

**Comité de définition et d'orientation
du portail des thèses (CDO)
Réunion du 30 septembre 2011 (Paris, MESR)
Relevé de décisions**

Présents :

Alain COLAS, MISTRD
Laurent REGNIER, DGESIP – MESR
Françoise PROFIT, DGESIP – MESR
Sandra JOSEPH, DGESIP – MESR
Nadia HILAL, CGE
Serge BAUIN, CNRS
Christiane STOCK, INIST/CNRS
Olivier ROUCHON, CINES
Mélanie RIBAS, Intelli'Agence Association Bernard Gregory
Sylvain COLLONGE, Institut de formation doctorale Paris 6
Jean-Marie BARBICHE, ADBU
Jean BERNON, AURA
Raymond BERARD, ABES
Yann NICOLAS, ABES
Isabelle MAUGER PEREZ, ABES

Absents :

Caroline WIEGANDT, Radio France
Francine PIERRE, THALES GROUP
Jean-Claude DUPAS, CPU

Excusés :

Michel MARIAN, MISTRD

1. Etat d'avancement du projet depuis la réunion du 25 mars 2011

L'application Star, dont la nouvelle version a ouvert il y a un peu plus d'un an, reçoit de plus en plus de thèses par mois. La barre des 6 000 thèses a été dépassée en août 2011.

2. Le volet interface publique : theses.fr

<http://www.theses.fr>, le moteur de recherche des thèses de doctorat, a ouvert le 11 juillet 2011 avec les données issues de l'application STAR. Presque 6 000 thèses soutenues dans les établissements ayant abandonné le dépôt papier au profit de l'électronique étaient référencées dans theses.fr le jour de l'ouverture ; le texte intégral était accessible pour plus de 4 000 d'entre elles.

Sur le net, une trentaine de blogs ont relayé l'annonce de l'ouverture et quelques critiques ont été publiées sur le web¹ mais aussi dans la presse professionnelle (un article dans la revue Bases n°282).

Les statistiques de consultation de theses.fr (fournies par Google analytics) sont les suivantes :

	Visiteurs uniques	Visites	Pages vues	Pages par visite
juil-11	12 711	15 815	86 810	5,49
août-11	11 868	15 288	75 383	4,93
sept-11	15 961	21 043	116 255	5,52

Les utilisateurs recourent à tous les modes de recherche, parfois en les combinant : explorer toute la base sans requête préalable, passer par l'encart de recherche en utilisant deux-trois mots-clés, affiner grâce aux facettes.

Le 27 septembre 2011, un deuxième onglet donnant accès aux thèses en préparation a vu le jour. 66 563 thèses en préparation issues du Fichier central des thèses ont été versées dans theses.fr.

Au préalable, l'ABES a essayé de lier chaque directeur de thèse présent dans la base du Fichier central des thèses à sa notice d'autorité dans le Sudoc. Le recours au référentiel IdRef a été systématique pour les établissements de soutenance. Désormais toutes les pages des personnes et organismes identifiés de manière univoque dans theses.fr peuvent contenir des informations sur les thèses soutenues mais aussi sur les thèses en préparation.

Exemples :

- Page de François Mancebo : <http://www.theses.fr/050524895>
- Page de l'université de Reims : <http://www.theses.fr/026403838>
- Page de l'École doctorale Sciences et Environnements (Bordeaux) <http://www.theses.fr/151367981>

Dans les prochaines semaines, theses.fr proposera un onglet pour chercher des personnes. L'internaute pourra par exemple avec la requête « urbanisme Grèce » trouver des noms de personnes (directeur de thèse et auteur) dont les thèses liées traitent de ces domaines.

¹ - Le point de vue de bibliothécaires-blogueurs : Etienne Cavalié (Lully) ["theses.fr : ce que ça change \(question de paradigme\)"](#); vu sur Bibliothèques

- le point de vue des juristes : Emmanuel Barthes, documentaliste juridique [« Theses.fr : comment chercher, le contenu, les innovations, les trous »](#); vu sur Précisément.org

- le point de vue des historiens : Franziska Heimburger et Émilien Ruiz [« Trouver des thèses en ligne : theses.fr et TEL »](#); vu sur La boîte à outils des historiens

- le point de vue des chercheurs en littérature jeunesse : Carnet de recherches de l'Association française de recherches sur les livres et objets culturels de l'enfance [« theses.fr »](#) vu sur Le magasin des enfants

3. Le volet thèses en préparation : STEP <http://step.theses.fr>

Les membres du CDO prennent connaissance des principaux écrans de l'application STEP, à la fois côté utilisateurs professionnels des établissements de soutenance mais aussi côté doctorant. L'Intelli'Agence et l'Université Paris 6 soulignent l'importance de mettre en valeur les thèses réalisées sous convention CIFRE ou réalisées en co-tutelle.

L'application STEP doit ouvrir début octobre [hors-réunion : le lancement a eu lieu le 4 octobre 2011] pour 84 établissements jusqu'alors utilisateurs du Fichier central des thèses. La liste de ces établissements ainsi que le nombre de thèses en préparation en leur sein sont présentés au CDO.

L'ABES, la Conférence des Grandes écoles et l'INIST-CNRS se sont mis d'accord sur le périmètre de STEP et celui de Thésa : les établissements membres de la CGE n'ont pas vocation à rejoindre STEP mais à rester dans le giron de Thésa, sauf à quelques exceptions près justifiées par l'histoire du Fichier central des thèses. Via l'API XML de theses.fr, Thésa pourra récupérer les données de ces établissements.

L'interfaçage avec Apogée » et l'Adum est prévu en 2012.

4. Prochain chantier : les données du Sudoc relatives aux thèses

L'algorithme qui permet d'interroger le Sudoc pour repérer les données des thèses est en phase bêta. Comme convenu lors du CDO du 28 septembre 2010, les efforts de traitement sont consacrés aux thèses postérieures à 1985 mais l'ABES espère pouvoir charger dans theses.fr toutes les données quelle que soit la date de soutenance.

Le Sudoc et theses.fr seront synchronisés : toute nouvelle thèse créée dans le Sudoc ou toute modification sera immédiatement visible dans theses.fr

Le chargement des données du Sudoc dans theses.fr (environ 450 000 notices) aura deux conséquences immédiates :

- le changement d'échelle : la description de plusieurs centaines de milliers de thèses soutenues sera désormais disponible.
- l'effondrement de la proportion de thèses pour lesquelles le texte intégral est accessible en ligne. L'enjeu sera alors pour l'ABES de récupérer le texte intégral de toutes les thèses qui ont fait l'objet d'une numérisation ; le signalement de ces thèses par chargement automatique dans le Sudoc est privilégié car il aura pour conséquence de les rendre accessibles dans theses.fr

5. Discussions autour des futurs services et des partenariats

Les utilisateurs de theses.fr pourront se créer un compte sur theses.fr en réutilisant un compte Gmail, LinkedIn ou autre.

L'ABES présente les services que theses.fr pourrait rendre à un utilisateur lambda (panier de notices et de requêtes..) et à un utilisateur reconnu par theses.fr comme étant l'auteur d'une thèse (statistiques de sa page dans theses.fr, service de mise en relation...)

Concernant les partenariats, l'ABES expose aux membres du CDO sa démarche :

- rencontrer les acteurs des services documentaires qui pourraient être mis en valeur dans theses.fr (accès au texte intégral, commande d'une impression à la demande...) : TEL, ANRT Lille...

La MISTRD indique que l'ANRT de Lille, dont la mission a été récemment confortée avec l'absorption des activités de l'ANRT de Grenoble, souhaite lancer un programme de numérisation. Ce dossier fera l'objet d'une rencontre avec l'ABES.

- être ouverte aux propositions portées par les acteurs des services non-documentaires, comme par exemple les services d'insertion professionnelle.

S'en suit une discussion notamment avec Intelli'Agence sur les axes de collaboration. Outre l'échange de données, déjà possible grâce aux API XML proposés par theses.fr, Intelli'Agence souligne l'intérêt de :

- faire figurer dans la description des thèses en préparation un résumé à destination du monde économique.
- valoriser tous les aspects de l'identité du chercheur, par exemple sa mobilité professionnelle.

Il est convenu qu'Intelli'Agence prendra contact début 2012 avec l'ABES pour voir dans quelle mesure les bases de données peuvent dialoguer. L'anonymat de la base de CV géré par Intelli'Agence ne semble pas être un problème.

6. Date de la prochaine réunion

28 septembre 2012 9h - 13h au Ministère.