



Valorisation de la production académique :

→ COLLECTE, CONSERVATION, DIFFUSION

ANNE-LAURENCE MENNESSIER
mennessier@abes.fr

SÉVERINE DABERNAT
dabernat@abes.fr

MARIANNE GILoux
giloux@abes.fr

ISABELLE MAUGER-PEREZ
mauger@abes.fr

Abes (Agence bibliographique de l'enseignement supérieur)

Conservateur des bibliothèques, Anne-Laurence Mennessier, en poste à l'Abes depuis 2005, est actuellement responsable des missions « Thèses » et « Pilotage national des opérations de rétroconversions ».

Bibliothécaire contractuelle à l'Abes, Séverine Dabernat est administratrice du Fichier central des thèses.

Ingénieur de recherche à l'Abes, Marianne Giloux est responsable fonctionnelle de l'application Star.

Conservateur des bibliothèques, Isabelle Mauger-Perez, en poste à l'Abes, est responsable fonctionnelle du projet de portail des thèses.

La mission « thèses » à l'Abes

Le ministère chargé de l'Enseignement supérieur mène depuis plusieurs décennies une politique volontariste en matière de thèses, qu'il s'agisse de leur signalement, de leur conservation ou de leur diffusion, de valoriser la recherche ou de favoriser l'insertion professionnelle des jeunes docteurs.

À la faveur de la révision générale des politiques publiques (RGPP), le MESR¹ s'est recentré sur sa mission de pilotage stratégique. Appelé à devenir « une administration de mission, stratégique, innovante, transparente² », ses activités opérationnelles ont été transférées à d'autres opérateurs, dont l'Abes, agence nationale. Parmi les activités transférées figure notamment le volet « thèses », dont l'Abes a officiellement la charge depuis septembre 2010.

Pourquoi les thèses ?

« Le vrai diplôme universel, c'est le doctorat. Faire une thèse, c'est faire preuve d'inventivité, de ténacité et de rigueur. C'est autre chose que de suivre des cours puis de valider des examens³. »

La thèse de doctorat⁴ est un document particulier, travail exigeant et

innovant de recherche et exercice académique donnant accès à un grade universitaire. Ces deux éléments (un document scientifique + un diplôme) forment un couple administrativement indissociable. On pourrait identifier au moins six éléments constitutifs d'une thèse, chacun de ces éléments soumettant la thèse à des textes, codes ou règlements distincts qui correspondent à autant de protections (de l'auteur par exemple) ou de garanties (de l'accès au document thèse). L'abondance des métadonnées relatives aux thèses en reflète bien la complexité (voir figure page 27).

En 2008, 10 700 thèses de doctorat ont été soutenues en France⁵. Les thèses conservent une importance prépondérante dans le paysage scientifique et au sein des établissements, universités ou grandes écoles, qui se positionnent comme des producteurs et diffuseurs d'informations scientifiques, dans un contexte d'intégration (PRES), d'autonomie des établissements (LRU) et de compétition internationale accrue (voir le classement de Shanghai⁶ et autres palmarès internationaux). Chacun se doit donc de valoriser sa recherche, son patrimoine scientifique et, singulièrement, ses thèses, grâce aux possibilités ouvertes

d'exercice, HDR (habilitation à diriger des recherches) et thèses sur travaux.

5. Source : DEPP (Direction de l'évaluation, de la prospective et de la performance) du ministère de l'Éducation nationale, « Repères et références statistiques sur les enseignements, la formation et la recherche », septembre 2010 : www.education.gouv.fr/pid316/reperes-et-references-statistiques.html

6. Version 2010 consultable ici : www.lemonde.fr/mmpub/edt/doc/20100812/1398457_5eeb_arwu2010.pdf

1. Pour le détail des sigles rencontrés dans ces articles, se reporter à l'encadré p. 28.

2. http://media.enseignementsup-recherche.gouv.fr/file/organigramme/29/0/MESR17-03-09_46290.pdf

3. Cédric Villani, médaille Fields 2010, au congrès de la Conférence des grandes écoles – 7 octobre 2010.

4. Nos propos se limitent aux thèses de doctorat – ne sont pas évoquées ici les thèses

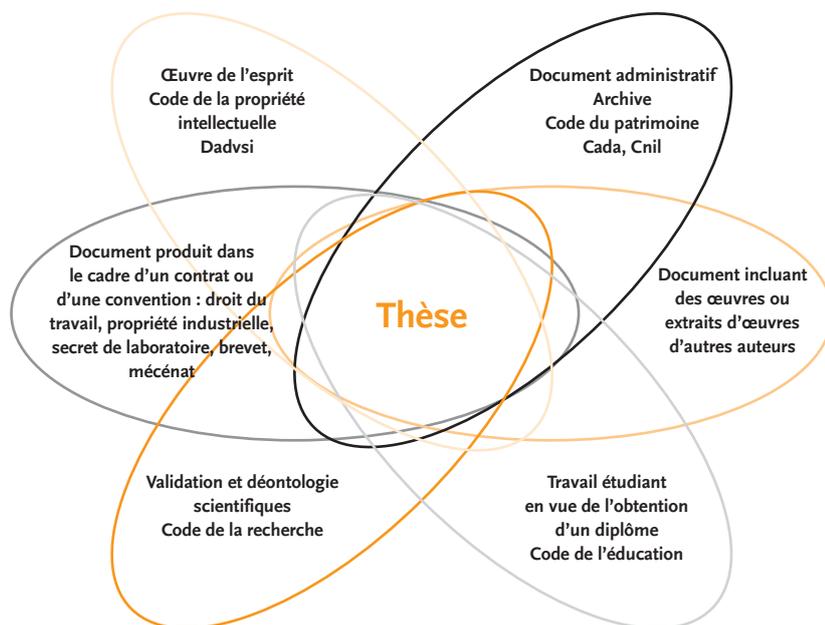


Figure illustrant l'abondance et la complexité des métadonnées relatives aux thèses (d'après un document du ministère de l'Enseignement supérieur et de la Recherche – Mission de l'information scientifique et des réseaux)

par les technologies actuelles. Celles-ci ne permettent pas seulement de dématérialiser, d'exposer et diffuser les documents, elles enrichissent aussi considérablement les fonctionnalités de recherche, de rebond ou de navigation, sans parler des nouveaux services possibles autour du document, de son auteur ou des différents intervenants (directeur, membres du jury, etc.).

Pourquoi l'Abes ?

La volonté de signaler les thèses soutenues est ancienne et continue, puisque le fichier qui recense les thèses soutenues existe depuis 1929 pour sa forme imprimée (répertoire bibliographique) et depuis 1985 sous forme informatique – accessible d'abord par télématique. En 1991, le ministère de l'Éducation nationale (et plus particulièrement la DPDU : Direction de la programmation et du développement universitaire) attribue une existence officielle⁷ au fichier

Téléthèses et en confie la gestion au centre du Catalogue collectif national. Ce fichier Téléthèses a pour objet de permettre le signalement des thèses soutenues en France :

- en lettres et sciences humaines et sociales : la saisie des bordereaux de thèses soutenues est confiée au CTIAG (Centre de traitement de l'informatique administrative et de gestion) ;
- en sciences : le traitement est effectué par l'Inist (Institut de l'information scientifique et technique) ;
- en médecine, pharmacie, odontologie et sciences vétérinaires : la saisie des bordereaux revient à la bibliothèque municipale et interuniversitaire de Clermont-Ferrand, la BMIU (actuelle BCIU, bibliothèque communautaire et interuniversitaire).

La base Téléthèses est chargée sur le Sunist (Serveur universitaire national pour l'information scientifique et technique), librement consultable par le grand public. À sa création en 1994⁸, l'Abes hérite d'une partie des

attributions du Sunist et se voit confier notamment le soin d'assurer la gestion et le développement de la banque de données Téléthèses, ainsi que l'édition sur tout type de support de produits dérivés. Elle produit donc Docthèses, version sur disque optique compact de Téléthèses, de 1992 à 2002. Depuis 2000, date de sa mise en production, le catalogue Sudoc prend le relais du disque compact, difficile à mettre à jour, et signale toutes les thèses françaises soutenues, de façon exhaustive depuis 1972 (depuis 1982 pour les thèses de médecine et de chirurgie dentaire). Pour les thèses plus anciennes, le catalogue est complété progressivement par des conversions rétrospectives de fichiers papier. Le Sudoc, géré par l'Abes, est donc la bibliographie nationale des thèses en France : le signalement des **thèses soutenues** y est en effet **obligatoire**⁹, qu'il soit effectué manuellement dans l'interface de catalogue Sudoc par la bibliothèque de l'université de soutenance dans le cas d'un dépôt papier, ou automatiquement via l'application Star¹⁰ dans le cas d'un dépôt électronique.

La préoccupation de l'Abes pour les thèses et sa responsabilité en ce domaine ne datent donc pas d'hier. Cet intérêt a même été renforcé récemment :

- par la forte implication de l'Abes dans la rédaction de la recommandation TEF¹¹ ;
- par la mise en production fin 2006 de l'application nationale Star ;
- par le transfert à l'Abes, le 1^{er} janvier 2010, du Fichier central des thèses¹² (FCT), auparavant hébergé et

9. Arrêté du 7 août 2006 relatif aux modalités de dépôt, de signalement, de reproduction, de diffusion et de conservation des thèses ou des travaux présentés en soutenance en vue du doctorat (NOR : MENS0602085A).

10. Voir page 30 l'article de Marianne Giloux, « Star : au-delà d'une application, un réseau ».

11. La recommandation TEF (Thèses Électroniques Françaises) définit un jeu de métadonnées pour les thèses électroniques soutenues en France. Son objectif est d'organiser de manière cohérente des métadonnées de thèses normalisées et riches, pour faciliter leur échange et leur diffusion, au niveau national ou international. www.abes.fr/abes/documents/tef/index.html

12. Voir page 29 l'article de Séverine Dabernat, « Le Fichier central des thèses ».

7. Arrêté du 10 septembre 1991 portant création d'un traitement automatisé des thèses soutenues en France (NOR : MENT9102183A).

8. Décret n° 94-921 du 24 octobre 1994 portant création de l'Agence bibliographique de l'enseignement supérieur (NOR : RESK9401199D).

Le circuit des thèses en sigles

Abes • Agence bibliographique de l'enseignement supérieur

www.abes.fr/abes/page,675,mission-theses.html

ANRT • Atelier national de reproduction des thèses

www.anrtheses.com.fr

CCSD • Centre pour la communication scientifique directe

www.ccsd.cnrs.fr

Cines • Centre informatique national de l'enseignement supérieur

www.cines.fr

DART-Europe E-theses Portal • Portail européen des thèses électroniques

www.dart-europe.eu/About/info.php?lan=fra

FRBR • Spécifications fonctionnelles de notices bibliographiques (Functional Requirements for Bibliographic Records) qui proposent une modélisation conceptuelle de l'information contenue dans les notices bibliographiques.

FCT • Fichier central des thèses

www.fct.abes.fr/FCT-APP/index.jsp

MESR • Ministère de l'Enseignement supérieur et de la Recherche

http://media.enseignementsup-recherche.gouv.fr/file/organigramme/29/0/MESR17-03-09_46290.pdf

MISTRD • Mission de l'information scientifique et technique et du réseau documentaire du ministère de l'Enseignement supérieur et de la Recherche. Elle est commune aux deux directions : Direction générale pour la recherche et l'innovation (DGRI) et Direction générale pour l'enseignement supérieur et l'insertion professionnelle (DGESIP).

ORI-OAI • Outil de référencement et d'indexation *open source* pour un réseau de portails OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) destiné à faciliter le partage, l'accessibilité et la diffusion des ressources numériques entre établissements.

www.ori-oai.org

TEF • Thèses électroniques françaises

Format TEF : jeu de métadonnées défini par la recommandation TEF, qui propose à la fois une modélisation des métadonnées et un format d'échange XML. Ce schéma XML, permettant de spécifier la structure et le contenu d'une notice TEF, a pour vocation d'échanger, sous une forme normalisée et validée, les métadonnées des thèses électroniques françaises.

www.abes.fr/abes/page,403,tef.html

TEL • Serveur de thèses en ligne du CCSD. TEL est un environnement particulier de l'archive ouverte pluridisciplinaire HAL (Hyper Articles en ligne) destinée au dépôt et à la diffusion d'articles scientifiques de niveau recherche, qui est dédié aux thèses de doctorat et HDR.

<http://tel.archives-ouvertes.fr>

Sudoc • Système universitaire de documentation

www.sudoc.abes.fr

Thesa • Base de données répertoriant les sujets de thèses et les PhD en cours dans les écoles d'ingénieur et de management membres de la Conférence des grandes écoles (CGE), ainsi que ceux des thèses soutenues depuis moins d'un an.

<http://thesa.inist.fr>

géré par l'université Paris 10 Nanterre, qui signale les **sujets de thèses en cours** dans les disciplines des lettres et sciences humaines et sociales.

Si la possibilité de localisation (voire consultation pour les thèses numériques) des « documents thèses » donnée par le catalogue Sudoc est essentielle, la connaissance des sujets de thèses en cours n'en est pas moins importante pour la communauté universitaire, voire pour les acteurs économiques¹³.

L'Abes bénéficie donc d'une vision globale – et, plus important, va pouvoir en faire profiter les usagers – du circuit de la thèse, depuis le dépôt du sujet – bien en amont de la soutenance donc – jusqu'à « l'après-thèse » : c'est tout l'enjeu du futur portail des thèses¹⁴, dont la réalisation, confiée à l'Abes, devrait voir le jour en 2011.

Pour quoi faire ?

Lors des réunions de travail préparant le transfert du « volet thèses » à l'Abes, les activités relatives aux thèses ont été listées et classées en trois catégories :

- *les activités transférées intégralement à l'Abes*. Y figurent notamment la gestion et la modernisation du Fichier central des thèses, la formation professionnelle et celle des usagers, la normalisation et les standards documentaires relatifs aux thèses (dans le cadre du transfert de l'ensemble de la mission « normalisation » du MESR à l'Abes) ;

- *les activités ou dossiers délégués à l'Abes en collaboration avec le MESR*. Se trouvent dans cette catégorie la mise en place d'une politique de valorisation et de diffusion des thèses électroniques, la mise en place d'un « portail des thèses » (si la partie technique et

fonctionnelle est assurée par l'Abes, un comité décisionnel et opérationnel en fixe les orientations), le conseil lors de la mise en place ou de la réorganisation de services, la réalisation, la mise à jour et l'hébergement de divers documents de référence (formulaire de dépôt de la thèse, guide du docteur...);

- *les dossiers qui restent du ressort du MESR*. Il s'agit de dossiers qui relèvent du pilotage politique : habilitations des établissements pour la délivrance du diplôme de doctorat, tutelle des partenaires impliqués dans la diffusion ou valorisation des thèses (ANRT, presses universitaires), questions juridiques (relatives à la propriété intellectuelle en particulier).

La « mission thèses » dévolue à l'Abes depuis septembre 2010 n'est donc pas liée à une application en particulier (Star, Sudoc, FCT, futur portail des thèses...) mais s'exerce à un niveau stratégique. Il s'agit en effet d'avoir une vision d'ensemble afin de :

- mettre en œuvre la politique générale décidée avec le MESR (en parti-

13. Les grandes écoles ont d'ailleurs instauré elles aussi un « fichier des sujets de thèses » en cours en leur sein, Thesa.

14. Voir page 32 l'article d'Isabelle Mauger-

Perez, « Le projet de portail des thèses : un moteur de recherche dédié à la production doctorale ».

culier avec la Mission de l'information scientifique et technique et du réseau documentaire – MISTRD) et avec les partenaires sur la gestion des thèses. Il convient pour cela de voir, au-delà des applications (pas seulement celles créées ou gérées par l'Abes), comment elles s'articulent ou pourraient s'articuler entre elles, comment elles sont complémentaires plutôt que concurrentes. Il s'agit également de voir au-delà des spécificités locales qui sont bien réelles et s'expliquent par la taille des établissements (de 10 thèses par an à plus de 700 – la complexité du circuit n'est forcément pas la même), les disciplines, le support matériel de la thèse, le mode de dépôt du document, etc. ;

- travailler avec les partenaires extérieurs (déjà identifiés ou à venir) : acteurs de la formation doctorale,

acteurs documentaires (des établissements et acteurs nationaux – Cines, CCSD), partenaires associatifs et économiques du domaine public ou privé, fournisseurs d'applications (Apogée, ORI-OAI, etc.) elles aussi impliquées dans le circuit des thèses ;

- assurer la liaison avec le Ministère pour les fonctions qui n'ont pas (ou pas intégralement) été transférées.

Ce transfert de missions entre le MESR et l'Abes ne bouleverse pas le paysage des thèses ni le dispositif en place. En confortant l'Abes dans son rôle et dans son expertise sur les thèses, il reste fidèle à une préconisation du groupe de travail dont le rapport a, entre autres, inspiré les arrêtés du 7 août 2006 :

« Il revient à l'État ou à un opérateur national de fixer le cadre général, d'éta-

blir les prescriptions indispensables à la cohérence du dispositif, de fournir les outils techniques nécessaires et de jouer un rôle d'assistance et de conseil ; il revient en revanche aux établissements de soutenir d'assurer le signalement, la valorisation et la diffusion de leurs thèses¹⁵. » C'est simplement le curseur qui a bougé, au profit de l'opérateur national qu'est l'Abes. ●

Anne-Laurence Mennessier
Novembre 2010

15. Rapport sur la diffusion électronique des thèses établi par un groupe de travail [rapport rédigé par Claude Jolly], 2000, p. 13 : www.cndwebzine.hcp.ma/IMG/pdf/La_Diffusion_electronique_des_theses.pdf

Le Fichier central des thèses

Des origines à l'Abes

Depuis 1968, le Fichier central des thèses, le FCT, signale les sujets de thèses en cours de préparation ou soutenues depuis moins d'un an dans plus de quatre-vingt-dix universités et établissements d'enseignement supérieur et de recherche français partenaires. Il est compétent dans les disciplines suivantes : lettres, sciences humaines et sociales, théologie, droit, sciences politiques, sciences économiques et sciences de gestion.

Le signalement des sujets dans une base de données accessible au public sur internet permet aux futurs doctorants, directeurs de thèses, et plus largement à l'ensemble de la communauté scientifique, d'avoir un aperçu de l'état de la recherche doctorale en cours et de pouvoir faire le point sur un thème de recherche ou une discipline.

En 2009, la mission de gestion du dispositif de signalement et valorisation des thèses de doctorat soutenues dans les établissements français d'enseignement supérieur a été confiée par le ministère à l'Agence biblio-

graphique de l'enseignement supérieur, l'Abes. C'est dans ce cadre que l'Agence a récupéré, le 1^{er} janvier 2010, la gestion du FCT, assurée jusque-là par l'université Paris-Ouest Nanterre La Défense. Le transfert d'application a été totalement transparent pour les utilisateurs du Fichier central des thèses et l'interface de la base de données a été conservée à l'identique.

État des lieux : aujourd'hui et demain

Aujourd'hui, le Fichier central des thèses¹ est alimenté par 160 personnes qui travaillent au sein des services de scolarité et des écoles doctorales des établissements partenaires en saisissant et en mettant à jour les informations relatives aux sujets de thèses des doctorants inscrits dans leur établissement.

Dans le cadre du portail des thèses, l'interface actuelle du Fichier central des thèses est appelée à disparaître et les données vont être intégrées à la future application.

Afin de préparer la migration des données vers le portail des thèses, un

travail de nettoyage et de mise à jour du FCT est effectué actuellement à l'Abes. En effet, sur 109 345 sujets de thèses déclarés comme en « cours » dans le Fichier central des thèses, 33 158 le sont depuis plus de dix ans. Les membres du comité de pilotage du projet de portail des thèses ont décidé que ces données datant de plus de dix ans ne seront pas migrées dans la nouvelle application.

Une fois versé dans le portail des thèses, le signalement des sujets ne sera plus limité aux disciplines visées initialement par le FCT : l'ambition est de s'ouvrir à toutes les universités. En ce qui concerne les grandes écoles, un partenariat a été noué avec Thesa², la base de signalement des thèses en cours de préparation dans ces établissements. Ainsi, la brique « thèses en cours » du portail des thèses donnera une visibilité à l'ensemble des domaines de la recherche doctorale française. ●

Séverine Dabernat
Novembre 2010

2. <http://thesa.inist.fr>

Thesa se développe dans le cadre d'une convention établie entre la Conférence des grandes écoles (CGE) et l'Institut national de l'information scientifique et technique du CNRS, l'Inist.

1. www.fct.abes.fr/FCT-APP/index.jsp

Star : au-delà d'une application, un réseau

Depuis la publication de l'arrêté du 7 août 2006 relatif aux modalités de dépôt, de signalement, de reproduction, de diffusion et de conservation des thèses ou des travaux présentés en soutenance en vue du doctorat¹, chaque établissement habilité à délivrer des diplômes de doctorat doit faire le choix entre opter pour le dépôt des thèses sous leur forme papier ou sous leur forme électronique. Les établissements ayant fait le choix du dépôt électronique se voient offrir le service Star (application nationale pour le *signalement des thèses, archivage et recherche* développée par l'Abes) leur permettant de procéder au signalement et à l'archivage pérenne de ces documents.

L'application a été ouverte en 2006 et, depuis, 96 établissements ont choisi d'être déployés dans le réseau Star et 37 sont effectivement en production et ont choisi officiellement le mode de dépôt électronique pour leurs thèses.

Un circuit à mettre en place

Le passage au dépôt électronique a souvent été l'occasion pour les établissements de revoir l'ensemble du circuit de la thèse en impliquant les différents acteurs concernés dans la réalisation du processus de signalement et d'archivage de la thèse électronique.

Par exemple, ont été impliqués à l'École centrale Paris² le service des affaires juridiques (convention de diffusion sur internet, droits d'auteur...), la direction des systèmes d'information (modes de diffusion, imports partiels...), l'école doctorale (nouveau circuit de la thèse) et la bibliothèque,

sans oublier la communication envers les doctorants eux-mêmes.

L'alimentation

Les établissements utilisent l'application pour venir y déposer les fichiers de la thèse (éditions d'archivage et de diffusion) et saisir les métadonnées associées, qu'il s'agisse des métadonnées descriptives ou des métadonnées de gestion (administration, archivage, droits). Chaque établissement possède un espace de travail collaboratif sur le web, dans lequel interviennent ses différentes composantes (doctorant, école doctorale, scolarité, SCD...), pour y effectuer les tâches pour lesquelles elles sont habilitées.

Les établissements ont également la possibilité d'importer les données déjà gérées dans leur propre environnement (système d'information, GED, ORI-OAI³...) afin d'éviter les saisies multiples de métadonnées. C'est le format TEF⁴ qui est utilisé pour cet échange de données, ce schéma XML ayant précisément pour vocation d'échanger, sous une forme normalisée et validée, les métadonnées des thèses électroniques françaises.

Les sorties

À la fin du processus d'alimentation, Star se charge, d'une part d'exporter les fichiers d'archives et certaines métadonnées vers la plateforme d'archivage pérenne du Cines⁵, d'autre part de convertir les métadonnées TEF en Unimarc pour alimenter en masse le catalogue Sudoc⁶ (voir figure 1 ci-contre).

Star alimente automatiquement un réservoir OAI-PMH⁷, permettant à l'établissement de soutenance de récupérer ces métadonnées et/ou les fichiers déposés, dans le format sou-

haité (TEF, mais aussi DC⁸, Unimarc en XML) afin de leur permettre de répondre à l'obligation légale qui leur est faite de communiquer la thèse en leur sein (physiquement ou virtuellement). Ce réservoir est également moissonné régulièrement par DART⁹ (Portail européen des thèses électroniques).

Enfin, Star propose à chaque établissement des options supplémentaires : export de la version de diffusion de la thèse vers le serveur de diffusion de l'Abes et/ou l'archive ouverte du CCSD, Thèses en ligne (TEL)¹⁰.

La diffusion des thèses

Dans l'application Star, les établissements doivent définir la politique de diffusion des thèses soutenues dans leur établissement, qu'il s'agisse d'une diffusion intranet et/ou internet, sur leur propre serveur de diffusion ou sur un serveur national (CCSD, Abes). Ensuite, thèse par thèse, il est toujours possible de modifier ces droits en fonction des autorisations de l'auteur. Il est également possible de définir une période de confidentialité édictée par le chef d'établissement ou une période d'embargo demandée par le docteur (voir figure 2 ci-contre).

L'année 2010

2010 est une année charnière pour le réseau Star, une nouvelle version de l'application ayant été déployée en octobre. Entièrement développée par l'Abes, cette version permet de répondre en grande partie aux lacunes de la première version : amélioration de l'ergonomie des formulaires de saisie, exemplarisation automatique des thèses dans le Sudoc, amélioration de l'interface de saisie des liens aux autorités Sudoc avec création possible

1. www.legifrance.gouv.fr/WAspad/UnTexteDeJorf?numjo=MENS0602085A%20

2. Retour d'expérience par deux établissements du PRES Universud Paris – Jean-Marie Barbiche, École centrale Paris, et Martine Coppet, université d'Évry : www.abes.fr/abes/page,564,journee-de-lancement-star.html

3. www.ori-oai.org

4. www.abes.fr/abes/page,403,tef.html

5. www.cines.fr/spip.php?rubrique219

6. www.sudoc.abes.fr

7. Open Archives Initiative Protocol for Metadata Harvesting : www.openarchives.org/OAI/openarchivesprotocol.html

8. Dublin Core : <http://dublincore.org>

9. DART-Europe E-theses Portal : www.dart-europe.eu

10. CCSD : Centre pour la communication scientifique directe.

Thèses en ligne : <http://tel.archives-ouvertes.fr>

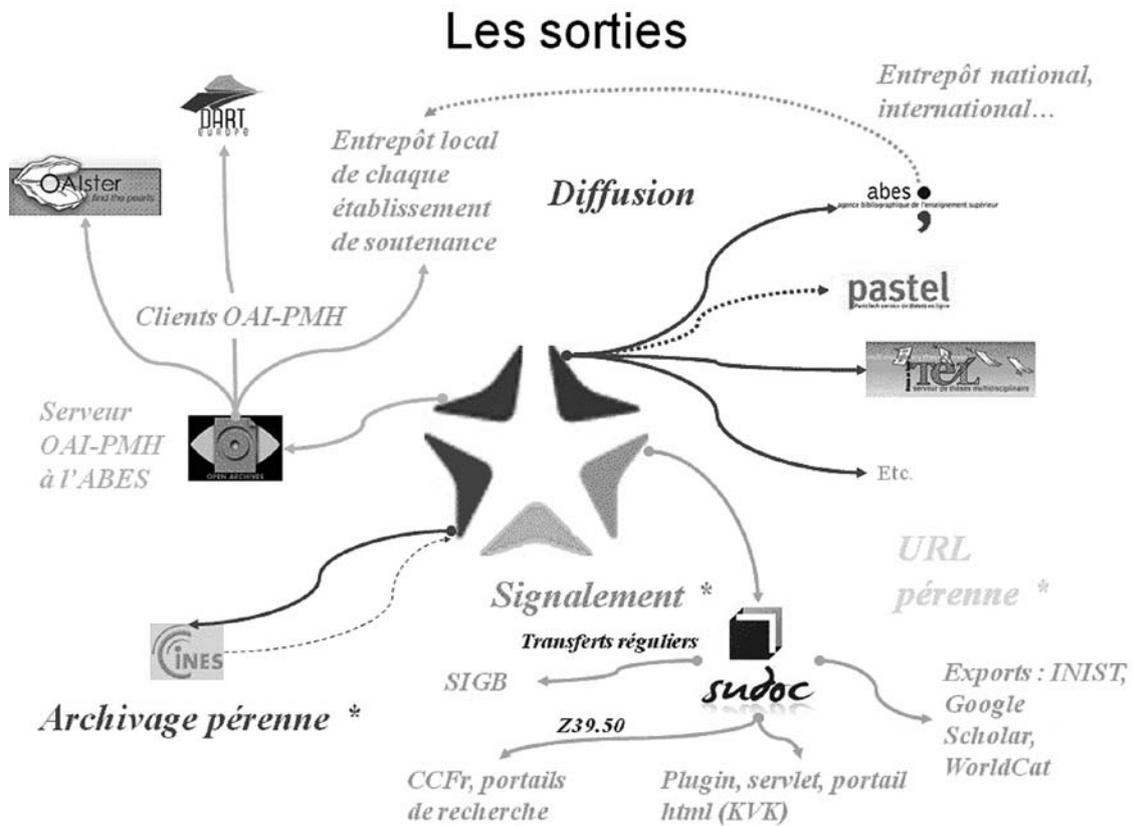
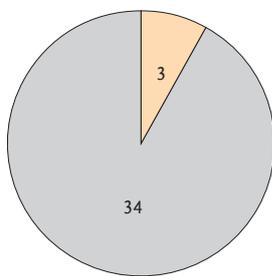


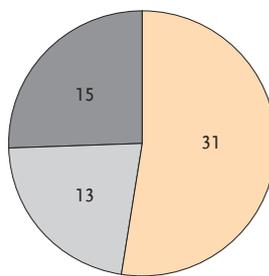
Figure 1
Schéma des sorties de Star

Choix du type de diffusion par défaut des établissements dans Star



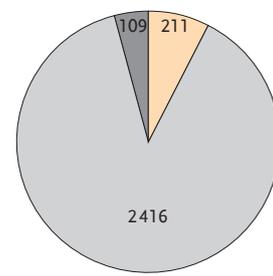
La diffusion internet
La diffusion intranet

Choix du serveur de diffusion par défaut des établissements dans Star



Établissement de soutenance
CCSD
Abes

Choix des options de diffusion, thèse par thèse, dans Star



Accessibles en intranet
Accessibles sur internet
Avec une restriction temporelle

Figure 2
La répartition actuelle des thèses déposées dans Star sur les différents sites de diffusion (certains établissements choisissent de diffuser sur plusieurs sites)

via la nouvelle application IdRef ¹¹ mise à disposition par l'Abes.

En effet, l'Abes s'est lancée dans un nouveau projet afin d'encourager la réutilisation des référentiels d'autorité du Sudoc par les applications documentaires relevant de l'enseignement supérieur et de la recherche : catalogues de bibliothèques, archives ouvertes et institutionnelles, plateformes pédagogiques, bibliothèques numériques, applications de gestion de la recherche, etc.

Autre nouveauté de la version, la mise en place d'une URL pérenne tel que prévu dans l'arrêté du 7 août 2006. La pérennité de l'accès à la thèse est ainsi garantie par l'attri-

bution d'une URL, à la fois adresse fiable et label d'authenticité. Ces URL pérennes sont un élément essentiel du projet de portail des thèses : la thèse pourra être citée, de façon pérenne, dans n'importe quel environnement (article, notice descriptive).

Un nouveau réseau

Certains établissements comme l'École centrale Paris et l'université d'Évry ont tiré un bilan assez positif de la mise en place de ce nouveau circuit : meilleure visibilité et diffusion des thèses, simplicité du *workflow* (peu d'intervenants), moins de « paperasse », ou encore appropriation rapide de l'application. Reste encore à

améliorer certains points non négligeables : meilleure communication vers les doctorants, feuille de style encore trop peu utilisée, méconnaissance du droit d'auteur, etc.

La mise en place d'un réseau comme Star permet également aux établissements de partager et d'échanger leurs expériences dans ce domaine. Certains s'associent dorénavant pour mettre en place des formations communes destinées aux doctorants et souhaitent également concevoir et partager leurs feuilles de style ou leurs chartes de diffusion des thèses. ●

Marianne Giloux
Novembre 2010

11. www.idref.fr

Le projet de portail des thèses : un moteur de recherche dédié à la production doctorale

Comment avoir une idée précise de la production doctorale en France ? De plus en plus d'informations sont disponibles librement sur internet, mais avoir une vue globale de la question reste souvent très difficile.

Où chercher une thèse ?

Vous cherchez une thèse en particulier ? Le Sudoc, parce qu'il inclut la bibliographie nationale des thèses, signalement exhaustif de toutes les thèses de doctorat soutenues en France depuis 1985, vous permet d'obtenir des références bibliographiques, de consulter sur place, de faire une demande de prêt entre bibliothèques ou parfois de consulter la thèse recherchée en ligne.

Vous cherchez à consulter le texte intégral d'une thèse électronique ? Avec un peu de chance, le docteur a lui-même déposé son document dans le réservoir d'archives ouvertes dédié aux thèses, le serveur thèses en ligne

(TEL) géré par le CCSD. L'établissement de soutenance peut aussi piloter une archive institutionnelle : après accord du docteur et des ayants droit, la thèse a été mise en ligne dans sa version de soutenance sous la responsabilité de l'établissement.

Vous souhaitez acheter une thèse ? Cette thèse a peut-être fait l'objet d'un travail d'édition et figure au catalogue des presses universitaires ou d'autres éditeurs commerciaux. Le docteur peut également avoir donné son accord pour que son document figure au catalogue des « Thèses à la carte » de l'Atelier national de reproduction des thèses (ANRT) de Lille, permettant ainsi l'édition à la demande de son document.

En résumé, bien qu'il n'existe pas de point d'entrée unique sur internet, il est de plus en plus facile d'obtenir des informations sur le « document-thèse » et même d'accéder à son contenu.

En amont et en aval de la thèse : un vrai jeu de piste

Si, par contre, vous vous intéressez à la thèse en tant que formation professionnalisante, vos recherches

seront plus laborieuses. En amont de l'obtention du diplôme, savoir qui encadre qui aujourd'hui reste difficile : les bases recensant les thèses en cours (Fichier central des thèses et Thésa notamment) ne couvrent pas tous les champs disciplinaires. En aval de l'obtention du doctorat, des services de mise en relation des docteurs et des employeurs potentiels existent, proposés notamment par l'Association Bernard Gregory, mais n'ont pas la visibilité qu'ils méritent.

Le portail des thèses : l'ambition d'un guichet unique

Le projet de portail des thèses est né de ce constat dressé en 2008 dans deux rapports distincts : le rapport du comité Information scientifique et technique présidé par Jean Salençon¹ et le rapport du groupe de travail « Adéquation public-privé » présidé par Germain Sanz dans le cadre de

1. Travaux du comité du 19 décembre 2007 au 18 mai 2008, remis à la DGR et à la DGES, www.enseignementsup-recherche.gouv.fr/cid21677/rapport-du-comite-ist-information-scientifique-et-technique.html

FutuRIS². Fin 2008, la MISTRD a confié à l'Abes la mission de proposer un guichet unique pour l'ensemble des thèses.

Interface de recherche destinée à tous les publics, le site donnera accès à des métadonnées relatives à la fois aux thèses et aux personnes et organismes impliqués dans le cycle de vie des thèses. Le périmètre du portail des thèses a été arrêté par le comité de pilotage du projet³ :

- les thèses de doctorat (les thèses d'exercices sont exclues du périmètre du portail) ;
- en cours et soutenues ;
- quel que soit le support exigé pour la soutenance (papier ou électronique) ;
- dans leur version de soutenance et dans leurs versions dérivées.

Les données du Sudoc relatives aux habilitations à diriger les recherches seront également accessibles dans le portail des thèses.

Le modèle de données du portail des thèses est basé sur une extension de la recommandation Thèses électroniques françaises (TEF) et, par voie de conséquence, sur le modèle FRBR. Le but du portail des thèses est d'exploiter les métadonnées de l'œuvre thèse, puis de présenter à l'internaute les différentes expressions (par exemple, version complète/version partielle), et les différentes manifestations (papier, électronique mis en ligne sur TEL, électronique mis en ligne par l'Abes, électronique accessible dans une archive institutionnelle, microfiche, édition commerciale...). Cet

effort de modélisation rendra plus aisée la recherche : les métadonnées qui permettent l'identification de la thèse comme diplôme (nom de l'auteur et des intervenants, organisme de délivrance, date de soutenance, discipline...) et comme production scientifique (indexation, résumés...) seront mises en avant et permettront de regrouper sous une même notice chapeau tous les supports matériels du document éventuellement disponibles.

L'interface publique permettra d'exploiter la richesse de la recommandation TEF : recherche par membres du jury, par école doctorale, par établissement de cotutelle, etc. L'internaute pourra se laisser guider par les rebonds par facettes et accéder aux informations par des nuages de mots clés. Lorsque le texte intégral des thèses sera accessible, il sera bien entendu indexé.

Le portail comportera également des services pour exploiter les données sans passer par l'interface publique : fils RSS, alertes mail, web services, intégration au web de données (web sémantique).

Les informations concernant chaque thèse mais aussi chaque personne ou chaque organisme lié à une thèse seront présentées dans une page web dotée d'une URL *ad hoc*. Le référencement par les moteurs de recherche classiques gagnera ainsi en efficacité.

Calendrier

Une première version du portail des thèses verra le jour à la fin du premier semestre 2011 avec les données provenant de l'application Star (thèses électroniques des établissements de soutenance ayant abandonné le papier comme support de la version de soutenance).

Au cours du second semestre 2011, une interface professionnelle du portail destinée à la saisie des thèses en cours sera mise en production. Les données du Fichier central des thèses seront versées dans le portail. Les recherches sur les thèses en cours et les thèses soutenues seront alors possibles.

Les chantiers suivants concerneront le versement de la bibliographie nationale des thèses du Sudoc (plus de 500 000 notices) et le signalement des thèses en ligne non recensées par l'application Star.

Le portail des thèses renforcera la visibilité de la production doctorale française. Il fédérera les informations sans bousculer le paysage institutionnel : chaque acteur des thèses en France garde ses compétences en matière de signalement, de mise en ligne ou de numérisation. ●

Isabelle Mauger-Perez
Novembre 2010

2. *Entreprise et recherche publique : développer les synergies*, Rapport final du groupe de travail FutuRIS « Adéquation public-privé », avril 2008. www.anrt.asso.fr/fr/futuris/pdf/rapport_futuris_gt3_2008.pdf

3. Le comité de définition et d'opération du portail des thèses rassemble des représentants de l'administration centrale (MISTRD, DGESIP, DGRI), de la Conférence des présidents d'université, de la Conférence des grandes écoles, des formations doctorales dans les établissements habilités, des représentants de la recherche privée et du monde économique, des représentants des opérateurs documentaires nationaux des thèses (CNRS-Inist, Atelier national de reproduction des thèses, Cines), et des représentants d'associations : Association Bernard Gregory, ADBU, Aura.