



## Diffusion électronique des thèses en Chine et ailleurs

**Yu Zhenglu et Pan Yuntao**

Centre de recherche sur les méthodes d'information  
Institut chinois d'information scientifique et technique  
Pékin, République populaire de Chine

*Traduit par : Isabelle Meuret et Jacques Hellemans  
Université Libre de Bruxelles, Belgique*

Date : 11/09/2006

Meeting:	<b>110 Bibliothèques de sciences sociales avec les bibliothèques gouvernementales</b>
Simultaneous Interpretation:	Non
WORLD LIBRARY AND INFORMATION CONGRESS: 72ND IFLA GENERAL CONFERENCE AND COUNCIL 20-24 August 2006, Seoul, Korea <a href="http://www.ifla.org/IV/ifla72/index.htm">http://www.ifla.org/IV/ifla72/index.htm</a>	

**Résumé :** *Les thèses constituent un précieux corpus pour les chercheurs. Leur numérisation et leur diffusion électronique, sous forme de MTE (mémoires et thèses électroniques) constitue une bonne manière de les mettre à disposition. Le gouvernement et les bibliothèques jouent un rôle important dans la conception d'une base de données des MTE en Chine. Cet article présente l'état actuel de la constitution de la base, la multiplicité des besoins des chercheurs ainsi que la tendance en matière de conception de bases dans d'autres pays et l'analyse de certains problèmes rencontrés aujourd'hui en Chine. Certaines recommandations sont faites afin qu'un meilleur service soit offert aux chercheurs.*

### 1. Mémoires et thèses

Les monographies, les articles de journaux, les actes de conférences ou les rapports de recherches et les thèses sont les sources de première importance pour la recherche. Les mémoires ou les thèses constituent une importante production des universités. Elle peut refléter le niveau de recherche et les domaines de spécialisation des organisations. En Chine, on en compte aujourd'hui plus de 700 qui peuvent accorder des doctorats ou des maîtrises. Aussi, présente-t-on chaque année de plus en plus de mémoires ou thèses. Leurs caractéristiques sont les suivantes :

- 1) Ils sont rédigés selon certaines règles, car les présenter signifie se soumettre à un examen en vue de l'obtention d'un diplôme.
- 2) La plupart se concentrent sur les théories récentes ou sur des problèmes urgents, importants ou très populaires. Ils sont donc des indicateurs des tendances scientifiques.
- 3) Ils contiennent nombre d'idées nouvelles et de résultats de recherche.
- 4) Sans restriction de longueur, l'auteur peut analyser et discuter un problème dans sa totalité.

- 5) On y trouve de nombreux schémas originaux, des analyses de données et des conférences.
- 6) La plupart des thèses restent inédites et sont relativement inaccessibles.

La Chine a restauré et fixé son système de diplômes à la suite de la réforme de 1978. Afin d'améliorer la collecte des thèses, l'Office des diplômes du Conseil d'État établit des règles depuis 1983, notamment que les mémoires et thèses soient envoyés à l'Institut d'information scientifique et technique, celles en sciences sociales, à l'Académie chinoise des sciences sociales et toutes, à la Bibliothèque nationale.

## **2. Les Mémoires et thèses électroniques ont la cote**

Le développement rapide de l'industrie de l'information mène à la version électronique des mémoires ou thèses. C'est la tendance. Des projet de MTE voient le jour. Les MTE facilitent la collecte et la conservation, le partage et la diffusion de la connaissance et de la technologie. Ils peuvent aussi présenter un avantage financier à long terme. Ils vont accélérer le rythme de la recherche scientifique et améliorer la communication entre les chercheurs.

Entretemps, les chercheurs économiseront temps et argent dans leurs efforts de repérage de thèses. Pour une recherche documentaire, ils peuvent obtenir plus d'informations sur les auteurs. Les données peuvent être exhaustives et continues. Les chercheurs peuvent être stimulés et, d'autre part, pouvoir éviter les répétitions.

Le programme MTE possède un bon mécanisme d'indexation, lequel facilite le repérage de données en plus des thèses elles-mêmes. Dans une certaine mesure, on peut trouver qui fait quoi dans le domaine, quel est le niveau des recherches et quels sont les projets similaires. On peut également apprendre combien de thèses sont dirigées par une même personne. Différentes combinaisons de recherche documentaire peuvent aussi mener à différents résultats et à différents indices.

En fait, bien des personnes ont besoin des MTE et pas seulement des spécialistes : des personnes travaillant pour l'État, dans le monde du commerce, dans l'enseignement, dans l'armée, etc.

D'un point de vue économique, on peut classer les MTE selon leur rentabilité ou non. Selon les opérateurs, on peut les répartir en MTE universitaires, en MTE d'organisations et en MTE à visée commerciale.

## **3. Les MTE en Chine**

La conception des MTE se développe très rapidement en Chine. De nombreuses universités ont commencé le projet et proposé une série de mesures. Mais la taille et le niveau de conception des MTE sont moindres que les instructions sur Internet. et il n'y a pas de bases de données MTE de même importance en Chine.

Il y a présentement quatre bases de données MTE très populaires en Chine :

### **1) Base de données des thèses chinoises**

L'Institut de l'information scientifique et technique de Chine (sigle anglais : Istic) est le bureau de dépôt légal des mémoires et des thèses. On a commencé à y rassembler les thèses

chinoises en 1980. Un bulletin signalétique, conçu par l'Istic, a vu le jour en 1984. Wanfang data l'a transformé en disque compact – « Base de données des thèses chinoises (en anglais, Chinese Dissertations Database ou CCDB) - en 1995.

Basé sur les données de l'Istic, Wangfang data a fait de grands et rapides progrès en matière de MTE. En mai 2006, la base de données a presque atteint les 710 000 titres. Maintenant, la base complète – CDDDBFT -, laquelle archive thèses et mémoires ultérieurs à l'an 2000, est en croissance. On y trouve actuellement plus de 300 000 titres et elle augmente de 100 000 chaque année.

Trois termes de recherche documentaires peuvent être combinés à l'occasion d'une recherche de base dans CDDDBFT. Le résultat détaillée comprend auteur, directeur de thèse, organisme subventionnaire, temps alloué, nombre de pages, langue, cote de classement en bibliothèque, numéro de conservation, mots clés et résumé. L'un des trois termes de l'index peut être le sujet. Le CDDDBFT possède une bonne fonction de repérage alphabétique.

En plus de la base de données des mémoires et thèses, il y a aussi des journaux en texte intégral, des actes de congrès, les brevets, les rapports de compagnies et d'autres bases. Wanfang peut fournir d'autres produits semblables. Pour les recherches documentaires des mémoires et thèses, il y a des liens automatiques par auteurs, professeurs, sujets et mots clés. Sous le lien automatique d'un auteur, on peut trouver tous ses écrits, mémoires ou thèses, articles de journaux, comptes rendus de conférences, rapports et autres informations. Le logiciel du navigateur de textes est Adobe Acrobat.

## **2) Bases sélectives de textes intégraux de thèses de doctorat ou de maîtrise**

La base de données est développée par la société de journaux électroniques China National Knowledge Infrastructure et la société TTOD. On y archive les mémoires et les thèses depuis 1999. On comptait 270 000 titres au 31 décembre 2005.

Il y a deux options de navigation pour la recherche documentaire : la navigation par la classification de connaissances et la navigation par sujet. La CDMD (sigle anglais) propose trois méthodes de recherche : élémentaire, avancée et spécialisée. Le logiciel du navigateur de textes complets est Cajviewer. Grâce à la mise à jour quotidienne, chaque site miroir peut également être actualisé quotidiennement par transmission internet ou satellite. Le disque compact, quant à lui, est mis à jour une fois par. La CDMD a ses propres règles de compensation monétaire.

## **3) Bibliothèque universitaire et système d'information de Chine – base de données de mémoires et thèses**

La base de données ETD est produite par Bibliothèque universitaire et système d'information (= China Academic Library & Information System ou CALIS). Elle contient à ce jour 70 000 documents, mémoires ou thèses. Les membres participants au projet sont 83 universités.

On rassemble les métadonnées selon les règles OAI (open archives initiative protocol for metadata...) et on lie les métadonnées au texte complet grâce à la technologie de l'URL. Son modèle est donc la recherche centralisée et la distribution du texte intégral. Les universités

participantes devraient établir leur propre base de données de thèses et la soumettre à l'université de Tsinghua en vue de l'élaboration d'une base unique.

Les seize premières pages des mémoires et thèses sont gratuites pour les membres. Les lecteurs peuvent obtenir un texte complet grâce au système d'emprunt commutatif de l'université de Tsinghua et par le réseau local des membres.

Afin d'encourager l'utilisation de ce canal, les organisateurs diffusent régulièrement sur le site de l'université de brèves informations à propos des mémoires ou des thèses.

#### **4) PQDD (ProQuest digital dissertations)**

PQDD est publié par la compagnie américaine ProQuest information and learning (autrefois UMI). En 2000, Dissertation abstracts database contenait plus de 1,6 million de thèses de doctorat et de mémoires de maîtrise. Ils proviennent de plus de 1000 universités à travers le monde. Le tout premier titre archivé remonte à 1861.

PQDD est la base de données de thèses la plus importante et la plus puissante du monde. Elle fonctionne selon une procédure détaillée. La plus grande partie de la base concerne les sciences de la santé. S'il y a plus de sept commandes d'une thèse, l'auteur peut obtenir 10 % du montant des ventes à titre de compensation. On peut obtenir gratuitement les 24 premières pages d'un certain nombre de thèses postérieures à 1997.

PQDD est une organisation commerciale. Après plusieurs années de progrès, la société peut procurer de très bons services aux étudiants, aux universités et aux lecteurs.

Depuis 2002, des universités et des établissements se sont regroupés pour acheter les thèses de PQDD. China national Sci Tech information I/E sert d'agence. Bibliothèque universitaire et système d'information est responsable de l'organisation. Tous les membres peuvent partager les documents commandés par le groupe.

#### **4. Problème actuel**

Le développement des MTE est très populaire en Chine aujourd'hui mais certains aspects devraient être améliorés et des problèmes, résolus.

- 1) Dépôt légal. Il y a un système de dépôt des mémoires et thèses. Cependant, les universités et les établissements ne sont pas soumis aux organisations d'archivage et, en conséquence, le travail de collecte n'est pas facile et certains mémoires et thèses sont oubliés.
- 2) Résumés. Certains résumés sont longs et contiennent des schémas et des formules. Ils ne respectent pas tous les critères internationaux.
- 3) Mots clés. Certains étudiants ne comprennent pas la signification des mots clés. Aussi les mots clés sont-ils parfois éloignés des sujets abordés. Cela nuit à l'efficacité de la recherche.
- 4) Recherche rétrospective. La conservation à long terme du papier étant difficile, la procédure rétrospective est compliquée et coûteuse.
- 5) Métadonnées. Bibliothèque universitaire et système d'information chinois a mis de l'avant un plan de standardisation des critères de métadonnées.
- 6) Droit d'auteur. Le processus de transfert complet devrait protéger le droit des auteurs, des bibliothèques et des utilisateurs. Le transfert des thèses sur Internet peut être

considéré comme une publication : les droits d'auteur devraient être protégés. Les bibliothèques constituent des bases de données et les diffusent : leurs droits devraient l'être aussi. Et les droits des lecteurs d'utiliser légalement les bases devraient être également assurés.

D'après la loi sur les droits d'auteurs en Chine, seule la base de données de départ peut être protégée. Cela suppose un choix original des ressources et, bien sûr, la constitution d'une base de données.

- 7) Étudiants. Certains étudiants refusent de diffuser leur mémoire ou thèse. Trois raisons sont invoquées : (1) les résultats vont être publiés dans une revue spécialisée, et donc on retarde la publication complète; (2) les mémoires et les thèses relèvent du secret ; (3) les résultats des recherches ont un rapport avec les brevets.

## **5. Les MTE dans les autres pays.**

### **1) Virginia Tech-ETD (Etats-Unis)**

Les États-Unis ont accordé beaucoup d'attention à la conception des MTE. Avant la Seconde Guerre mondiale, des bibliothèques et universités prirent une série d'initiatives dont la reproduction et le microfilmage des thèses. Virginia Tech-ETD est l'un des pionniers dans le domaine. L'université Virginia Tech, avec la société UMI et d'autres organisations, avaient commencé un projet de recherche sur les MTE en 1987.

À présent, la soumission ou le dépôt des thèses est devenu un souhait élémentaire des étudiants. La raison qui a justifié la mise sur pied de VT-ETD était le petit nombre de mémoires de maîtrise envoyés à la société UMI, la plus grande organisation de collecte dans le domaine. La base a été créée par Virginia Tech University Library and Graduate School. À ce jour, on y trouve 7941 mémoires ou thèses.

VT-ETD attache beaucoup d'importance à l'analyse des faits. Dans le tableau ci-dessous, on peut voir que les accès des entreprises commerciales ont atteint le nombre de 405 302 en 2002/03, soit 49,4 % de l'ensemble. Les consultations proviennent du milieu de l'enseignement, des entreprises, du gouvernement, de l'armée, de réseaux, etc.

Le tableau suivant montre la disponibilité croissante des VT-ETD. De 1998 à 2004, il y eut de plus en plus d'étudiants qui acceptaient que leur mémoire ou thèse soit consultable en ligne pendant que les récalcitrants diminuaient.

### **2) Dissonline (Allemagne)**

La Bibliothèque nationale allemande a commencé à regrouper les MTE sur Internet et à concevoir un index avec métadonnées en 1998. Elle se chargea du transfert des métadonnées et des accès à long terme vers les MTE. Plus tard, la Bibliothèque a fondé une agence de coordination pour promouvoir la conception de MTE. Le programme a été soutenu par le comité central de la recherche (DFG). Le Dissonline a archivé 97 MTE dès 1998, 20 957 en 2003 et 24 158 en 2004. Le travail de l'agence de coordination implique :

- point de contact et échange d'informations
- encouragement à la coopération et transfert de savoir-faire
- coordination des avancées technologiques et ajustements
- coopération avec les organisations nationales et internationales

### **3) Royaume-Uni**

La British Library n'est pas l'organisation de dépôt légal des mémoires et thèses mais elle en archive un grand nombre. La bibliothèque a fondé le centre de service des mémoires et des thèses, lequel en a archivé 150 000. Avec le développement de la technologie de l'information, le Royaume-Uni a mis un lot de mémoires et thèses en ligne en 1994. On peut y retrouver 460.000 résumés de thèses écrites depuis 1716.

#### **4) Le Canada**

La Bibliothèque nationale du Canada a mis sur pied un programme de MTE et un comité de consultation sur les modèles d'opération des MTE et elle prend en charge les problèmes liés à la mise en place.

### **6. *Recommandations***

Tout d'abord, l'établissement de critères suffisants pour la constitution de MTE, tels que le format d'écriture, le format de la table des matières et des schémas, le résumé, les mots clés et la classification par sujet. Secundo, consacrer plus d'efforts dans le domaine de la technologie de l'information. Il y a aujourd'hui de nombreuses recherches sur les métadonnées partout dans le monde. Une équipe des États-Unis a mis en avant les EDT-ms : Interoperability metadata standard for electronic theses and dissertations. Des métadonnées d'application plus amples et puissantes dans le domaine des bibliothèques et l'industrie de l'information constituent le DC (Dublin Core) Metadata Element Set. Et troisièmement, avoir recours à des ressources humaines mieux formées, établir un système de services compétents et coordonner les organisations. Ces mesures peuvent assurer la collecte de MTE et les rendre plus facilement repérables.