

Systeme universitaire de documentation

Le portail documentaire du Sudoc

Un nouvel outil pour une utilisation à variables multiples

Ouvert au public depuis le 1^{er} mars 2005 à l'adresse <http://www.portail-sudoc.abes.fr/>, le portail documentaire du Sudoc, développé à l'aide de logiciels de la société Archimed, s'inscrit dans le cadre plus global du système d'information qu'est le Sudoc. Il vient compléter les outils déjà existants pour offrir à la communauté de l'enseignement supérieur et de la recherche de nouveaux services bibliographiques, dont l'accès direct au texte intégral des documents primaires (articles scientifiques, thèses, rapports, prépublications, etc.) et la recherche sur le dit texte intégral. En accès libre pour tous sur Internet, différents types de connexion sont possibles : l'utilisateur « anonyme » a accès à l'ensemble des ressources libres de droit ; l'utilisateur « anonyme pisté », par la reconnaissance de l'adresse IP de son poste, peut effectuer des recherches et avoir accès aux ressources payantes pour lesquelles son établissement est abonné, enfin, l'utilisateur « identifié » (pour l'instant l'ensemble des identifiés au sein du Sudoc) offre des services personnalisés du type « historique des requêtes ». Un accès « professionnel », bientôt ouvert à l'ensemble des correspondants de la documentation électronique des établissements, permettra la gestion des droits associés aux différentes ressources électroniques. Conçu comme outil de valorisation de la production scientifique universitaire, comme relais de diffusion des fonds patrimoniaux des bibliothèques universitaires mis en ligne, comme outil pour la gestion mutualisée des pages de « signets », le portail du Sudoc est donc également une interface d'accès aux ressources électroniques payantes ainsi qu'un outil de gestion pour l'administration de ces ressources.

Un accès simultané multibase, une recherche fédérée multicritère

Grâce à l'utilisation de différents protocoles, le portail du Sudoc donne accès de façon simultanée à plusieurs bases. Cible naturelle de cette interconnexion : les catalogues de bibliothèques mais aussi différentes bases externes dont le nombre, dans les mois à venir, va s'étoffer. Bases bibliographiques, bases de sommaires d'éditeurs commerciaux ou d'agrégateurs, bases de thèses en ligne, bases de fonds numérisés, etc. forment un ensemble sur lequel porte chacune des recherches.

Par ailleurs, principalement par le biais du système de « moissonnage » OAI-PMH, la récupération des métadonnées descriptives – voire des textes intégraux – d'importants corpus d'articles est effectuée auprès de différents producteurs de données (éditeurs scientifiques, archives ouvertes, archives institutionnelles, fonds numérisés d'établissement), ce qui constitue le socle des bases gérées en interne par les équipes de l'ABES. Pour ce qui concerne la recherche, le portail documentaire du Sudoc permet différents types de recherche – fédérée – sur l'ensemble des bases interconnectées. Dès la page d'accueil, l'utilisateur a la possibilité de lancer une « recherche globale simple »* qui cible l'ensemble des bases. Cette recherche globale peut être affinée grâce à une « recherche avancée » multicritère. Par ailleurs, les bases étant organisées selon différents types de ressources, les recherches peuvent être plus spécifiques. À cet effet, les bases interconnectées

sont organisées en cinq catégories – ou sous-portails : catalogues et bibliographies, production scientifique, thèses, fonds numérisés, sites de références. À tout moment, l'utilisateur peut affiner ou modifier le périmètre de ses recherches en sélectionnant les bases sur lesquelles il souhaite lancer une requête.

Vers une politique documentaire concertée

L'un des objectifs énoncés du portail du Sudoc en termes de services visait la complémentarité avec les portails locaux développés dans nombre d'établissements. À ce titre, afin de remplir au mieux cette vocation, depuis septembre 2004, un groupe de travail composé d'une dizaine de professionnels des bibliothèques – BIUM (bibliothèque interuniversitaire de médecine), BMNHN (bibliothèque du Muséum national d'histoire naturelle), SCD des universités Aix-Marseille-III, Lille-III, Paris-VIII, SCD des universités de Metz, de Nice, de Savoie, de Valenciennes, SICD (Toulouse) – se réunit régulièrement pour réfléchir de concert aux éléments de politique documentaire nécessaires à cette complémentarité. Certaines des orientations préconisées par le groupe de travail, présentées lors des dernières Journées réseaux (mars 2005) seront discutées et, le cas échéant, validées par un comité scientifique, dont la création a été décidée lors du dernier conseil d'administration de l'ABES (janvier 2005). Entre autres orientations, il est rappelé qu'en tant qu'outil national, le portail a bien entendu vocation à la valorisation des ressources des établissements, l'accès aux fonds numérisés constituant l'un des pôles de cette démarche. L'indexation et la recherche, à partir du portail du Sudoc, sur l'ensemble du fonds *Medic@* (fonds d'histoire de la médecine numérisé par la BIUM), sur les bases de *Persée* (portail d'accès aux revues numérisées en sciences sociales et humaines) ou prochainement sur les fonds *PôliB* et *NordNum* (fonds patrimoniaux numérisés issus de la région Nord) en sont des exemples. Le groupe de travail se prononce également pour une valorisation plus systématique (voire exhaustive) de la production scientifique universitaire francophone. À ce titre, l'ensemble des archives institutionnelles, des archives ouvertes provenant d'organismes de recherche, les différentes bases de thèses en ligne, etc. trouveraient une place circonstanciée au sein du portail du Sudoc. Pour l'heure, la collaboration avec le CCSD qui inclut la base *TEL* (Thèses en ligne) ainsi que l'archive Jean-Nicod (travaux des chercheurs en sciences de l'éducation) va dans ce sens. En outre, de façon à être au plus proche des besoins respectifs des établissements, l'attention devra être portée sur l'équilibre entre les différents domaines scientifiques. Outil attendu par les établissements documentaires, la base « sites de références » a fait l'objet d'une réunion spécifique. Pour permettre le suivi des liens URL (dont on connaît la stabilité relative) ou la recherche sur les sites signalés, une gestion mutualisée d'une base de signets commune à tous semble une solution adéquate. Pour ce faire, une collaboration, tant technique que fonctionnelle, est à l'étude. Conformément à la circulaire concernant les nouveaux circuits de validation des thèses, texte qui officialise le statut du document électronique « thèse », la base des thèses trouvera dans un avenir très proche son utilité spécifique. En tant qu'interface d'accès au texte intégral de l'ensemble des thèses françaises, cette base constituera un miroir d'une partie de la production scientifique des universités.

Au-delà d'un portail d'entrée : créer une passerelle d'accès ?

Autre fonction du portail du Sudoc : faciliter la gestion des périodiques électroniques. À cet effet, en collaboration avec le consortium COUPERIN, un module spécifique, l'APE (Accès aux périodiques électroniques), sera prochainement implémenté sur le portail du Sudoc. Il sera alors possible aux correspondants de la documentation électronique des établissements qui le souhaitent de gérer leurs bouquets d'abonnements via le portail et de répartir les droits d'accès aux sources pour lesquelles leur établissement est abonné. De plus, en tant qu'opérateur financier pour certains groupements d'achats, l'ABES semble l'interlocuteur tout désigné pour faciliter l'accès aux articles fournis par les éditeurs commerciaux. Différentes discussions (avec les éditeurs Swets, Elsevier, Springer notamment) sont en cours.

Dans cet esprit de collaboration, les portails locaux pourraient accéder à des bases de données en ligne à travers le portail du Sudoc, mutualisant ainsi les outils de connexion (connecteurs) développés par l'ABES. De même, il serait possible d'ouvrir tout ou partie des bases internes du portail du Sudoc aux établissements qui souhaitent enrichir leur propre outil de recherche fédérée. À suivre...

Christine Fleury
 fleury@abes.fr

CCSD
 Centre de communication scientifique directe - site créé par le CNRS
<http://www.ccsd.cnrs.fr>

COUPERIN
 Consortium universitaire de périodiques numériques
<http://www.couperin.org/>

* Cet encart de recherche est dès à présent téléchargeable sur les systèmes d'information des établissements qui le souhaitent.

Point sur les contenus du portail Sudoc en juin 2005

Deux sortes de bases sont branchées sur le portail.

Bases externes : accessibles via une passerelle Z39.50 ou via les protocoles SRU, SRW ou autres Webservices. Pour ces bases, l'ABES utilise les index mis à disposition par la source interrogée, ce qui a une incidence au moment de la recherche. En effet, lors d'une recherche multibase, les index doivent être communs.

Bases internes : constituées à partir de chargement ou de moissonnage OAI d'autres bases issues d'éditeurs scientifiques, d'archives ouvertes, d'archives institutionnelles, de fonds numérisés... Pour ces bases, l'ABES conçoit les index-mots, index-texte intégral, les filtres, les options spécifiques. Un accès direct au document primaire, dans son environnement d'origine, est ensuite proposé.

Bases externes				
Nom de la base	Type de ressource	Droits	Type de connexion	Volumétrie
Catalogue Sudoc	Catalogue de bibliothèque	libre	Z39.50	5 800 000 notices
BN Opale Plus	Catalogue de bibliothèque	libre	Z39.50	9 millions de notices
Library of Congress	Catalogue de bibliothèque	libre	Z39.50	12 millions de notices
PCI (Periodicals Contents Index)	Base de données de sommaires (accès aux métadonnées de chacun des articles)	selon abonnement / sur reconnaissance IP	Z39.50	Sommaire de centaines de revues (couverture : de 1770 à 1995)
PCI full text (Periodicals Contents Index)	Base de données de sommaires plus texte intégral des articles	selon abonnement / sur reconnaissance IP	Z39.50	Collection complète de 122 revues (couverture : de 1770 à 1995)
Science Direct	Métadonnées d'articles proposées par Elsevier	libre (index restreint) selon abonnement pour l'accès au texte intégral	connexion Webservice	Environ 1 800 titres
RDN (Resource Discovery Network)	Base de données de ressources électroniques (sites web)	libre	SRU	Environ 100 000 sites
XTCAT- Experimental Thesis Catalog	Catalogue international / accès aux thèses en ligne	libre	SRU	4 millions d'enregistrements
Bases internes				
Nom de la base	Bases d'origine	Droits	Spécificités	Volumétrie
Thèses	Thèses en ligne (TEL) du CCSD (Centre de communication scientifique directe)	libre	moissonnage OAI - recherche sur le texte intégral via le portail Sudoc	2 800 documents
Production scientifique	Archive Jean-Nicod	libre	moissonnage OAI	426 documents
	Scientific literature digital library (Citeseer)	libre	moissonnage	95 256 documents
	Biomed Central	libre	moissonnage OAI	16 000 documents
Persée	Portail de revues numérisées en sciences humaines et sociales	libre	moissonnage OAI	Collection complète de 7 revues SHS numérisées : 6 000 documents
Fonds numérisés	Gallica (fonds numérisés de la BNF)	libre	moissonnage OAI possible sur une partie du fonds Gallica	27 000 documents
	Medic@ (fonds numérisés par la BIUM d'ouvrages anciens de médecine)	libre	chargement	2 700 documents