche (online et offline), annuaires généralistes et spécialisés et portails de type « open directory project »⁷.

⁷ Voir la liste, non exhaustive, des URL dans la bibliographie.

Quantité de bases gratuites ont été trouvées et 300 d'entre elles figurent désormais dans l'annuaire, portant le nombre de références proposées à 1000.

Les bases de données scientifiques sont les plus nombreuses car nombre d'universités et d'organismes de recherche mettent à disposition de tous les chercheurs le fruit de leurs travaux via Internet. Les ministères sont aussi de gros producteurs de bases et en particulier le ministère de la Culture en France.

On trouve une large majorité de bases de données en anglais (environ les deux tiers du total).

2.3. Mise en place d'une veille avec Copernic

Copernic est sans doute l'un des agents de veille les plus connus au monde. C'est un logiciel canadien convivial et performant.

La version de Copernic 2000 sur laquelle j'ai travaillé est en fait un métamoteur offline qui permet d'interroger plus de 85 sources d'informations (moteurs, annuaires et banques de données) grâce à la création de requêtes.

J'en ai testé plusieurs dans plusieurs langues et il apparaît que la simple combinaison des termes DATABASE + le thème recherché et DATA + BASE + le thème recherché ramène de nombreux résultats intéressants.

Mais la plus grande surprise est venue d'un essai « naïf » : BASE(S) + DONNEES (également testé en anglais DATABASE(S)) : cette requête a été une source quasi inépuisable de références souvent très pertinentes que j'ai mis plusieurs semaines à exploiter.

2.4. Nouvelle indexation

Après un examen des catégories servant à indexer les bases de données, ainsi que des thèmes des bases elles-mêmes, il est apparu qu'il était nécessaire de mettre en place une classification plus fine et débarrassée de toute ambiguïté.

En effet, les catégories, au nombre de 39, semblaient trop larges pour indexer plus de 700 bases de données.

La principale question qui s'est posée au moment de faire le choix de la classification a été : faut-il en exploiter une existante ou bien convient-il d'en créer une nouvelle ? Cette dernière option a rapidement été écartée pour deux raisons : la

première est une question de temps car la création d'un index ordonné qu'il aurait inmanquablement fallu mettre en œuvre aurait pris énormément de temps au point de faire l'objet d'une bonne partie du stage. La deuxième raison tient au fait que l'annuaire des bases de données gratuites étant généraliste et donc pluridisciplinaire, une classification de type bibliothèque pouvait très bien faire l'affaire.

Une fois ces considérations faites, j'ai pensé qu'étant donné la volonté du site de s'adresser au plus grand nombre et pas uniquement à un public de spécialistes, il fallait trouver un système qui soit identifiable rapidement et dans lequel chacun puisse rapidement trouver l'information qu'il recherche.

La classification décimale Dewey a été adoptée pour ces raisons.

La classification décimale Dewey :

Créée aux Etats-Unis au 19^{ème} siècle, il s'agit d'une méthode d'organisation des connaissances qui porte le nom de son inventeur Melvil Dewey et qui répartit le savoir en dix grandes classes (qu'on appellera ici **catégories primaires**), à leur tour divisées en cent sous-classes (ici, **catégories secondaires**), puis en une multitude de sections et sous-sections⁸.

Employée dans la majorité des bibliothèques universitaires et municipales en France et dans le monde, cette classification est la plus répandue et la plus connue, même implicitement, de tous. Quiconque a visité une bibliothèque a certainement utilisé ce classement pour trouver l'ouvrage recherché.

La popularité de cette classification constituait donc un atout promotionnel pour le futur site en réduisant le temps d'adaptation des usagers.

A la fois larges et précises, ses catégories la rendaient ergonomique au genre de bases de données qu'il convenait d'organiser.

Au cours de ces quatre mois de stage, j'ai effectué l'indexation des nouvelles bases trouvées (environ 300) et la réindexation de celles qui étaient déjà référencées (environ 700).

Le choix d'adopter une entrée unique a été fait pour faciliter l'administration de l'annuaire. Lorsque la base de données est polysémique, le thème dominant sert à l'indexer et les thèmes secondaires sont décrits dans les mots-clés. Ainsi, on améliore la précision des requêtes lors de l'interrogation.

3. La nouvelle administration de l'annuaire

3.1. Création d'une nouvelle notice

Afin de perfectionner le système, il a été décidé d'étoffer la notice renseignant chaque base de données entrée dans l'annuaire en créant de nouveaux champs et en modifiant le type de données d'autres.

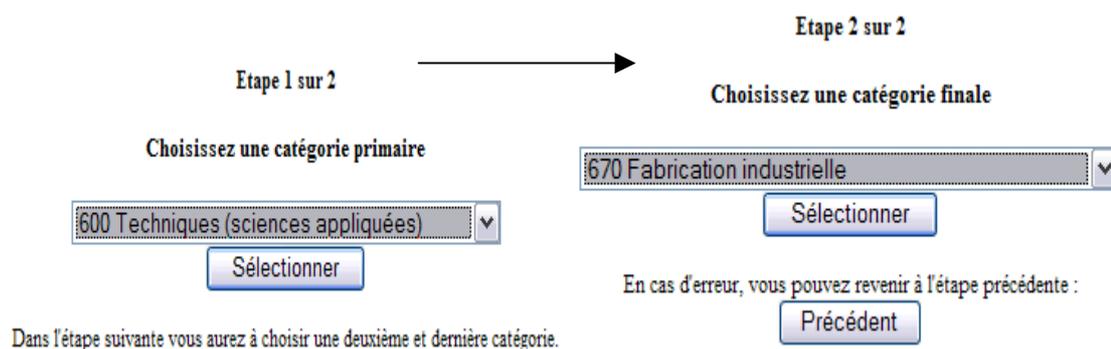
Rappelons les anciens champs : catégorie, nom, description, URL, langue, volume et producteur. Leurs fonctions ont été décrites lors de l'analyse de l'existant.

Voici les modifications, champ par champ, qui ont été apportées :

(Les champs marqués d'une astérisque sont obligatoires)

Catégorie (*):

Ce champ n'est pas libre et fait appel à deux listes de choix accessibles grâce au bouton « choisir » accolé. Une fenêtre apparaît permettant de choisir une catégorie primaire de la classification Dewey dans un premier menu puis une catégorie secondaire dans un second menu. La catégorie secondaire ainsi que son numéro s'inscrit alors dans le champ « catégorie » après validation :



⁸ Le détail des cent premières classes, utilisées ici, se trouve en annexe I.

Nom (*):

@nnuaire des Professionnels du Textile

Ce champ est libre et le nombre de caractères n'est pas limité. Il a été établi des règles de catalogage qui régissent la saisie du titre de chaque base⁹.

Description (*):

Consultez gratuitement l'annuaire on-line où sont référencées plus de 12000 entreprises textiles. L'annuaire est multilingue et les informations apportées sont mises à jour régulièrement. Vous recherchez un produit, un savoir-faire ou le partenaire commercial idéal, consultez l'Annuaire.

Ce champ n'est pas limité en caractères et est totalement libre. Il doit cependant comporter un maximum d'informations sur la base et, de préférence, pas uniquement le résumé succinct écrit par son producteur.

Mots clés en français (séparateur = virgule) (*):

textile, annuaire professionnel

Les mots-clés sont choisis en raison de leur pertinence par l'administrateur et doivent décrire au mieux la base de données nouvellement indexée. Leur nombre n'est pas limité mais un au moins est obligatoire.

Ce champ n'est pas limité en caractères mais n'est pas totalement libre car il doit y avoir correspondance entre le nombre de mots en français et en anglais. Un total divergent empêche la validation de la référence et affiche un message d'erreur.

Le séparateur entre les mots-clés est la virgule et là encore les règles de catalogage sont strictes et détaillées dans l'annexe II.

Mots clés en anglais (séparateur = virgule) (*):

textile, professional directory

⁹ Voir annexe II

Idem. Le site web devenant bilingue, la recherche par mots-clés le devient aussi. Il convient donc de tous les traduire.

URL (*):

Champs libre et non limité.

Nombre de référence(s) en français :

Ce champ n'est pas limité en caractères et laisse la possibilité d'écrire en toutes lettres, en chiffres ou de combiner les deux. Il n'est pas obligatoire car l'information n'est pas toujours donnée par le producteur de la base et impossible à vérifier par soi-même.

Nombre de référence(s) en anglais :

Idem. Le site web devenant bilingue, le nombre de références s'inscrit en français ou en anglais selon l'interface choisie. Il convient donc de traduire toutes les expressions.

Producteur(s) (séparateur = virgule) :

Ce champ n'est pas limité en caractères. L'index offre la possibilité de sélectionner un producteur parmi une liste alphabétique si celui-ci est déjà connu.

Liste des producteurs référencés dans notre base



- A&E Television Networks
- AAS
- Abbott Laboratories
- ABES Parc Euromédecine 25, rue Guillaume Dupuytren 34196 Montpellier Cedex 5
- ABES Parc Euromédecine 25, rue Guillaume Dupuytren 34196 Montpellier Cedex 5
- Accessible Archives
- Accessible Archives Inc. 697 Sugartown Road Malvern PA 19355
- Accessible Archives, Inc. 697 Sugartown Road Malvern, PA 19355
- ACCPUF
- AccuWeather, Inc. 385 Science Park Road State College, PA 16803
- Accessible Archives
- Adminet
- AFNOR FR
- Agricultures
- Agence pour l'énergie nucléaire (AEN)
- Agence Rhône-Alpes pour le livre (ARAL)
- Agence Rhône-Alpes pour le livre (ARAL)
- Agence Universitaire de la Francophonie
- Agences de l'eau - Direction de l'Eau du Ministère de l'Environnement
- AgNIC
- Agora21 ARMINES, Ecole Nationale Supérieure des Mines de Saint Etienne
- Agrovoc

En cliquant sur son nom, l'information s'affiche automatiquement dans le champ. Il peut y avoir plusieurs producteurs et, en ce cas, il convient d'utiliser la virgule en guise de séparateur.

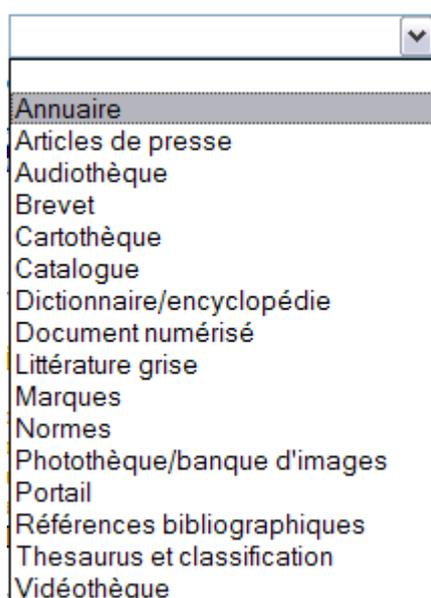
Langue(s) (*) : on peut cocher autant de cases que nécessaire. « Langue » signifie celle(s) dans laquelle ou lesquelles la base est produite et peut-être interrogée. C'est pourquoi certaines bases multilingues (par exemple produites par l'Union Européenne) peuvent comporter six langues ou plus. L'option des cases à cocher est apparue comme une alternative à la liste de choix dans laquelle on ne peut sélectionner qu'une seule langue.

Langue(s) (*) :

<input type="checkbox"/>		(Allemand)
<input checked="" type="checkbox"/>		(Anglais)
<input type="checkbox"/>		(Arabe)
<input type="checkbox"/>		(Chinois)
<input checked="" type="checkbox"/>		(Espagnol)
<input checked="" type="checkbox"/>		(Français)
<input checked="" type="checkbox"/>		(Italien)
<input type="checkbox"/>		(Japonais)
<input type="checkbox"/>		(Portuguais)
<input type="checkbox"/>		(Russe)

Type de ressource (*) :

Ce champ est contrôlé par une liste de choix modifiable de catégories de ressource.



A dropdown menu with a scroll bar, showing a list of resource categories. The 'Annuaire' category is highlighted. An arrow points from the text above to the dropdown menu.

- Annuaire
- Articles de presse
- Audiothèque
- Brevet
- Cartothèque
- Catalogue
- Dictionnaire/encyclopédie
- Document numérisé
- Littérature grise
- Marques
- Normes
- Photothèque/banque d'images
- Portail
- Références bibliographiques
- Thesaurus et classification
- Vidéotheque

C'est, outre la gestion des mots-clés, la grande nouveauté de la nouvelle version du système. Il est attribué à chaque base de données un type de ressource selon son type de données : annuaire, articles de presse, référence bibliographique, brevets, carte, catalogue, dictionnaire/encyclopédie, document numérisé, littérature grise, marques, normes, photo, portail, thesaurus ou classification, vidéothèque... la liste n'est pas exhaustive et augmentera certainement au fil des besoins.

Cette double indexation permet au visiteur d'affiner considérablement sa requête si il a une idée précise de ce qu'il cherche.

3.2. Une gestion simplifiée

L'administration de la base a été simplifiée car accessible par mot de passe de n'importe quel poste relié à l'Internet¹⁰. Toutes les fonctionnalités apparaissent dans un menu interactif.

Catégories : [[lister](#)] [[ajouter](#)]
Types de répertoire : [[lister](#)] [[ajouter](#)]
Langues : [[lister](#)] [[ajouter](#)]
Notices : [[lister](#)] [[ajouter](#)]
Traductions : [[lister](#)] [[ajouter](#)]
Statistiques : [[toutes](#)]
Accueil : [[administrateur](#)] [[visiteur](#)]

On peut ainsi en un simple clic :

- Lister, modifier, ou ajouter une **catégorie**. Bien que la Dewey soit un système de classification fini, des modifications d'intitulé peuvent survenir (et sont déjà survenues) au fil du temps. Ainsi, il m'a semblé plus prudent de conserver l'option de la modification. Dans le cas d'un ajout, les noms des catégories primaires et secondaires devront être renseignés en français et en anglais.
- Lister, modifier, ajouter ou supprimer un **type de ressource**. Dans le cas d'un ajout, le nom du type de ressource devra être renseigné en français et en anglais. Dans le cas d'une suppression, le système prévient que toutes les références correspondant à ce type de ressource seront logiquement également supprimées. Mais quoi qu'on fasse, un message récapitule toujours l'action entreprise et demande une confirmation avant suppression. Le risque d'erreurs irrémédiables est donc très limité.

- Lister, modifier, ajouter ou supprimer une **langue**. Dans le cas de l'ajout d'une langue il sera demandé d'écrire son nom en français et en anglais et de choisir un drapeau, pour la symboliser, parmi un fichier d'images enregistrées en local¹¹.
- Lister, modifier, ajouter ou supprimer une **notice**. Un système de navigation dans la liste des références a été créé : par page et par ordre alphabétique.

Titre	Mise à jour
1. @nuaire des Professionnels du Textile	[supprimer] [modifier]
2. AAT	[supprimer] [modifier]
3. Abbott Laboratories	[supprimer] [modifier]
4. Abbreviations of Chemical Compounds	[supprimer] [modifier]
5. Access to Asian vegetables	[supprimer] [modifier]
6. Accessible Archives	[supprimer] [modifier]
7. ACCPUF	[supprimer] [modifier]
8. Acteurs culturels en Rhône-Alpes	[supprimer] [modifier]
9. Acute Toxicity Database	[supprimer] [modifier]
10. Ad*Access Project	[supprimer] [modifier]
11. Adhémar	[supprimer] [modifier]
12. Adminet Africa	[supprimer] [modifier]
13. African-American Newspapers : the 19th century	[supprimer] [modifier]
14. Afriphoto	[supprimer] [modifier]
15. Agence US du médicament (FDA)	[supprimer] [modifier]
16. Agenda des événements culturels africains	[supprimer] [modifier]

Il y a 16 bases affichées.
[page 1 sur 54]

[suivante]

[1][2][3][4][5][6][7][8][9][10][11][12][13][14][15][16][17][18][19][20][21][22][23][24][25][26][27][28][29][30][31][32][33][36][37][38][39][40][41][42][43][44][45][46][47][48][49][50][51][52][53][54]

Il y a 864 base(s) référencé(s) pour la lettre sur un total de 864.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

- Lister, modifier, ajouter ou supprimer une **traduction** : tous les textes et leurs traductions contenus dans le site visiteur sont listés dans un fichier d'administration ce qui permet, sans connaître le langage de programmation, de modifier directement le site web.

¹⁰ <http://urfist.univ-lyon1.fr/gratuits/admin.php>

¹¹ Pour plus de précisions, rendez-vous dans la partie traitant de l'architecture informatique en troisième partie.

4. Le nouveau site

4.1. Un nouveau nom

Le nouveau site a été rebaptisé car « les bases de données gratuites sur Internet » était une appellation d'usage qui avait le mérite de décrire le produit de façon explicite mais était trop longue et difficile à manier. Un sigle s'imposait.

« **DADI** » a été adopté en raison de sa consonance sympathique et de sa facilité de mémorisation ; en outre, ce sigle a une signification aussi bien pour l'utilisateur anglophone que francophone, ce qui était une contrainte dans le choix à effectuer :

**répertoire des bases de Données
grAtuites Disponibles sur Internet**

DAtabases DIrectory

4.2. Schéma des liens

(Cf au dos)

4.3. Le plan du site

4.3.1. Le plan du site

Le menu prend place, dans la langue demandée, dans une barre verticale fixe à gauche de l'écran et reste apparente dans chaque page du site.



- **Accueil** présente le site : introduction au site et à la classification décimale Dewey, liens vers des sites pointant sur DADI et vers des outils de recherche dans lesquels le site est bien positionné, liste des technologies utilisées, quelques statistiques, quelques liens vers des rubriques complémentaires.

❑ Recherche

- Par **navigation** dans la classification Dewey
- Par **recherche avancée**

(Ces deux fonctions seront détaillées ultérieurement)

- Notice **aléatoire**

Affichage aléatoire



Voici une notice tirée au hasard dans DADI.

Document numérisé Un problème ? 

➔ Les arts > Arts graphiques, Gravures

Base Estampes

La base Estampes de la Bibliothèque municipale de Lyon a pour but de reproduire et décrire les estampes conservées dans la collection de la bibliothèque municipale de Lyon, qui contient plusieurs dizaines de milliers de pièces.

<http://sgedh.si.bm-lyon.fr/dipweb2/esta/estampes.htm>
Bibliothèque municipale de Lyon Plusieurs milliers de références

 Référencée le 16/07/2002 - Dernière mise à jour le 16/07/2002

- Le **top 10** des notices les plus populaires

Le top 10 des bases les plus populaires



Voici le classement des 10 bases gratuites les plus populaires de notre site. Cliquez sur le nom pour en afficher la notice via le module de recherche avancée.

DADI Top 10

Rang	Nom	Décompte
1	Asiaweek	6
2	Free e-mail provider guide	4
3	Les bibliothèques	2
4	Base de données relative au droit anglais	1
5	Black History	1
6	California Ballot Propositions	1
7	Castro Speech	1
8	Child Abuse and Neglect	1
9	Deep Sky Database	1
10	Extra-solar Planets Catalog	1

□ Contact

- **Un problème ?** qui affiche un formulaire permettant de signaler un problème affectant une notice

Un problème ?

Malgré nos nombreuses vérifications périodiques, il se peut que certains liens affichent une erreur de type "Erreur 404, fichier non trouvé". Cela provient souvent d'un changement d'adresse du site en question, ou bien celui-ci n'existe plus. Si vous avez trouvé une erreur de ce type, vous pouvez nous la rapporter en remplissant le formulaire ci-dessous.

Nom de la base de données : **Agora21**

Adresse du site non trouvé : **<http://www.agora21.org/recherche/>**

Si c'est un changement d'adresse et que vous connaissez la nouvelle, merci de nous l'indiquer ici :

S'il s'agit d'un autre type d'erreur, merci de nous la rapporter brièvement :

Votre nom :

Votre adresse e-mail :

- **Commentaires** permettant de faire des commentaires divers

Vos commentaires et questions

Nous recherchons continuellement à améliorer DADI alors si vous avez des suggestions ou des remarques à faire, remplissez le formulaire ci-dessous. Merci pour votre contribution !

Votre nom :

Votre adresse e-mail :

Commentaire (*) :

(*) Les champs marqués d'un astérisque (*) sont obligatoires.

- **Proposer une base de données** permettant de suggérer à l'administrateur l'ajout d'une nouvelle base de données

Proposer une nouvelle base de donnée

Nous souhaitons voir augmenter perpétuellement le nombre de bases de données répertoriées dans DADI alors c'est avec plaisir que nous accueillons vos propositions dans le formulaire ci-dessous. Merci pour votre contribution.

Nom de la base de données : (*) :

URL (*) :

Sujet :

Votre nom :

Votre adresse e-mail :

(*) Les champs marqués d'un astérisque (*) sont obligatoires.

- **Aide** sur la navigation, la recherche avancée et un modèle de notice
- **Notice légale** listant le personnel participant à la création et à l'administration du site. Contient aussi une liste de bannières diverses pour la promotion du site

Notice légale

Organisme : URFIST de Lyon - Université Claude Bernard Lyon 1 

Directeur de la publication : Jean Pierre LARDY  

Conception du système documentaire et mise à jour : Marianne QUIOC 

Développement de l'application : Hugo ETIEVANT  



4.4. Trois options de recherche

Comme nous l'avons vu, les possibilités de recherche étaient limitées.

Le logiciel de gestion de base de données Access a été abandonné afin de rendre possible, entre autres aspects documentaire et technique, l'exploitation d'un index de mots-clés. La nouvelle base de données a été créée sous MySQL qui offre entre

autres avantages un champ description exploitable illimité et la possibilité d'en gérer les mot-clés .

Cela constitue le principal progrès du nouveau système documentaire de recherche.

4.4.1. La recherche avancée

DADI est insensible aux accents et à la casse (majuscules et minuscules peuvent être utilisées indifféremment). On peut combiner tous les critères. Tous les champs sont facultatifs, mais au moins un doit être saisi.

Sept champs sont dédiés à la recherche :

Catégorie : 070-079 Journalism, édition, journaux

Type de répertoire : Document numérisé

Vous connaissez le nom de la base de données : archidoc

Un mot contenu dans la description : bible

Producteur : ministère de la culture

Mots clés (avec booléens ET OU SAUF) :

architecture

ET art contemporain

OU art moderne

Langues (opérateur OU implicite) :

<input type="checkbox"/>		(Allemand)	<input checked="" type="checkbox"/>		(Français)
<input checked="" type="checkbox"/>		(Anglais)	<input type="checkbox"/>		(Italien)
<input type="checkbox"/>		(Arabe)	<input type="checkbox"/>		(Japonais)
<input type="checkbox"/>		(Chinois)	<input type="checkbox"/>		(Portuguais)
<input type="checkbox"/>		(Espagnol)	<input type="checkbox"/>		(Russe)

1- Catégorie : Sélectionner une catégorie primaire ou secondaire de la classification Dewey parmi les 100 proposées. Liste de choix est obligatoire.

2- Type de ressource : Le type de ressource attribué par l'équipe de DADI à une base de données ; par exemple, base de données bibliographique, photothèque, brevets...

3- Vous connaissez le nom de la base de données : Nom, titre ou sigle de la base de données, même incomplet.

4- Un mot contenu dans la description : Taper un mot ou une expression.

Cette recherche d'une de caractères, qui existait déjà dans la précédente version du site, était intéressante afin de garder une vision générale du fonds documentaire. La recherche a cependant été volontairement limitée au champ « description » en raison de l'organisation des autres champs, facilement interrogeables.

Mais ce système génère beaucoup de **bruit** et de **silence** comme le montre l'exemple suivant :

Imaginons un utilisateur tapant « *inde* » pour trouver des bases de données relatives à l'Inde. Voici les résultats qu'ils obtient : des notices de bases dont le champ description contient les mots : *index*, *indépendance*, *indemne*, etc. Aucune des 38 bases affichées ne répond au critère demandé. Or, il existe bien une base de données sur le thème de l'Inde mais la description parle de « *péninsule indienne* » et ne comprend donc pas la chaîne de caractères recherchée.

Ce type de recherche de chaîne de caractères a tout de même son utilité mais pour pallier ses lacunes, j'ai instauré un système de traitement de mots qui sont considérés comme des entités et permettent donc une recherche « intelligente »

5- Producteur : on peut taper un nom de producteur (organisme, université, association...) ou en choisir un dans la liste consultable en cliquant sur « index »

6- Mot(s)-clé(s) : Recherche un mot ou une combinaison de mots parmi ceux choisis par l'équipe de DADI pour décrire chaque base de données. Afin de guider la requête, nous avons mis à disposition un index alphabétique des descripteurs utilisés par DADI.

Descripteurs utilisés pour indexer les bases de données répertoriées par DADI



abréviation
Academy Award
acoustique
acteur
actrice
actualité
adult literacy
affaires
affiche
africain-américain
Afrique
Afrique francophone
agence de presse
agriculture
agroalimentaire
Albert Einstein
Algérie
alimentation
Allemagne
Amérique centrale
Amérique coloniale
Amérique du sud
analyse sociale

Architecture
architecture
architecture civile
architecture religieuse
archive
archives
Archives Nationales
archéologie
ARIELS
armement
Art
art
art contemporain
art du spectacle
article
article numérisé
articles
artiste
arts décoratifs
ASCA
Asie
astrométrie
astronaute

En cliquant sur l'un d'entre eux il s'affichera automatiquement dans le champ "mot-clé" associé.

Cette option autorise la construction de requêtes incluant au maximum trois conditions et utilisant les opérateurs booléens ET, OU et SAUF ; cependant si deux mots consécutifs sont tapés, seules les notices de bases de données contenant ces deux mots seront retrouvées (i. e. l'opérateur par défaut est ET).

Les critères peuvent être modifiés dans le formulaire même après l'affichage des notices.

Les troncatures à droite et à gauche sont autorisées. Aucun caractère n'est nécessaire pour les signaler. Par exemple :

Psych trouvera psychiatrique, psychiatrie, psychologie, etc.

ologie trouvera biologie, muséologie, psychologie, etc.

Le système de recherche de DADI ne permet pas la recherche par adjacence.

Lorsque deux mots sont tapés consécutivement, par exemple **bibliothèque sciences** DADI l'interprète comme signifiant **bibliothèque ET sciences**.

7- Langue(s) : Cocher la case correspondant à la langue dans laquelle on souhaite que DADI trie les résultats. On peut sélectionner autant de langues qu'on le souhaite.

4.4.2. La navigation dans les catégories de la classification Dewey

Deux possibilités s'offrent à l'utilisateur :

1) Sélectionner l'une des 10 catégories primaires disponibles de la classification Dewey en cliquant sur son intitulé.

000 Généralités (9)	[nouveautés]
010 Bibliographie (1)	[nouveautés]
020 Bibliothéconomie et sciences et de l'information (1)	[nouveautés]
030 Encyclopédies générales (2)	[nouveautés]
040 (Cette catégorie n'est plus renseignée) (0)	[nouveautés]
050 Publications en série d'ordre général (0)	[nouveautés]
060 Organisations générales et muséologie (0)	[nouveautés]
070 Journalisme, édition, journaux (8)	[nouveautés]
080 Recueils généraux (0)	[nouveautés]
090 Manuscrits et livres rares (7)	[nouveautés]

L'utilisateur est alors dirigé vers la liste des 10 catégories secondaires qui lui est associé. Il choisit ensuite la catégorie secondaire en cliquant sur son intitulé pour voir s'afficher toutes les notices des bases de données indexées sous ce thème.

900 Géographie et Histoire (2)	[nouveautés]
910 Géographie générale. Voyages (3)	[nouveautés]
920 Biographie universelle et généalogie (1)	[nouveautés]
930 Histoire générale du monde ancien (4)	[nouveautés]
940 Histoire générale de l'Europe (0)	[nouveautés]
950 Histoire générale de l'Asie (7)	[nouveautés]
960 Histoire générale de l'Afrique (3)	[nouveautés]
970 Histoire générale de l'Amérique du Nord (11)	[nouveautés]
980 Histoire générale de l'Amérique du sud (0)	[nouveautés]
990 Histoire générale des autres régions du monde (0)	[nouveautés]

Ce mode de recherche étant très général, la possibilité est donnée d'affiner les résultats en cliquant sur les différents liens (textes ou icônes) des notices. Ces fonctionnalités sont répertoriées dans le modèle de notice.

Si le visiteur revient régulièrement et connaît déjà certaines références de DADI, il peut demander à n'afficher que les notices des bases de données référencées depuis 10 jours au plus en cliquant sur l'étiquette [nouveautés] associée à une catégorie secondaire de son choix.

2) On peut aussi lister toutes les bases de données référencées dans une des 10 catégories primaires de la classification Dewey. Il convient alors de cliquer sur l'étiquette [toutes] associée à une catégorie primaire. Ce mode de recherche étant très général, on peut affiner les résultats en cliquant sur les différents liens (textes ou icônes) de la notice ; nous en détaillerons le mode d'emploi par la suite.

Dans tous les cas, l'affichage des notices est précédé de la liste alphabétique interactive des noms de bases de données. En cliquant sur ces liens, le visiteur est directement dirigé vers la notice de son choix.

Il y a 1 résultat(s).

- American Indians of the Pacific Northwest Digital Collection

Version imprimable

précédente | suivante

1

Une version imprimable, c'est à dire du texte sans image, est présente à chaque page afin de faciliter, le cas échéant, le tirage et la lecture sur papier des notices.

répertoire des Bases de données grAtuites Disponibles sur Internet

<http://urfist.univ-lyon1.fr>

Affichage limité aux 100 premières notices

Généralités > Manuscrits et livres rares (Document numérisé)

Biblioteca digital Dioscórides

Cette bibliothèque numérique pour l'histoire de la médecine présente des notices bibliographiques de livre de médecine du XVIe au XIXe siècle appartenant aux collections de la bibliothèque de l'Universidad Complutense de Madrid. Pour le moment, en accès internet, la consultation intégrale des ouvrages n'est pas encore possible alors qu'elle l'est sur l'intranet des bibliothèques de médecine et de pharmacie de l'Université Complutense. Chaque notice est liée en général à une image la page de titre, rarement plus. Un seul livre intégralement numérisé est disponible : Le Compost et Kalendrier des Bergères, Paris, Guy Marchant, 1499, choisi pour son bon état de conservation et ses 70 gravures.

<http://www.ucm.es/BUCM/diosc/00.htm>

Universidad Complutense de Madrid
1 400 références
Espagnol Anglais

Référencée le 05/07/2002 - Dernière mise à jour le 12/07/2002

4.4.3. Une notice interactive

L'affichage d'un ensemble de résultats sous la forme d'une liste de notice ne constitue pas le point final d'une recherche ni la réponse unique à une requête.

En effet, chaque notice est interactive et offre donc la possibilité d'élargir ou d'affiner sa requête par un simple clic :

Version graphique :

The diagram shows a notice for a database titled '@nuaire des Professionnels du Textile'. It is categorized under 'Techniques (sciences appliquées) > Fabrication industrielle'. The notice includes a description: 'Consultez gratuitement l'annuaire on-line où sont référencées plus de 12000 entreprises textiles. L'annuaire est multilingue et les informations apportées sont mises à jour régulièrement. Vous recherchez un produit, un savoir-faire ou le partenaire commercial idéal, consultez l'Annuaire.' and a URL: 'http://www.b2btextile.ws/annuaire.php3?idlink=0&lang=2'. It also shows flags for French, Spanish, and English, and mentions '12000 entreprises référencées' and 'Référencée le 05/07/2002 - Dernière mise à jour le 05/07/2002'. Callouts explain various elements: 'Type de répertoire', 'Catégorie primaire', 'Catégorie secondaire', 'Retour à la première étape', 'Nom du/des producteur(s)', 'Langue(s) disponible(s)', 'Nombre de références', 'Dates de référencement', 'Description/résumé', 'Adresse du site internet', and a 'Un problème?' icon.

Partie 3 : Nouvelle Architecture informatique et pérennité du projet

1. Architecture informatique

La programmation de l'application a été confiée à Hugo Etiévant, étudiant en licence MIAGE de l'Université Lyon 1.

1.1. Le site visiteur

1.1.1. L'environnement technique

Objectif : création d'une application web dynamique reprenant le contenu de la base de données Access « Les bases de données gratuites sur Internet ».

Existant : un site statique en HTML généré par le programme GDIDb à partir de la base Access.

Pré-requis : Internet, HTML, HTTP/1.1, langage PHP, bases de données SQL, administration UNIX, design et retouche d'image.

Technologie utilisée : PHP4, MySQL et le serveur Apache sous Linux Mandrake 8.

Logiciels utilisés : kEdit, PhpEdit, Adobe PhotoShop 6.0 fr.

Sites testés avec les navigateurs : Internet Explorer 5.5 et Netscape 6.1 sous Microsoft Windows 98, 2000 et XP ; Mozilla 0.8 et Konqueror 2.1.1 sous Linux Mandrake 8.

1.1.2. Charte graphique et maquette

1.1.2.1. *Charte graphique*

Libre initiative nous a été donnée à Hugo Etiévant, étudiant à l'IUP de Lyon en licence informatique réseau (MIAGE) et moi-même pour la création de la charte graphique du site. Il m'a semblé intéressant, dans un souci d'exhaustivité de l'exercice, de participer à l'élaboration de cette charte. En effet, l'habillage du site et la présentation des interfaces constitue le premier élément qui apparaît au visiteur, avant même qu'il ne commence sa recherche. C'est, outre la fonctionnalité du site, ce qui lui donnera envie d'utiliser l'outil documentaire développé, d'y revenir et de le conseiller à d'autres utilisateurs. Il s'agit donc de ne pas ignorer l'importance de la charte graphique et en définitive son rôle d'ambassadeur auprès des usagers.

- La couleur dominante choisie est l'orange. Le menu, le logo, la bannière de titre, les notices et autres éléments reposent sur des graphismes originaux homogènes.
- Les images sont créées au format PSD sous Photoshop et converties en .GIF
- Les images des drapeaux (libres de droit) utilisés pour figurer les langues proviennent du site <http://www.blueflash.de/flaggen/topten2r1.php> (en langue allemande). Les images déjà téléchargées en local sur le serveur ont été placées dans un répertoire.



répertoire des bases de Données gratuites Disponibles sur Internet

- Accueil
- Recherche
 - navigation dans la classification Dewey
 - avancée
 - aléatoire
 - le top 10
- Contact
 - commentaires
 - proposer une base de données
- Aide
- Notice légale



 English version

Accueil

Introduction

Internet change petit à petit la donne pour la diffusion des bases de données bibliographiques et factuelles.

Les universités, laboratoires de recherche et même les offices de brevets proposent leurs bases de données en accès gratuit. Sachez que ces ressources sont mal signalées par les moteurs de recherche car l'information est diffusée sous forme de pages dynamiques pas toujours indexées par les robots. On en parle sous le nom de **web invisible**.

Bien que gratuites, les bases de données sont protégées (voir [Database protection activities](#)).

Dadi est un répertoire en évolution constante de bases de données en consultation gratuite ; certaines fournissent aussi le document primaire mais le plus souvent ce dernier est payant.

NB : les catalogues de bibliothèques ne sont pas signalés ici, on peut consulter le [Catalogue des bibliothèques francophones](#) de l'ENSSIB, [Gabriel](#) pour les bibliothèques nationales européennes ou [LibDex](#) au niveau mondial.

Le lien pointe en général directement sur l'interface web d'interrogation; quelques bases restent accessibles par Telnet.

Vous pouvez aussi consulter :

- [CompletePlanet](#)
- [Invisible web produit par IntelliSeek](#)
- [The Invisible Web Directory](#)
- [Those Dark Hiding Places : The Invisible Web Revealed](#)
- [A collection of special search engines](#)
- [infomine](#)

qui complètent ce travail.

NB : Ces services proposent beaucoup de bases de données et de ressources internet en anglais repérées automatiquement. Leur nombre est donc plus important que DADI qui ne pointe que vers des bases de données bibliographiques validées manuellement par l'équipe de DADI .



DAtabases DIrectory

- Home
- Search
 - browse the Dewey class
 - advanced search
 - random
 - the top 10
- Contact
 - comments
 - submit a database
- Help
- Rights notice



 Version française

Home

Introduction

Internet changes bit by bit the broadcasting of bibliographical and factual databases.

Universities, research laboratories and even offices of patents propose their databases in free access.

These resources are not always found by search engines because information is spread through dynamic html pages, so not always locatable. It is called the "invisible web".

But although free, these last ones are protected (see [Database protection activities](#)).

DADI is a directory of databases in constant evolution in free consultation ; some also supply free primary documents but mostly this last one is paying.

NB : library catalogs or OPACS are not referenced

here, one can consult the ENSSIB [directory for libraries of French-speaking countries](#), [Gabriel](#) for the European national libraries or [LibDex](#) at the world level.

Generally, the link points directly towards the interface of interrogation ; some databases remain accessible by Telnet.

You can also browse :

- [CompletePlanet](#)
- [Invisible web produit par IntelliSeek](#)
- [The Invisible Web Directory](#)
- [Those Dark Hiding Places : The Invisible Web Revealed](#)
- [A collection of special search engines](#)
- [infomine](#)

which complete my work.

NB : These services propose many databases in English from internet resources tracked down automatically so they are larger than DADI which points only towards bibliographical databases selected one by one by DADI staff.

1.1.2.2.

Maquette

Logo	Bannière
Menu avec lien vers le site de l'URFIST	<i>Contenu de la page</i>
Langue	

L'apparence des pages du site est permanente : un logo en haut à gauche, une bannière centrée en haut, un menu à gauche, un bouton pour changer de langue et le contenu de la page dans l'espace restant.

1.2. Structure du site

Le site est constitué de deux parties distinctes : la partie visiteur accessible librement par tous via le web et la partie administrateur sécurisée par mot de passe.

1.2.1. Partie visiteur

Elle est formée de deux fichiers :

- ❑ `index.php` dont le paramètre **page** détermine la partie du site à laquelle accéder et le paramètre **action** l'action à réaliser.
- ❑ `imprim.php` qui peut être appelé – dans certains cas – par le fichier précédent pour afficher la version imprimable (sans aucun graphisme) du site.

Elle permet aux internautes d'accéder au contenu du site.

1.2.2. Partie administration

Elle correspond au fichier `admin.php`. Elle permet à des utilisateur identifiés d'administrer le site et d'en changer le contenu.

1.2.2.1. Description de la base de données

- **Une notice** possède un identifiant numérique unique (Id), un nom (Name), une description (Description), une adresse pointant vers son site Internet (Url), un nombre de références écrit en français (Nb_refs) et en anglais (Nb_refs2), une date de référencement (Linked) et une date de dernière modification (Last_modified), une liste de mots clés français (Keywords) et anglais (Keywords2) et un nombre de clics¹² (Hits).

A chaque fois qu'un visiteur va sur le site en passant par la notice, il incrémente le compteur de clics.

A une notice sont associées une catégorie, un type de ressource, une ou plusieurs langues, un ou plusieurs producteurs via les tables de jointures.

- **Une langue** possède un identifiant numérique unique (Id), un nom français (Caption), un nom anglais (Caption2) et une image (Image).
- **Une catégorie** possède un identifiant numérique unique (Id), un numéro dans la classification Dewey (Num), un nom français (Caption), un nom anglais (Caption2) et le numéro Dewey de la catégorie parent (Parent_numx).
- **Un producteur** possède un identifiant numérique unique (Id) et un nom (Caption).
- **Un type de ressource** possède un identifiant numérique unique (Id), un nom français (Caption) et un nom anglais (Caption2).
- **Une traduction** possède un identifiant numérique unique (Id), un texte français (Fr) et un texte anglais (Uk). Cette table contient tous les textes statiques du site lui permettant d'être bilingue.

Visiteurs : cette table contient la date de début de chacune des sessions des visiteurs. Elle permettra plus tard de faire des statistiques de visite détaillées.

Authentification : cette table contient la liste des utilisateurs (User) et leur mot de passe (Pass) autorisés à accéder à la partie administrateur.

1.2.2.2. *Scripts généraux*

Fichier de configuration : config.php

Il contient quelques fonctions génériques : connexion à la base de données, extraction de la traduction adéquate d'un texte, incrémentation du compteur de visites, affichage normalisé d'un message d'information... ainsi que des paramètres tels que le nombre maximum de notices à afficher par page, etc.

tt3() : Permet la traduction instantanée d'un texte. Cette fonction de traduction retourne le deuxième paramètre si la langue en cours est l'anglais, sinon retourne le premier. Ce qui fait du français la langue par défaut.

mytt() : Permet la traduction instantanée depuis le contenu d'une des données. Retourne le texte correspondant à l'identifiant numérique passé en paramètre dans la langue en cours.

- Fichier d'entêtes HTTP : headers.php

Il envoie au client les entêtes relatives au cache, date d'expiration et date de dernière modification afin de forcer le rafraîchissement des pages.

- Fichier d'entêtes HTML : metas.php

Contient les balises META du document HTML, le titre et joint la feuille de style CSS.

- Fichier de redirection : redirect.php

Prend en paramètre l'id de la notice vers le site de laquelle le visiteur va être redirigé (après incrémentation du compteur de clic de ladite notice).

- Fichier de gabarit (modèle de document) : index.php

Accueil du site auprès des internautes. Ce fichier contient la mise en page et les graphismes permanents (logo, bannière, menu, choix de la langue) (voir Maquette). Il appelle la configuration, les entêtes http et HTML, inclut les classes d'objets,

¹² C'est-à-dire le nombre de fois qu'on a cliqué sur le lien qui envoie à l'adresse de la base de données

créé une session (nécessaire pour le décompte des visites et la gestion de la langue).

Chaque objet de la base de données est matérialisé par une classe :

Notice	Item.class.php
Catégorie	Category.class.php
Type de ressource	Directory_type.class.php
Langue	Language.class.php
Producteur	Productor.class.php

Selon le paramètre **page** il appelle la classe de l'objet concerné car une page est matérialisée par l'instance d'un objet. Et cet objet se comporte suivant la valeur du paramètre **action**.

Accueil	Home.class.php
Navigation	Navigation.class.php
Recherche Avancée	Search.class.php
Notice aléatoire	Random.class.php
Top 10	Top10.class.php
Contact	Contact.class.php
Aide	Help.class.php
Notice légale	Legale.class.php

- Statistiques : stats.class.php

Affiche des statistiques détaillées sur le contenu de la base.

1.2.2.3. *Scripts des pages*

- Version imprimable : imprim.php

Ce fichier constitue une alternative au modèle de document graphique `index.php` en n'affichant aucune image. Il appelle les pages à afficher en

version imprimable à la demande de l'internaute. Cette fonctionnalité n'est disponible que pour la page des résultats de la Navigation et ceux de la Recherche avancée.

- Accueil : `home.class.php`

Cette classe matérialise la page Accueil et se contente d'afficher du texte :

- ❑ Titre
- ❑ Texte de présentation du site
- ❑ Présentation de la classification Dewey
- ❑ Liste des sites pointant vers le nôtre
- ❑ Liste des technologies utilisées
- ❑ Quelques statistiques (nombre de visiteurs, de notices, de catégories)
- ❑ Liens vers quelques rubriques connexes du site de l'URFIST
- ❑ Liste des navigateurs testés

- Navigation : `navigation.class.php`

Cette classe permet de naviguer dans la hiérarchie des catégories Dewey.

printFirst()

Listage des catégories primaires (de numéro de parent nul).

printSub()

Si on a cliqué sur une catégorie primaire, alors on en affiche les catégories secondaires.

Si on a cliqué sur « toutes les notices », alors on n'affiche toutes les notices qui en découlent sans afficher de navigation pour les catégories secondaires.

printItems()

On arrive ici lors du clic sur une catégorie secondaire. On affiche la liste des notices correspondantes. En cliquant sur « nouvelles », on affiche que celles référencées il y a moins de `$NB_DAYS_RECENT` (10) jours.

L'affichage des résultats de la recherche se fait par pages de `$NB_NOTICES_BY_PAGE` (10) notices avec la fonction **printPage()** qui s'oriente

vers la version graphique ou imprimable. En fin de page de résultats, apparaissent des liens vers les pages précédentes, suivante et vers toutes les autres pages grâce à la fonction **navByPages()**. En début de page de résultats s'affiche un lien vers la version imprimable ainsi que la liste sur deux colonnes des noms des notices trouvées avec un lien interne à la page pour atteindre la notice, grâce à la fonction **printNames()**.

- Recherche avancée : `search.class.php`

printForm() : Affichage du formulaire avec extraction des catégories, types de répertoire, langues.

Le formulaire de recherche avancée est affiché en permanence en bas de page et contient les derniers critères entrés (pour permettre éventuellement d'affiner la recherche).

checkForm() : Normalise les valeurs de tous les champs en utilisant sur chacun d'eux la fonction **protect()** qui limite leur taille, enlève les slashes, les espaces inutiles et les caractères spéciaux de MySQL.

findItems() : Prépare les conditions de sélection de la requête en fonction des champs du formulaire, formule la requête et l'exécute. Puis affiche les résultats par pages de la même manière que pour la navigation.

Index des mots clés : `getkey.class.php`, **getkey.php**

Dans le formulaire de recherche avancée, on peut sélectionner un mot (ou une expression) clé dans un index qui s'affiche dans une fenêtre pop-up.

Index des producteurs : `getprod.class.php`, `getprod.php`

Dans le formulaire de recherche avancée, on peut sélectionner un producteur dans un index qui s'affiche dans une fenêtre pop-up (idem que précédemment).

Notice aléatoire : `random.class.php`

Affiche une notice au hasard.

Le top 10 : top10.class.php

Affichage des 10 notices des bases les plus populaires de DADI, c'est à dire celles qui ont enregistré le plus de clics sur le lien hypertexte menant au produit lui-même (et non simplement à la notice). Ce clic se fait soit sur le nom, soit sur l'URL et renvoie vers le script `redirect.php` qui incrémente le compteur et redirige le navigateur vers le site de la notice de façon transparente.

Contact : contact.class.php

Il permet l'affichage de trois formulaires :

- Un problème ? prend en paramètre l'id d'une notice et en rappelle le nom et l'URL. Permet de signaler les liens morts.
- Commentaires & questions permet la libre expression.
- Proposer une nouvelle base permet de suggérer l'ajout d'une nouvelle base.

Ces trois formulaires reposent sur la messagerie électronique du client. L'adresse du destinataire est spécifiée par `$CONTACT_EMAIL`.

Aide : help.class.php

Permet l'affichage d'un texte détaillé sur : la navigation, la recherche avancée, les modèles de notice.

Notice légale : legale.class.php

Affichage des renseignements suivants : organisme, directeur de la publication, conception du système documentaire et mise à jour, développement de l'application.

Affichage de bannières et vignettes aux formats suivants (selon les recommandations des standards IAB¹³ 2002) :

¹³ L'IAB France est « la première association de dimension internationale dédiée à l'optimisation de l'usage et de l'efficacité de la publicité sur le média Internet ». <http://www.iabfrance.com>

Bannières 468x60 pixels - GIF

16/9 320x180 pixels - GIF

Demi-Skyscraper 160x320 pixels - GIF

Skyscraper 160x600 pixels – GIF

Elles peuvent être utilisées librement pour signaler DADI.

Vignettes de diverses dimensions - GIF

Informations complémentaires : Email, URL, copyrights

1.2.2.4. *Scripts des objets*¹⁴

Catégorie : category.class.php

- addForm()

Formulaire d'ajout / modification. Si le paramètre id n'est pas nul, alors extraction de la catégorie et remplissage des champs avec leurs valeurs, sinon les champs sont vides.

Choix facultatif d'une catégorie parente parmi celles qui existent. Saisie d'un numéro Dewey, d'un nom français et d'un nom anglais.

- checkAddForm()

Vérification de la validité des valeurs saisies : la catégorie parente sélectionnée doit exister dans la base. Le numéro Dewey saisi ne doit pas déjà exister et être homogène avec celui de la catégorie parente si elle existe (le premier chiffre doit être le même, le troisième chiffre doit être zéro), ce numéro est composé de trois chiffres. Si les données sont valides, alors insertion / modification dans la base.

- showList()

Affichage hiérarchique trié de toutes les catégories de la base et affichage d'un lien vers le formulaire de modification.

- showMenu()

Affichage d'un petit menu : lister les catégories, ajouter une catégorie.

¹⁴ Toutes les fonctions ne sont pas ici présentées car les mêmes méthodes reviennent tout le temps : addForm() et modifyForm() pour ajout / modification, checkForm() pour contrôle des données saisies, printList() pour lister les objets d'un même type, confirmDelete() et delete() pour supprimer un objet...

Notice : item.class.php

- addForm()

Affichage du formulaire d'ajout / modification.

- checkAddForm()

Contrôle des données saisies et ajout / mise à jour dans la base avec appel des autres objets (langue, catégorie...) pour les jointures.

- printItem()

Appel de la fonction d'affichage de la notice **show()** avec les boutons « précédente », « suivante ».

- showList()

Affiche la liste des noms des notices par page de \$NB_ITEMS_BY_PAGE noms avec un lien vers le formulaire de modification, ou vers un formulaire de suppression, ou vers l'affichage de la notice. Permet de n'afficher que les notices commençant par une lettre de l'alphabet en particulier.

- show()

Affichage graphique d'une notice (visiteur ou administrateur).

Extraction de la notice à partir de son Id passé en paramètre. Création de liens sur certains des champs pour actionner une recherche avancée. Troncature de l'URL avec la fonction **truncateUrl()** afin de gérer le problème des adresses trop longues. Conversion des dates au format anglophone MySQL vers le format francophone avec la fonction **dateToFrench()**.

Langues : language.class.php

Ajout / modification / suppression d'une langue.

Ajout / suppression d'une jointure entre une notice et une langue.

- checkImage()

Permet le contrôle et l'ajout / modification d'une image (drapeau) associée à une langue.

Contrôles : existence du fichier joint, vérification de la bonne extension « .GIF », existence dans le répertoire dédié \$IMAGES_LANGUAGE_PATH d'un fichier du même nom. Si c'est une modification, suppression de l'ancien fichier. Copie du fichier dans le répertoire dédié. Retourne le nom du fichier qui sera inséré / MAJ dans la base.

2. Difficultés rencontrées

2.1. La question légale

Ce problème s'est posé lors de l'adoption de la classification Dewey. En effet, elle est protégée par un copyright géré par l'OCLC¹⁵ (Online Computer Library Center, Incorporated). Nous avons donc demandé l'autorisation de l'exploiter en présentant notre projet et elle nous a été accordée.

La notice légale du site comporte d'ailleurs une mention indiquant le copyright :

DADI - répertoire des bases de Données gratuites Disponibles sur Internet utilise
la classification décimale Dewey qu'il intègre dans son interface de recherche.
La classification décimale Dewey est (c) 1996-2001
OCLC

2.2. Les contraintes informatiques

Nous l'avons vu, la gestion de la base de données sous Access ne donnait plus satisfaction en raison de difficultés techniques et d'incompatibilités qui ont rendu obsolète le système documentaire élaboré.

Il a donc fallu trouver une solution informatique compatible avec nos exigences documentaires. Cet enjeu fut l'objet de nombreuses discussions et d'ajustements avec l'informaticien, Hugo Etiévant.

Mais cet aspect du travail a été très enrichissant en mettant en relief tous les aspects de la conception d'un système documentaire informatisé, indispensable complément aux cours théoriques dispensés pendant l'année de formation du DESSID.

¹⁵ <http://www.oclc.org>

La formation de trois jours au web dynamique (MySQL et php) que j'ai eu la chance de suivre dans les locaux de l'URFIST au début du mois de juillet m'a grandement sensibilisée au travail de l'informaticien qui a développé l'application.

3. Promotion et avenir

3.1. Promotion

Une fois le site terminé, la phase de promotion a pu commencer :

Hugo Etiévant, qui possède des talents d'infographe, a composé une affiche (format A3 et A4) informant de la nouvelle version du site :

DADI

Recherche par :

- navigation dans la hiérarchie des catégories
- formulaire de recherche avancée
- notice aléatoire
- top 10 des bases les plus populaires

répertoire des bases de données gratuites disponibles sur Internet

L'unique répertoire généraliste francophone à proposer un millier de notices sur les bases de données gratuites de l'Internet

Un site de l'Unité Régionale de Formation et de promotion pour l'Information Scientifique et Technique de Lyon

<http://urfist.univ-lyon1.fr/gratuits/>

Jean Pierre LARDY,
URFIST - UCBL

Conclusion

Le bilan de ce stage s'avère très positif sur différents plans.

D'un point de vue professionnel, le projet qui m'a été confié a été mené à bien.

Il est encore un peu tôt pour dire que le nouveau système de recherche correspond aux attentes des visiteurs du site, mais l'espoir est permis si l'on en juge par les réactions positives des personnes ayant testé l'application.

Cette expérience a été d'autant plus riche que Jean-Pierre Lardy, en m'accordant une autonomie totale mais vigilante, m'a donné l'occasion de faire et d'assumer mes premiers choix documentaires «grandeur nature ».

La collaboration étroite avec l'informaticien chargé de la programmation du système a été extrêmement enrichissante tant au point de vue technique qu'au point de vue méthodologique.

Sur le plan pratique, toutes les compétences acquises au cours de la formation du DESSID m'ont été fort utiles. Elles ont été approfondies par l'aide et les conseils précieux des personnes m'ayant entourée tout au long de ce stage.

Bibliographie

Ouvrages :

BIBLIOTHÈQUE NATIONALE DU CANADA, BIBLIOTHÈQUE NATIONALE DU QUÉBEC, BIBLIOTHÈQUE NATIONALE DE FRANCE. *21ème classification décimale Dewey.* Montréal, 1999 : Asted. Volume 4, index guide. 1276 p. ISBN : 2-921548-38-0.

FOENIX-RIOU Béatrice. *Recherche et veille sur le web visible et invisible.* Paris, 2001 : éditions tech et doc. 234 p. ISBN : 2-7430-0450-9.

LARDY Jean-Pierre. *Recherche d'information sur Internet : méthodes et outils.* Paris, 2001 : ADBS. 123 p. ISBN : 2-84365-052-6 – ISSN : 1159-7666.

Sites Internet :

- <http://www.sup.adc.education.fr/bib> : ce site de l'Education Nationale donne des renseignements sur les URFIST.
- <http://www.oclc.com/dewey/index.htm> : le site web de l'OCLC qui, entre autre, gère les droits de la classification Dewey.
- <http://urfist.univ-lyon1.fr/gratuits/index.html> : site de « Bases de données gratuites sur Internet », ancienne application.
- <http://bubl.ac.uk/link/> : Cet annuaire de ressources internet couvrant tous les sujets universitaires utilise la classification Dewey.
- <http://www.lii.org/> : index des ressources de l'Internet à l'usage des bibliothécaires.
- <http://www.phpdebutant.com/>

Moteurs de recherche : <http://www.yahoo.fr> - <http://www.google.fr> - <http://www.aol.fr> - <http://www.voila.fr> - <http://www.altavista.fr>

Métamoteurs : <http://www.ixquick.fr>, Copernic

Portails : <http://dmoz.org/> (Open Directory Project)

Table des annexes

ANNEXE 1 : LA CLASSIFICATION DEWEY.....	I
ANNEXE 1 : RÈGLES DE CATALOGAGE.....	III

Annexe 1 : La classification Dewey

<p>000 Généralités 010 Bibliographie 020 Bibliothèque économie et sciences de l'information 030 Encyclopédies générales 040 La vingtième édition n'attribue aucun sujet à ces indices 050 Publications en série d'ordre général 060 Organisations générales et muséologie 070 Journalisme, édition, journaux 080 Recueils généraux 090 Manuscrits et livres rares</p>	<p>100 Philosophie et psychologie 110 Métaphysique 120 Connaissance, cause, finalité, l'homme 130 Phénomènes paranormaux, parapsychologie 140 Les divers systèmes philosophiques 150 Psychologie 160 Logique 170 Éthique (Philosophie morale) 180 Philosophie ancienne, médiévale, orientale 190 Philosophie occidentale moderne</p>	<p>200 Religion 210 Religion naturelle 220 Bible 230 Théologie doctrinale chrétienne 240 Théologie morale et spirituelle chrétienne 250 Églises locales, congrégations religieuses 260 Théologie et société, ecclésiologie 270 Histoire et géographie de l'Église 280 Confessions et sectes chrétiennes 290 Autres religions et religions comparées</p>	<p>300 Sciences sociales 310 Statistiques 320 Science politique 330 Économie politique (Économique) 340 Droit 350 Administration publique 360 Services sociaux; associations 370 Éducation 380 Commerce, communications, transports 390 Coutumes et folklore</p>
<p>400 Langage 410 Linguistique 420 Langues anglaise et anglo-saxonne 430 Langues germaniques Allemand 440 Langues romanes. Français 450 Langues italienne, roumaine, rétho-romane 460 Langues espagnole et portugaise 470 Langues italiennes Latin 480 Langues helléniques. Grec classique 490 Autres langues et groupes de langues</p>	<p>500 Sciences de la nature et mathématiques 510 Mathématiques 520 Astronomie et sciences connexes 530 Physique 540 Chimie et sciences connexes 550 Sciences de la Terre 560 Paléontologie. Paléozoologie 570 Sciences de la vie 580 Botanique 590 Zoologie</p>	<p>600 Techniques (sciences appliquées) 610 Sciences médicales 620 Art de l'ingénieur, activités connexes 630 Agriculture et techniques connexes 640 Arts ménagers et sciences ménagères 650 Sciences de la gestion et techniques connexes 660 Génie chimique et techniques connexes 670 Fabrication industrielle 680 Fabrications diverses 690 Bâtiments</p>	<p>700 Les arts 710 Urbanisme et art du paysage 720 Architecture 730 Arts plastiques, Sculpture 740 Dessin, arts décoratifs et arts mineurs 750 La peinture et les peintures 760 Arts graphiques, Gravures 770 La photographie et les photographies 780 Musique 790 Loisirs spectacles, jeux, sports</p>
<p>800 Littérature et rhétorique 810 Littérature américaine 810 Littérature canadienne-anglaise 820 Littératures anglaise et anglo-saxonne 830 Littératures des langues germaniques 840 Littératures des langues romanes Littérature française 840 Littérature canadienne-française. Littérature québécoise 850 Littératures italienne, roumaine, rétho-romane 860 Littératures espagnole et portugaise 870 Littératures des langues italiennes Littérature latine 880 Littératures des langues helléniques Littérature grecque 890 Littératures des autres langues et groupes de langues</p>	<p>900 Géographie et histoire 910 Géographie générale. Voyages 920 Biographie universelle et généalogie 930 Histoire générale du monde ancien 940 Histoire générale de l'Europe 950 Histoire générale de l'Asie 960 Histoire générale de l'Afrique 970 Histoire générale de l'Amérique du Nord 980 Histoire générale de l'Amérique du Sud 990 Histoire générale des autres régions du monde</p>		

Annexe 2 : Règles de catalogage

Règles de catalogage pour l'administration de DADI

Renseignement d'une notice

(*) champ obligatoire

Catégorie (*): Liste de choix obligatoire. Choisir une catégorie primaire puis une catégorie secondaire. La consultation de l'index de la classification Dewey est indispensable.

Il faut choisir le thème le plus évident, les thèmes secondaires étant indexés dans le champ « mot-clés ». Par exemple : Afriphoto. Il s'agit d'une base de photographies sur les animaux d'Afrique.

- Catégorie : 590 zoologie (les animaux constituent le thème des photos donc de la base)
- Mot-clés : Afrique, animal et le détail des animaux si c'est précisé dans la description)
- Type de ressource : photothèque

Nom (*): il s'agit du nom original de la base de donnée, tel qu'il a été donné par le producteur. On ne traduira donc aucun intitulé.

- Le nom doit être écrit en minuscule et commencer par une majuscule :
 - **Il movie guide : film finder**
- sauf lorsqu'il s'agit d'un sigle :
 - **gri g ; nimal enome atabase** (sigle décomposé)
- ou que le nom de la base nécessite de conserver son orthographe originale :
 - **gri g ; nimal enome atabase** (sigle décomposé)
- **Pas de point en fin de ligne**

Description () : En minuscule. Se méfier des copier-coller à partir de pages web : souvent, les apostrophes ne sont pas correctement retranscrites et les espaces doublent après un point. Il faut étoffer la description autant que possible en donnant des détails sur le projet à l'origine de la mise en ligne de la base, le type de public à qui elle s'adresse, etc.

La mise en forme est reproduite telle quelle dans la notice qui apparaît au visiteur.

Mots-clés en français (séparateur = virgule) () : tous les noms communs doivent être écrits en minuscule **même en début de liste**. Seuls les noms propres comportent une majuscule en début de mot. Ils doivent être séparés par une simple virgule, avec ou sans espace. Pas de point à la fin de la liste.

Le nombre de mots-clés n'est pas limité mais un seul n'est pas suffisant. Il faut choisir les descripteurs d'une part dans la description que l'on a fait de la base de données auparavant et d'autre part en inventer pour affiner ou au contraire élargir la recherche.

Par exemple, pour une base de données sur les maladies de peau, on pensera à affiner (en reprenant les maladies énumérées dans la description : varicelle, eczéma...) et à élargir (dermatologie, peau, médecine...).

Mots-clés en anglais (séparateur = virgule) (') : idem. Il doit y avoir correspondance entre le nombre de mots anglais et français.

URL (°) : l'adresse doit toujours avoir la forme « http:// ». Pour réduire les erreurs fréquentes de transcription, on peut opter pour un copier-coller depuis la barre d'adresse de la page web. Le lien doit de préférence pointer vers le formulaire de recherche de la base de données sauf lorsque la lecture de consignes antérieures s'avère nécessaire.

Nombre de référence(s) en français : Le champ n'est pas obligatoire. Il faut écrire en toute lettre de quoi il s'agit (références, photographie, œuvres, références, adresses...). Le texte apparaît tel quel dans la notice donc il faut toujours le commencer par une majuscule. Les chiffres s'écrivent : **12 000 000** (avec espaces) ou en toute lettre **12 million**.

Nombre de référence(s) en anglais : Le champ n'est pas obligatoire. Les chiffres s'écrivent : **12,000,000** (avec virgules) ou en toute lettre **12 million**

Quelques traductions courantes :

accessible en ligne	Available online
base de données	Database
des centaines de	Hundreds of
des milliers de	thousands of
Dossier	File
Environ	About
Image	Picture
Œuvre	Work
Photographie	Photography(ies)
plus de	Over
Référence	Record
page	Page
10 000 références	10,000 records
plus de 10 000 références	over 10,000 records

Producteur(s) (séparateur = virgule) : ce champ n'est pas obligatoire car l'information n'est pas toujours disponible. Un index a été fait pour éviter les erreurs de retranscription. En règle générale, il faut vérifier dans cette liste si le producteur est déjà référencé et si c'est le cas, cliquer sur son nom afin de garder la même orthographe. C'est important pour la recherche avancée par producteur :

Par exemple : ministère de la culture, ministère de la culture et de la communication et ministère culture sont considérés comme trois producteurs différents. La nécessité d'harmoniser les orthographes est donc évidente.

Langue(s) (*) : Il s'agit des langues disponibles pour l'interrogation de la base de données. Leur nombre n'est pas limité.

Type de ressource (*) : Liste de choix obligatoire. En règle générale, tout ce qui n'est pas dans la liste appartient au type de ressource « références bibliographiques » à moins que l'ajout d'une catégorie ne s'avère nécessaire suivant le type de bases rencontrées.