



Numérisation des thèses

Elisabeth Fichez

► **To cite this version:**

Elisabeth Fichez. Numérisation des thèses. Du livre à Internet : quelles universités ?, Jun 2002, Paris, France. pp.56-60. edutice-00000473

HAL Id: edutice-00000473

<https://edutice.archives-ouvertes.fr/edutice-00000473>

Submitted on 18 Jun 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

"Numérisation des thèses"

Élisabeth Fichez

Domaine de compétences (enseignement):

Nouveaux dispositifs technologiques et industries du contenu (champs de la culture et de la formation) : approche des modes de structuration socio-économique ; rôle des acteurs ; approche par le biais des usages

Modélisation des processus de communication dans le contexte des organisations

Thématique de recherche générale :

Etude des mutations à l'œuvre dans les secteurs de la formation, des loisirs et de la culture liées à une industrialisation et à une marchandisation accrues. Ces mutations, qui ne s'effectuent pas de manière homogène, nécessitent des études empiriques tenant compte des spécificités des différents domaines. L'objectif est aussi de mettre en relation les conditions de transformation matérielle de ces champs et l'évolution de leur mode de fonctionnement symbolique.

Mots clefs :

Logiques socio-économiques, industrialisation, rationalisation

Dispositifs médiatisés, pratiques d'usage, appropriation, médiation sociale

Projets en cours

L'offre de formation multimédia et à distance dans les établissements d'enseignement supérieur du Nord Pas-de-Calais : offre, logiques, devenir. Projet dans le cadre du programme TIDSC de l'IFRESI (voir ci-dessous).

« Transfert des dispositifs de formation ouverte et à distance ». Préparation de la conférence de consensus organisée par l'Ecole Nationale Supérieure des Télécommunications, Paris, novembre 2001.

Emergence et structuration des campus numériques

Centres d'intérêts et responsabilités :

Séminaire Industrialisation de la Formation (SIF) : groupe de recherche national à l'origine de la problématique dans le champ des Sciences de la communication

Co-animation avec P. Grevet du programme de recherche « Technologies d'Information et Dynamique des Sociétés Contemporaines » dans le cadre de l'Institut Fédératif Régional pour l'Etude des Sociétés Industrielles (associé au CNRS)

Co-pilotage du groupe de recherche de l'IUFM de Lille sur « Innovation et recherche en formation »

Animation de l'équipe « Loisirs, Culture, formation : processus d'industrialisation et médiations » du laboratoire GERICO (Groupement des Equipes de Recherche Interdisciplinaires en Communication) de Lille 3 (10 chercheurs + 4 doctorants)

Membre du Comité de rédaction de la revue « Etudes de Communication » à Lille 3 et de la nouvelle revue « Distances et savoirs »

Membre de la commission nationale sur la diffusion électronique des thèses et Directrice de l'Atelier National de Reproduction des Thèses de Lille

Responsable du DESS « Management de la communication dans les organisations de services aux publics » à l'Université Lille 3 depuis 1996

Chef du projet « Campus numérique » à l'Université Lille 3 depuis janvier 2001

De mai à octobre 2000, Mission d'Administration Provisoire de l'Université Lille 3

DIFFUSION ELECTRONIQUE DES THESES : LE DISPOSITIF FRANCAIS

J'interviens dans ce colloque surtout au titre de la connaissance du dossier que me permet d'avoir une responsabilité exercée depuis janvier 1997 comme Directrice d'un des deux ateliers nationaux de reproduction des thèses en France, celui de Lille.

Il y a 5 ans, le dispositif de valorisation des thèses était celui défini dans l'arrêté du 25 septembre 1985, relatif au dépôt, au signalement, à la reproduction et à la diffusion de celles-ci sous forme micro-fichée. Mais au cours des années 1998 et surtout 1999, ce dispositif a commencé à être remis en question. Une commission nationale à laquelle j'ai participé, a travaillé deux ans, suite à quoi deux circulaires ministérielles, datant respectivement du 21 septembre 2000 et du 17 juillet 2001 ont préconisé aux universités françaises un dispositif complètement nouveau de diffusion en ligne. Celle du 17 juillet a en outre lancé un appel à candidatures aux universités volontaires pour expérimenter la mise en œuvre des nouvelles dispositions. J'ai été chargée en septembre de coordonner la réponse des universités de mon académie à cet appel et un groupe de délégués s'est mis au travail et a donc pu faire le point sur les difficultés et les réponses à leur apporter.

On se trouve donc en France dans une situation de transition qui risque de durer encore quelques années. Mon propos ce matin sera de donner un aperçu de cette situation et de dégager quelques conclusions concernant les avancées, les difficultés et les enjeux.

LE SYSTEME ANTERIEUR

Reproduction Off-set puis microfichage

Le système antérieur, toujours en vigueur d'ailleurs, mérite qu'on en souligne d'emblée un trait important : la France n'a pas attendu la technologie numérique pour se préoccuper de la valorisation et de la diffusion du savoir scientifique produit sous la forme des thèses dans les institutions universitaires.

Le Ministère de l'Éducation Nationale a mis en place dès 1971 un service public de reproduction des thèses en créant deux ateliers nationaux, l'un à Grenoble, l'autre à Lille pour les lettres. Jusqu'en 1983, la mission était très restreinte : ainsi, l'ANRT de Lille ne reproduisait en off-set que les thèses d'État de lettres, à la demande exclusive des auteurs et à titre onéreux. Puis, la technologie évoluant, le micro-fichage est devenu, à partir de 1983, systématique pour toutes les thèses de lettres, sciences humaines et droit soutenues dans les universités françaises et leur diffusion assurée gracieusement auprès des BIU et d'un certain nombre d'organismes français et étrangers, chaque auteur recevant personnellement un lot de 5 micro-fiches.

Ce système, institutionnalisé par la circulaire de 1985, a donc permis la constitution de deux catalogues assez complet de l'ensemble des thèses soutenues en France depuis près de trente ans : 37900 thèses enregistrées à Lille et 99200 à Grenoble, un stock catalogué et archivé qui suscite bien des envies en tant que patrimoine numérisable.

(Exploitation des fichiers numérisés : le système thèses à la carte)

Avec l'accord de son autorité de tutelle, l'Atelier de Lille s'est d'ailleurs lancé dans l'aventure de la numérisation pour les thèses soutenues depuis septembre 1996, avant même qu'il soit question de diffusion en ligne, pour offrir un service commercial dit de « Thèses à la carte » : sous réserve d'un contrat signé avec l'auteur de la thèse, celle-ci est numérisée et reproduite sur Docutech, intégralement et sans remaniement, dans un format réduit (16x24) au fur et à mesure des commandes provenant de personnes privées ou d'organismes (librairie, bibliothèques, service d'archives ...), français ou étrangers. Le grand succès que rencontre ce service, qui s'appuie actuellement sur un catalogue de 3700 thèses numérisées, montre qu'il correspond à une vraie demande.

Les raisons de l'évolution

Mais le problème essentiel que pose le système de diffusion par micro-fiches est son caractère institutionnel quasiment confidentiel : un audit mené en novembre 1997 à la demande du MEN (*Enquête sur les pratiques des utilisateurs du signalement des thèses et des utilisateurs de thèses, Société SCP Communication*) avait montré que le taux de consultation des thèses sous forme de

micro-fiches était extrêmement bas dans les bibliothèques et que celle-ci joue plus un rôle de conservation de la thèse (depuis une centaine d'années qu'il existe, le support argentique a fait ses preuves) que de moyen de diffusion.

Par ailleurs, la dissémination des moyens informatiques et le développement des réseaux ont changé assez radicalement la donne à la fois pour les auteurs de thèse et pour les institutions universitaires : Les thèses sont désormais produites « nativement » par leur auteur sous forme numérique et partir de l'exemplaire papier pour le traitement ultérieur paraît une perte de temps ; Quant aux universités, elles se positionnent de plus en plus à l'échelle locale comme producteurs et diffuseurs d'informations électroniques (revues, littérature grise, rapports etc.), les thèses soutenues apparaissant dès lors comme faisant partie de leur production propre.

LA MISE EN PLACE DU NOUVEAU SYSTEME

La préparation par une Commission nationale

Un groupe de travail, associant de nombreux partenaires institutionnels et des experts a donc été mis en place de février à juin 2000 et ses conclusions ont fait l'objet d'un rapport (<http://www.sup.adc.education.fr/bib/>) sur trois grands points : les objectifs recherchés, les options techniques (problème des formats) le mode d'organisation entre l'auteur, l'établissement et l'État relayé par des opérateurs nationaux.

Ce groupe de travail s'est transformé l'année suivante en groupe de projet et a élaboré deux documents mis à disposition de la communauté universitaire en juin 2001 : un guide pour la diffusion et l'archivage électroniques à l'attention des établissements d'enseignement supérieur ; un guide pour la rédaction et la présentation des thèses à l'usage des doctorants

Principales orientations du nouveau dispositif

En résumé, l'idée est que le doctorant dépose sa thèse avant la soutenance sous forme d'un document électronique prêt à être diffusé et archivé électroniquement par les soins de sa propre université. Si l'on regarde les choses plus en détail du point de vue des différents acteurs, cela signifie :

Pour le doctorant

- qu'il devra respecter des prescriptions techniques minimales (qui seront récapitulées dans une feuille de style qui lui sera proposée), le but étant que toutes les thèses puissent être traitées automatiquement pour un mode de lecture en ligne tirant parti des fonctionnalités de navigation (mode hypertexte).
- que, pour ce faire, il devra être formé, le plus tôt possible dans le cadre de l'école doctorale
- qu'il remettra sa thèse sous forme numérique avant soutenance pour vérification de la conformité aux prescriptions techniques (dans une période de transition, l'étudiant dépose sa thèse sous les deux formes - papier et disquette- pour ne pas retarder la soutenance).
- qu'il aura le droit d'accepter ou de refuser la diffusion sur Internet, (en toute connaissance de cause sur la visibilité qu'il donne ainsi ou pas à ses travaux). S'il n'accepte pas, sa thèse, sera consultable et empruntable sous forme papier dans la bibliothèque de l'université de soutenance et mise sur l'intranet de l'université (sauf cas de confidentialité signalé par le jury), mais elle ne sera plus diffusée du tout ailleurs.

Pour l'université

- qu'elle doit mettre en place les formations régulières dans le cadre de l'école
- qu'elle doit prévoir, seule ou en association avec d'autres universités, les ressources humaines et techniques nécessaires à la vérification des thèses sous forme numérique (deux heures par thèse en moyenne quand les étudiants ont été formés, jusqu'à trois jours quand ce n'est pas le cas) et au traitement en vue de la diffusion et de l'archivage
- l'archivage proprement dit est confié à la bibliothèque de l'université qui assure les opérations de signalement dans les catalogues local et national (Système Universitaire)

L'appui d'autres acteurs

Les établissements universitaires ne sont pas laissés seuls face à cette évolution :

Un groupe-projet sous la responsabilité du MEN regroupant notamment des équipes universitaires qui ont une certaine avance dans le domaine (Lyon 2, Marne-la-Vallée ...) organise des actions de formation et propose des logiciels de traitement qui ont déjà été expérimentés

Régionalement, les établissements peuvent coopérer et mettre leurs moyens en commun en faisant appel, comme c'est le cas dans le Nord Pas de Calais, à un prestataire de services pour la partie traitement technique

Deux agences nationales, l'ABES (Agence Bibliographique de l'Enseignement Supérieur) et le CINES (Centre Informatique National de l'Enseignement Supérieur) gardent des fonctions collectives, notamment d'archivage, nécessaires à la cohérence de l'ensemble. L'ANRT de Lille est invitée à se positionner comme opérateur technique sur le terrain régional ou national, à la demande des intéressés et garde la mission de produits dérivés.

Au final, il s'agit donc d'un dispositif qui donne la responsabilité centrale aux établissements en leur laissant le soin de trouver les modes d'organisation les plus adéquats sur le plan local et régional, mais qui leur offre un soutien logistique basé sur un partenariat actif de la part des membres du programme (mutualisation des compétences). Par ailleurs, l'État fixe les prescriptions indispensables à la cohérence générale du dispositif. C'est notamment lui qui validera pour chaque établissement le passage définitif de l'ancien système ou nouveau.

LA SITUATION SUR LE TERRAIN EN 2001-2002

Le choix fait par les responsables politiques pour la mise en œuvre du dispositif est celui d'une mise en route progressive, sur la base du volontariat.

Une quinzaine d'universités, dont certaines regroupées autour de projets collectifs (BRAIN pour la Région Rhône-Alpes, les universités de la région Nord Pas de Calais, des universités parisiennes) se sont déclarées partantes en novembre 2001. Elles ont été invitées en deux vagues à des journées de formation qui concernaient surtout les aspects techniques du traitement des disquettes pour les convertir aux normes SGML ou XML et les valider selon la DTD retenue (Document Type Definition = définition de la structure et des types d'éléments d'un document).

Pour l'instant peu d'établissements en dehors de ceux qui avaient déjà pris une certaine avance (Lyon 2 avec le projet Cyberthèses www.cybertheses.org, l'INSA avec le projet Cither <http://csidoc.insa-lyon.fr>) * ont effectivement mis en ligne des thèses consultables selon les normes du nouveau dispositif. C'est l'année universitaire prochaine qui devrait permettre de voir apparaître des mises en ligne plus significatives.

L'expérience que nous avons eue pour le collectif des Universités du Nord nous a montré que cette lenteur s'explique de plusieurs façons :

la nouveauté de la problématique pour les acteurs concernés (directeurs des écoles doctorales, responsables informatiques, responsables des SCD et ... thésards) qui ont besoin non seulement de connaître le dispositif, mais surtout de se l'approprier et de lever leurs appréhensions : en effet, des objections restent fortes au départ concernant l'intérêt pour le docteur de donner son accord pour une diffusion en ligne perçue comme concurrente de la diffusion sous forme d'articles ou d'ouvrages et comme dangereuse du point de vue du piratage intellectuel qu'elle peut favoriser

l'importance de l'effort à engager pour la formation des doctorants, surtout dans les universités de lettres où la culture scientifique est moins forte. Certaines universités, dans lesquelles un petit nombre de thèses sont soutenues, ne peuvent pas toujours engager elles-mêmes cet effort

l'état actuel des disquettes natives (compatibles avec le format RTF ou pour certaines disciplines scientifiques, au format L^AT_EX), dans l'échantillon d'une dizaine que nous avons retenue cette année, ne sont pas toujours réalisées avec une feuille de style, n'ont pas toujours mis en œuvre la gestion automatique du sommaire, des notes, des tableaux, des légendes sous les graphiques, des index, ...

Le temps à passer pour aboutir à un document numérique convenablement structuré et balisé est très long (3 jours de travail au moins, alors que l'objectif à terme pour que le projet de mise en ligne soit viable est de ne passer que deux ou trois heures maximum).

C'est donc une mise en œuvre à la fois dynamique qui se dessine (le nombre de réponses positives à l'appel d'offres a surpris), mais en même temps progressive pour la visibilité des résultats. On peut aussi penser que les universités qui ne se sont pas mises sur les rangs cette année le feront plus lentement par la suite, car ce sont celles qui ont moins de moyens humains et techniques.

CONCLUSIONS

La France s'est indéniablement engagée dans le mouvement international de diffusion en ligne du savoir scientifique, sans précipitation, en tenant compte de l'expérience acquise à travers le système antérieur de valorisation (intérêt d'une assistance et d'une validation nationale pour donner sa cohérence au système), tout en engageant la décentralisation imposée par l'évolution des techniques.

Les avantages du nouveau système sont indéniables :

le choix fait de la diffusion du texte intégral de la thèse, dans un format structuré, de façon telle que la consultation en ligne exploite les potentialités de lecture hypertextuelle.

le numérique va par ailleurs favoriser le développement des fichiers multimédia à côté des fichiers texte ... bien que, pour le moment, le traitement de ce type de fichiers pose encore beaucoup de problèmes.

la disponibilité des thèses devrait à terme être plus rapide que dans l'ancien système (facilement deux ans avant la mise à disposition des micro-fiches dans les bibliothèques à cause du retard mis par les universités à les transmettre)

l'accès à ces thèses se fait sous le régime de la gratuité

le signalement est amélioré grâce à un traitement affiné des métadonnées (SU DOC et moteurs de recherche)

intérêt indirect : la formation des doctorants à l'utilisation des technologies de publication et d'édition électronique constituera un réel élément de professionnalisation.

Toutefois, bien que les enjeux à l'échelle internationale **soient clairs** (présence de la littérature scientifique française sur Internet), **on regrettera** que le nouveau système va de fait exclure de toute diffusion les thèses dont les auteurs n'auront pas donné leur accord. D'où l'importance de maintenir différents modes de valorisation et notamment la version papier sous une forme maniable ou des versions sur supports numériques (CD, DVD) pour permettre aux auteurs d'avoir plusieurs choix possibles.