



HAL
open science

Dissemination and preservation of French print and electronic theses

Pierrette Paillassard, Joachim Schöpfel, Christiane Stock

► **To cite this version:**

Pierrette Paillassard, Joachim Schöpfel, Christiane Stock. Dissemination and preservation of French print and electronic theses. *The Grey Journal*, 2007, 3 (2), pp.77-93. sic_00380488

HAL Id: sic_00380488

https://archivesic.ccsd.cnrs.fr/sic_00380488v1

Submitted on 1 Jun 2009

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.

Dissemination and preservation of French print and electronic theses

Paillassard, P., Schöpfel, J. & Stock, C.

Bionotes

Pierrette Paillassard is a librarian in charge of theses and dissertations in the field of the Communication and Information Sciences, and conferences in Humanities and Social Sciences. She is also administrator of the open archive "mémSIC". She was member of the AFNOR expert group who prepared the TEF metadata scheme.

Contact address: Pierrette Paillassard, INIST-CNRS, 2 allée du Parc de Brabois, CS 10310, 54519 Vandoeuvre Cedex, France. E-Mail: pierrette.paillassard@inist.fr

Joachim Schöpfel is Head of the E-publishing and Document Supply Department at INIST-CNRS and Lecturer on Scientific Information at the University of Nancy.

Contact address: Joachim Schöpfel, INIST-CNRS, 2 allée du Parc de Brabois, CS 10310, 54519 Vandoeuvre Cedex, France. E-Mail: joachim.schopfel@inist.fr

Christiane Stock is the Head of the Monographs and Grey Literature service at INIST, in charge of repositories like LARA (reports), mémSIC (master's theses in information sciences) and OpenSIGLE. She was member of the AFNOR expert group who prepared the TEF metadata scheme.

Contact address: Christiane Stock, INIST-CNRS, 2 allée du Parc de Brabois, CS 10310, 54519 Vandoeuvre Cedex, France. E-Mail: christiane.stock@inist.fr

Abstract

How do you discover and locate a French thesis, how do you get hold of a paper copy and how do you access the full text of electronic theses and dissertations (ETDS)? What are the catalogues and databases referencing theses? Where are the archives, and are they open? What is the legal environment that rules the emerging structures and tools?

This paper presents the former database for print theses "Téléthèses" that merged with the national academic union catalogue "Sudoc", and gives an overview on initiatives for open archives and repositories for electronic theses and dissertations as well as the national program for these documents (STAR). Practical council is provided for the search of French theses, legal and metadata aspects are discussed, seven OAI-projects are presented in detail (CITHER, Cyberthèses, IRIS, INRIA, Mathdoc, PASTEL, TEL-HAL), and a glossary is added for some French acronyms.

The following article is an updated and revised version of our communication presented at the GL6 conference in 2004: Paillassard, P., Schöpfel, J. & Stock, C.: "How to get a French doctoral thesis, especially when you aren't French". - In: Farace, D. & Frantzen, J. (ed.): *GL6 Conference Proceedings. Sixth International Conference on Grey Literature: Work on Grey in Progress*. New York, 6-7 December 2004. - Amsterdam: TextRelease 2005.

Introduction: What is a French doctoral thesis?

Considered as scientific publications, French doctoral theses constitute an important part of scholarly communication. Following scientometrics, they represent 10-20% of indexed academic research in STM (OST* 2002).

Theses are often the result of 3-4 years of research. At the same time they are an administrative document necessary to obtain the doctoral degree. In some disciplines they are regarded as a result of teamwork and appear in the list of publications of the research laboratory (Mermet et al. 1998).

French universities are autonomous; each one delivers its own degrees and preserves the theses in its library. In the past, before 1985, the graduate student had to deposit a certain number of copies that varied according to local rules (30-180) for library interlending and exchange purposes. There are more than 100 universities in France, each one with one or more catalogues and with a specific logic of preservation and supply. Furthermore, academic communities – sciences, humanities, medicine, law etc. - hold different views and have different practices and traditions. And last but not least, local autonomy and responsibility are "counterbalanced" by a national framework structure, the French interlibrary loan network.

So, how to find a French thesis? And once found, how to get it? The following communication tries to give some practical hints and perspectives, imbedded in a larger description of the development of the production, processing and preservation of French doctoral theses and an overview of the principal actors, catalogues and databases.

Dissemination and preservation of French print theses (1985-2000)

The French government published in 1985 a decree that regulated and improved the deposit and dissemination of doctoral theses. Its main purpose was to guarantee the deposit of the doctoral thesis, to harmonize the number of copies to submit, to facilitate the identification and availability of the documents, and to move the format of preservation and dissemination from paper to microfiche in order to gain shelf-space, ensure long-term conservation and provide easier access.

Subsequent to the 1985 decree, the French Ministry of Education created a four-level national network for theses:

1. Registration: each university had to create a special service for doctoral theses ("service de doctorat").
2. Editing and reproduction: two public institutions (ANRT*) in Lille and Grenoble transformed the print originals into microfiches.
3. Recording: three public input centres (INIST* for sciences and technology) centralized the creation of bibliographic records from the registration form.
4. Dissemination: all records were loaded into a national online database called "Téléthèses".

(a) Deposit, registration and dissemination

Three weeks before the date of defense the candidate fills in two copies of a registration form and submits several print copies of his thesis at the "service de doctorat": one copy for each member of the jury, and three copies for the library.

The registration form contains personal, administrative and bibliographic data (including a French abstract, French keywords and, in later years, an English translation of title and abstract) and is used for the examination process as well as for the input into the national database.

The jury may ask for modifications of the thesis to be finished within three months after the date of defense. Once the final official version submitted, the university president authorizes its reproduction and dissemination.

The print copies and registration forms are transmitted to the university library. Depending on the scientific domain (social sciences and humanities, including economics and law; medicine; and sciences), the registration form is sent to one of the three input centres.

If authorized for reproduction, a print copy is shipped to one of the national theses reproduction services (ANRT) that produces a microform version. All university libraries and some other academic institutions receive a copy on microfiche. The students' guide mentions an average dissemination of 200 microform copies per thesis (Ministère 1994).

If the thesis has been published, the graduate student must deposit 10 sample issues of the publication at the university library (30 if the student received public funding for the publication). In this case, the thesis is not converted into a microform.

French theses are not deposited at the National library (BNF*), and they are not included in its national bibliography.

(b) Referencing – from print bibliography to online catalogue

Up to 1996, the Ministry published an annual print bibliography “Inventaire des thèses”. This catalogue was divided into three sections, “social sciences and humanities”, “medical sciences” and “sciences”. In 1986, a national database called “Téléthèses” was created, hosted on a university server and accessible through “Minitel”, a very popular Videotex online service launched in France in 1982 but inaccessible from foreign countries.

Records in the online database referred to theses going back to 1972 for sciences, social sciences and humanities, to 1983 for medical sciences and pharmacology and 1990 for veterinary sciences. Each record contained minimal bibliographic data, an abstract and keywords in French and later also in English. Authority lists were used for the university, type of degree and scientific domain. From 1986 on the university based “services de doctorat” attributed a unique identifier (national identification number) that was included in the database record. An ISBN was only attributed if the thesis had been published.

Between 1995 and 2003 the Téléthèses database was also published in a CD-Rom version called “Docthèses”, making the database available to foreign countries. The following table contains the number of French doctoral theses referenced by “Docthèses”:

| Year | total | SS&H | Sciences | Medical Sc |
|-------|--------|-------|----------|------------|
| 1993 | 18813 | 2844 | 6335 | 9634 |
| 1994 | 19344 | 3307 | 6613 | 9424 |
| 1995 | 18449 | 3019 | 6422 | 9008 |
| 1996 | 19542 | 3460 | 7181 | 8901 |
| 1997 | 19886 | 3637 | 7083 | 9166 |
| 1998 | 18444 | 3398 | 6655 | 8391 |
| 1999 | 17826 | 3530 | 6011 | 8285 |
| 2000* | 10662 | 3572 | 4977 | 2113 |
| 2001* | 8922 | 2612 | 6310 | ? |
| 2002* | 10720 | 3209 | 7511 | |
| total | 162608 | 32588 | 65098 | 64922 |

(*) 2000-2002 are transition years and the number of theses is not complete

Table 1: Theses referenced in the “Docthèses” CD-Rom database (1993-2002)

In 2000 the Téléthèses database moved from Videotex to a web server hosted by ABES*. At the same time, all records were uploaded into the new national academic union catalogue Sudoc*. Today, the all university libraries create their records directly in the Sudoc, and the online and CD-Rom databases of theses have disappeared together with the four-level national network. The Sudoc catalogue contains actually more than 530,000 theses and will be searchable through Google in 2007.

The most important former input centre, INIST, preserves more than 100,000 French STI theses in its collections. Most of them are searchable through its online database "Article@INIST"* and via the database "cat.inist" through Google and Google Scholar.

(c) Limits and critics

The 1985 decision facilitated recording and availability of French theses. These rules have been applied until 2001. Nevertheless, essential critics arose especially from academic librarians:

- Workload: initially, university libraries couldn't download the records from the database but had to key them again for their own catalogue.
- Incomplete information: especially in humanities and social sciences, librarians wanted to increase reference quality by adding national subject headings (RAMEAU*).
- Delay: the time between the date of defence and the integration of the record into the union catalogue was often rather long.
- Supply price: the cost of dissemination of theses through print copies from microforms was generally considered as too high.

Improvements were made since 1996, in particular as mentioned before through the development of the Sudoc functionalities. But it was above all the advancement of electronic theses and dissertations in France and other countries that led from 1998 on to the creation of ETD repositories and moved the government to a change in national politics (see below).

Electronic theses: legal aspects and metadata

In the 1980's a thesis was considered as a university document that should be disseminated as widely as possible. According to their examination regulations, the universities considered the jury's authorization sufficient for dissemination.

With the appearance of ETDs and the evolution of the author's rights, a thesis is no longer seen as a "university document" but as a work subject to intellectual property rights.

Today the explicit authorization by the author of the thesis (= copyright holder) is necessary for the electronic dissemination, in addition to the jury's decision. This authorization should be requested when the thesis is submitted (Jolly 2000). Furthermore, some universities ask for a declaration of conformity between electronic and print version and/or between the native deposit format and the XML version (Six&Dix 2004).

Some universities (Metz, for instance) already started to search for their former graduate students in order to obtain an authorization for retro-digitisation and online access of older theses.

Following the results of the Jolly report, AFNOR*, the French standardization organization charged an expert group to define the metadata required with the national deposit of ETDs. The recommendation was published in its extended version TEF* 2.0 in spring 2006 (TEF 2006).

Based initially on the Dublin Core, the new scheme TEF is far more detailed. In addition to the "traditional" bibliographic metadata, the scheme includes data for administrative needs as well as information related to the life cycle of ETDs and to the rights management (METS rights). Data of local interest (e.g. name of the research unit) are optional. TEF offers a match to Unimarc fields for the union catalogue Sudoc and is OAI-PMH compliant.

Electronic theses present new challenges: segmentation may be necessary to restrict access to confidential parts; it is necessary to distinguish archival versions and versions for dissemination. Therefore TEF includes metadata that will allow version management and migration.

For a detailed presentation of TEF see the papers of Boudia and Gomez de Regil presented at GL7 (Boudia 2006) and Nicolas at ETD 2006 (Nicolas 2006).

French ETD archives in 2007

The following chapter offers an overview of the seven most representative digital archives¹ that give free access to French ETDs. These archives were developed since 1997 and 1998 by French universities, engineering schools, national institutes and the CNRS*. Figures and data are from January, 2007. The appendix contains more detailed information for each of these ETD archives.²

(a) Physics, mathematics, chemistry and engineering sciences

- **CITHER**: produced by INSA* Lyon with 571 theses in the engineering field.
- **PASTEL**: produced by the Paris Institute of Technology (ParisTech with 11 independent engineering, management and business schools). PASTEL contains 968 theses with online access to the full text.
- **MathDoc***: developed by the University of Grenoble-1 and the CNRS, MathDoc is one of the oldest French archives with more than 1000 theses in mathematics.

¹ Digital archives does not necessarily mean open archive. By digital archive we understand platforms, institutional servers, repositories who give access to ETD full text in a permanent way.

² See a list with French ETD full text repositories: ABES (2006), Quelques sites pour trouver des thèses en texte intégral [Online]. -<<http://www.abes.fr/abes/page.430.sites-concernant-les-theses.html>>

- **INRIA***: the INRIA archive gives access to more than 600 theses in computer science and control between 1985 and 2005. Since 2005, students are encouraged to submit their document to TEL-HAL (about 300 deposits until January 2007)

(b) Multidisciplinary archives

- **Cyberthèses**: a common project between Canadian and French universities (Montreal, Lyon), gives for example access to 883 multidisciplinary ETDs for Lyon 2 University.
- **IRIS**: a multi-type institutional archive, produced by the University Library of Lille 1. It contains about 500 full-text digital theses submitted at Lille 1 (sciences, technology and social sciences). IRIS is a partial successor of Grisemine a multidisciplinary and multitype repository which was presented at GL'5 (Claerebout, 2003)
- **TEL - HAL**: created by the CCSD* and MathDoc*. It is today the most comprehensive French repository with 6028 ETDs in full text, covering all domains but mostly physics, mathematics and engineering sciences. TEL-HAL is also searchable through HAL, a repository including eprints and other document types.

(c) Typology of archives

Four different types of archives can be distinguished, even if these types are not exclusive:

- The **institutional archive** contains all theses of one (CITHER) or more than one structure (PASTEL).
- The **domain-specific archive** gives access to ETDs from different establishments but of the same scientific domain (MathDoc).
- The **collaborative or multi-side archive** offers facilities to different structures (International program Cyberthèses).
- The **multi-type archive** contains ETDs but also other academic literature - preprints, conference papers, courseware and so on (INRIA).

The most frequent type seems to be the collaborative or multi-side archive. The cooperation can be realized on different levels:

- **Management and administration**: Cyberthèses is co-managed mainly by the universities of Montreal and Lyon-2 and a French foundation for information highways (Fonds Francophone des Inforoutes).
- **Coverage**: Multilingual research interfaces are more and more frequent. TEL-HAL offers French, English and German versions. ETDs are in different languages and come from different European, African and American countries.

This willingness to cooperate is reinforced by the use of metadata harvesting through the OAI-PMH protocol and the use of open source software. PASTEL, TEL-HAL and Cyberthèses are declared OAI compliant.

(d) Other services and functionalities

Some archives offer more than full text access and include special and complementary services, for instance:

- Complete editorial chain: Cyberthèses proposes a complete editorial chain called “Cyberdocs” going from a document model to the conversion into a fully structured XML document using the TEI lite DTD. Discussion lists and downloadable tools complete the offer.
- Links to online services: MathDoc offers links to different special portals and online services such as the Zentralblatt-MATH (FIZ Karlsruhe), the MathSciNet (American mathematical society), or Springer Link.
- Online submission: TEL-HAL and PASTEL permit online submission by the author (self-archiving). Even so, in most cases the institution controls metadata and documents before making them available. Changes in the workflow may be possible with new national organisation.
- Technical progress: evolution of technical platforms includes multilingual user interfaces furthering still the access to ETD's. RSS feeds are added to numerous sites.
- Inclusion of course material : IRIS starts to cite course material following the LOMfr³ metadata scheme.

Up to now, we found no study on usage patterns of the different French archives and systems comparable to Zhang, Lee & You 2001 for the Korean KISTI system.

Detailed aspects can be found in the individual presentations of each archive (see appendix).

The national program for electronic theses and dissertations (1998-2007)

As mentioned above, motivated by the increasing number of electronic theses and dissertations, the French Ministry of Education published in 1998 the outlines of a national ETD server (Okret 1998). The project was meant to substitute the 1985 network, though preserving its underlying doctrine, a centralized structure based on the national academic union catalogue, the Sudoc, and similar software⁴ and procedures for all universities. Three other assumptions were made:

³ Educnet (2007), Métadonnées, normes et standards [Online]. -<
<http://www.educnet.education.fr/dossier/metadata/lom2.htm>>

⁴ “Cyberthèses” developed by the University Press of Montreal and the University of Lyon 2.

- Each ETD record in the Sudoc should be linked to the full-text (URL link from the 856 field).
- Each ETD should be archived on a local, campus-based server.
- A national backup server should contain part or all of French ETDs.

Between 1999 and 2000, a working group addressed the technical, organisational and financial features of this ambitious project (Jolly 2000). In spite of nationwide incentive action and promotion, the results were limited. Four years later, only 360 ETDs were compliant with the governmental guidelines, corresponding to hardly 0.5% of the theses recorded between 2001 and 2004 and only 8% of French ETDs in 2004. An audit ordered by the Ministry addressed some reasons for this situation (Six&Dix, 2004):

1. A unique ETD model was unrealistic and non-adapted to the heterogeneous needs, traditions and initiatives of scientific and academic communities that had started to develop their own and often less complicated ETD solutions.
2. The need for new technical knowledge and procedures, training of graduate students and investment for new soft- and hardware was underestimated. Generally, both academics and librarians considered the technical requirements as too complicated.
3. The Ministry's initial evaluation of human and budget resources was too optimistic, governmental funding was insufficient, and local investment by universities was often limited or inexistent.

The Six&Dix report recommended a modular network based on mixed deposit (print/native format), PDF/XML preservation and PDF/HTML supply and on a combination of centralized software (Sudoc and CINES* ETD archive) and campus- or community-based solutions.

Following these recommendations, the Ministry reconsidered its first project and elaborated a new and more realistic program. The related decisions were published 2005 and 2006 and can be summarized in four points:

1. **Deposit:** Print or electronic format, depending on the choice of each university.
2. **Recording:** Metadata TEF. Conversion into UNIMARC by ABES.
3. **Dissemination:** Conversion by the university into HTML and/or PDF.
4. **Preservation:** Conversion by the university into XML or PDF.

Each university decides which server to use or recommend for the preservation and dissemination of ETDs, for instance a campus-based server, the CNRS TEL-HAL open archive or the Sudoc portal, while the CINES guarantees the perennial preservation.

Since 2005 ABES and CINES prepare a new software tool to help with ETD logistics. STAR* (Signalement des thèses, archivage et recherche = referencing, archiving and retrieval of ETDs)

(STAR 2007) is the new “hub” through which electronic theses must pass in transit before further dissemination. On the input side STAR allows universities who have not chosen their own platform to manage the workflows of deposit, metadata (using the TEF metadata scheme) and administrative validation of theses with different validation levels.

As an output STAR proposes the following services:

- conversion of metadata to the Unimarc format and integration into the Union catalog Sudoc and its related authority files (national bibliography for theses),
- creation of a unique persistent identifier for each thesis,
- export of the archive version of the document and appropriate metadata (bibliographic, preservation) to the perennial archive on the CINES host,
- dissemination of a public version (if parts of the theses are confidential). STAR can provide other archives and servers with a public version and metadata in different formats: Unimarc/XML, DC, ETD-MS, TEF.

A future version of STAR will allow the integration of data from local systems through the OAI metadata harvesting protocol. Keying in of metadata will no longer be necessary. A combination of different input modes will also be possible, offering interesting solutions for different cases.

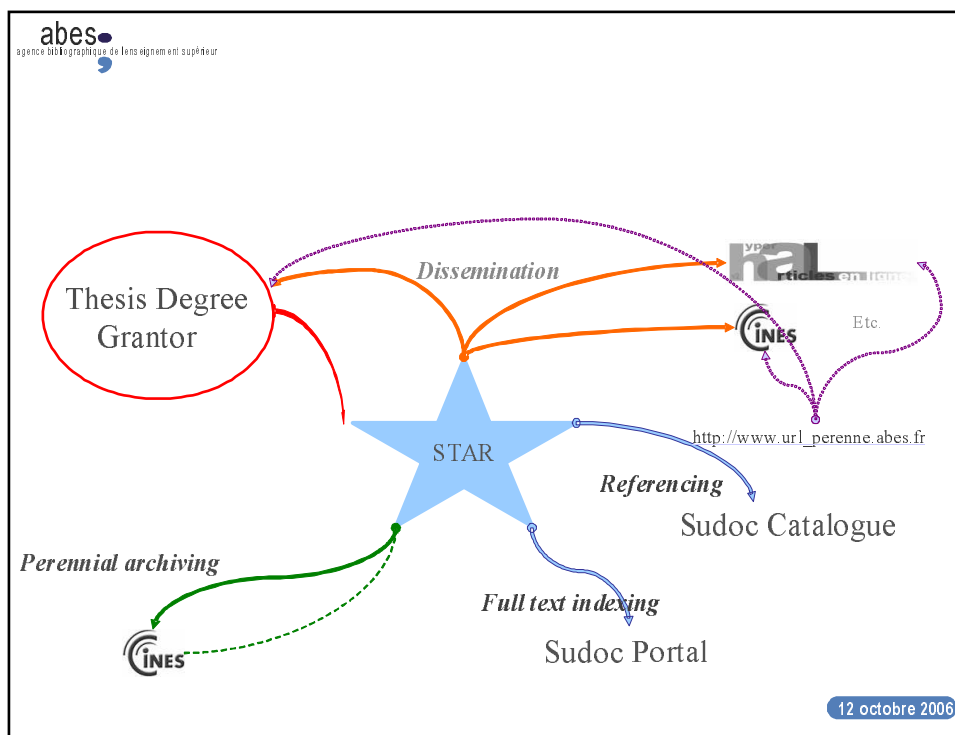


Figure 1

Source : adapted from: STAR – introduction générale / ABES. Journée de lancement STAR, Montpellier, 12 octobre 2006. (STAR 2007)

Based on the OAIS concept, the CINES developed a platform for the long-term preservation of French ETDs called "PAC" with three different functionalities for the management of deposit, archiving and access.

The new ETD program started at the end of 2006. Since then, the input of electronic theses into the Sudoc increased steadily and attains nearly 2,500 ETDs in March 2007. Nevertheless, it is too early to make assumptions about the results.

Conclusion: some practical tips to search and order a French thesis

From a clearly structured network in the 80s and 90s with defined roles, actors and services, the French theses landscape has changed into a heterogeneous mixture of national structures and local initiatives with centralized tools emerging again (STAR, TEF). This may be characteristic for a transitional period from a traditional "print circuit" to a networked digital library of ETDs. In the meantime, searching for French theses has to adopt a double strategy, based on an interrogation of the academic union catalogue Sudoc and a web-based search in ETD archives and repositories.

How to find a thesis in the Sudoc catalogue:

Choose the "Extended Search" interface.

De-select all types of publication except for theses.

Choose or select a subject.

Limit the publication year or range.

Add keywords with the index "subject words".

For formal information select the index "thesis note". This index contains formal information about the type of theses, the domain, the university and the date.

Each bibliographic record in Sudoc is linked to a holding record that lists the university libraries in possession of the document, with details on loan/copy conditions (PEB*).

In some special cases it is difficult or impossible to obtain a thesis referenced in the Sudoc:

(1) Confidential theses are referenced in the databases or university catalogues, but are not available.

The principal reasons for confidentiality are:

- The research has been conducted on a subject where patents have been submitted.
- The author plans to publish his work commercially.

If the confidentiality is time-limited, the document becomes available after this period.

(2) The jury/commission may ask the candidate to revise parts of his thesis. If this isn't done then the thesis may not be disseminated officially and be excluded from microform reproduction. Even if it could eventually be retrieved from a personal website, its scientific value should be considered with prudence.

(3) "Thèses d'exercice" in medicine normally are not reproduced in microform. They are available at the student's university and at the central library for medicine in Paris (BIUM*) where they can be retrieved through the BIUM catalogue.

Print copies from French theses can also be ordered directly via the INIST document supply service.

The Lille ANRT offers a service called "Thèse à la carte" where theses can be searched by subject or domain and ordered in book format; presently, the ANRT catalogue contains roughly 4500 theses.

Even if the Sudoc catalogue remains the point of access to all French theses in print format, ETDs should be searched in the different local and networked archives and databases to obtain full text access, since the Sudoc still offers a rather small number of records with hyperlinks to documents.

The search for a French ETD can start in some digital libraries or portals that offer updated selections of web links to repositories and archives.

Web links to ETD archive information:

Agence bibliographique de l'enseignement supérieur, Thèses [Online]. –
<<http://www.abes.fr/abes/DesktopDefault.aspx?Loupe=Moin>>

Bibliothèque Nationale de France, Thèses francophones en ligne [Online]. -
<http://signets.bnf.fr/html/categories/c_011theses_fra_ligne.html>

Ecole nationale supérieure des sciences de l'information et des bibliothèques (ENSSIB), Sibel. Thèses [Online]. -
<<http://sibel.enssib.fr/index.php?m=c&c=479>>

Maison des Sciences de l'Homme –Alpes, Thèses [Online]. -
<<http://www.msh-alpes.prd.fr/veille/actualitercherche.htm>>

Ministère de l'Education nationale, de l'Enseignement supérieur et de la Recherche, Thèses en ligne [Online]. -
<<http://www.educnet.education.fr/dossier/rechercher/these1.htm>>

Another way is to search directly in the ETD archives (see appendix) or on the universities' websites and catalogues. Nevertheless, in spite of these initiatives and services, searching French ETDs still remains a more or less difficult task.

References

- Boudia, D. and Gómez de Regil, R.M. (2006), TEF : Metadata for French thesis and dissertations. – In: GL7 Seventh International Conference on Grey Literature : Open Access to Grey Resources, 5-6 December 2005 / compiled and edited by D.J. Farace and J. Frantzen; GreyNet, Grey Literature Network Service - Amsterdam : TextRelease, January 2006. – pp. 108-113
- Claerebout, M.-F. (2003), Grisemine, a digital library of grey university literature. – In: GL5 Conference Proceedings, Amsterdam, 4-5 December 2003. – pp.27-31
- Jolly, C. (2000), Rapport sur la diffusion électronique des thèses [Online]. Ministère de l'éducation nationale. Direction de l'enseignement supérieur. Sous-direction des bibliothèques et de la documentation. –
<<http://www.sup.adc.education.fr/bib/>>
- Mermet, J.-M., Prudhomme, B., Joly, M., Pinon, J.-M. (1998), CITHER: un modèle de bibliothèque numérique de mémoires de thèses. - In: Document numérique. -Vol. 2, no. 3-4 (1998), pp.51-72.
- Ministère de l'Enseignement Supérieur et de la Recherche (1985), Arrêté du 25 septembre 1985 relatif aux modalités de dépôt, signalement et reproduction des thèses soutenues ou travaux présentés en soutenance en vue du doctorat [Online] –
<<http://www.sup.adc.education.fr/bib/Acti/These/textregl.htm#25sept>>
- Ministère de l'Enseignement Supérieur et de la Recherche (1994), Le signalement et la valorisation de la thèse. Guide du candidat au doctorat.
- Ministère de l'Education Nationale & Ministère de la Recherche (2000), Circulaire du 21 septembre 2000: Diffusion électronique des thèses. Circulaire no. 2000-149 [Online]. - In: Bulletin Officiel de l'Education Nationale. -No. 34, 28 September 2000. -
<<http://www.education.gouv.fr/bo/2000/34/sup.htm>>
- Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche (2005). Circulaire n° 05-094 du 29 mars 2005 sur le dépôt, signalement, diffusion et archivage des thèses sous forme électronique. [Online]. - Available at :
<<http://www.sup.adc.education.fr/bib/Acti/These/circulaire.rtf>>
- Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche (2006). 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 [Online]. - In: Journal Officiel n° 195 du 24 août 2006, page 12471. – Available at :
<<http://www.legifrance.gouv.fr/WAspad/UnTexteDeJorf?numjo=MENS0602085A>>

- Nicolas, Y. (2006). TEF : métadonnées pour les thèses électroniques françaises [Online]. – In : ETD 2006, The 9th International Symposium on Electronic Theses and Dissertations, June 7 - 10, 2006, Quebec City, Quebec, Canada. – Available at :
<http://www6.bibl.ulaval.ca:8080/etd2006/pages/papers/SP6_Yann_NICOLAS.pdf>
- Okret, C. (1998), Un élément de la rénovation du dispositif de diffusion des thèses françaises : un serveur national de thèses numérisées. In: Lettre d'information du CNUC. -No. 67 (1998).
- OST - Observatoire des Sciences et Techniques (2002), Indicateurs bibliométriques des institutions publiques de recherche. Année 1997 – hors sciences humaines et sociales. Essai méthodologique [Online]. –
<<http://www.obs-ost.fr>>
- Six&Dix (2004), Etude de la mise en œuvre du dispositif national de diffusion des thèses par voie électronique dans les Etablissements d'enseignement supérieur et de recherche. Note de synthèse [Online]. –
<<http://www.sup.adc.education.fr/bib/Acti/These/rapportV5.doc>>
- STAR (2007). Carrefour des thèses électroniques françaises [Online]. - Available at :
<<http://www.abes.fr/abes/page,428,star.html>>
- TEF (2007). Métadonnées des thèses françaises [Online]. - Available at :
<<http://www.abes.fr/abes/documents/tef/index.html>>
- Zhang, Y., Lee, K. and You, B.-J. (2001), Usage patterns of an electronic theses and dissertations system. –In: Online Information Review. -Vol. 25, no. 6 (2001), pp. 370-377.

Glossary

ABES: Agence Bibliographique de l'Enseignement Supérieur (operating agent of the French academic union catalogue and ILL system): <http://www.abes.fr>

AFNOR: Association Française de Normalisation (French standardisation organisation): <http://www.afnor.fr/>

ANRT: Atelier National de Reproduction des Thèses (national service for the reproduction of theses). The Lille ANRT is hosted at the Charles de Gaulle university at Lille-3: <http://www.anrtheses.com.fr>

Article@INIST: Online database with over 8 million references of articles and monographs for document supply: <http://services.inist.fr/public/eng/consit.htm>

BIUM: Bibliothèque Interuniversitaire de Médecine (largest medical library in France and most important supplier in the academic sector): <http://www.bium.univ-paris5.fr/>

BNF: Bibliothèque Nationale de France (French national library): <http://www.bnf.fr/>

CCSD: Centre pour la Communication Scientifique Directe (CNRS Centre for Direct Scientific Communication): <http://ccsd.cnrs.fr/>

CINES: Centre Informatique National de l'Enseignement Supérieur (national academic computer and data-processing centre, called CNUC until 1999): <http://www.cines.fr/>

CNRS: Centre National de la Recherche Scientifique (French National Research Organisation): <http://www.cnrs.fr/>

HAL: HAL - Hyper Article en Ligne (HAL - Hyper Article on Line): <http://hal.archives-ouvertes.fr/index.php?langue=en&halsid=13aa66ff4f10ccfb0298c0d5f4ef2860>

INIST: Institut de l'Information Scientifique et Technique (CNRS institute for scientific and technical information): <http://www.inist.fr>

INRIA: French national institute for research in computer science and control: <http://www.inria.fr>

INSA: Institut National des Sciences Appliquées de Lyon (one of the top French engineering universities): <http://www.insa-lyon.fr/>

MathDoc: French network for documentation in mathematics and server for the management of ETDs run by the university of Grenoble-1 and the CNRS: <http://math-doc.ujf-grenoble.fr/Theses>

OST: Observatoire des Sciences et des Techniques (Observatory for Science and Technology, production of quantitative indicators for S&T activity): <http://www.obs-ost.fr/fr/>

PEB: Prêt entre bibliothèques (French academic interlibrary loan)

RAMEAU: Répertoire d'autorité-matière encyclopédique et alphabétique unifié (academic subject headings authority list): <http://rameau.bnf.fr>

STAR: Signalement des thèses, archivage et recherche (referencing, archiving and retrieval of ETDs): <http://www.abes.fr/abes/page,428,star.html>

TEF : Thèses Electroniques Françaises (French Electronic Theses) : <http://www.abes.fr/abes/documents/tef/index.html>

Sudoc: Système Universitaire de Documentation (academic union catalogue of serials and monographs): <http://www.sudoc.abes.fr/>

All websites visited in January 2007

Appendix

The following sheets contain for each ETD archive four sets of data:

- General information: content, domains, coverage (anteriority) and number of ETDs.
- Technical information: software and formats.
- Archiving: self-archiving and legal aspects.
- Institutional information: institution, website, contact, complementary information.

All data were collected in January 2007.

CITHER

General information

| | |
|------------------|---|
| URL | http://csidoc.insa-lyon.fr/these/ |
| Content | Theses from INSA |
| Type | Institutional |
| Language | French |
| Domains | Engineering sciences |
| Input | INSA Lyon |
| First release | 1997 |
| Number of theses | 571 |
| First thesis | 1995 |

Technical information

| | |
|-------------------|---|
| Platform/software | CEN (DELPHI) |
| Deposit format | Word, Latex, PostScript, odt (OpenOffice.org) |
| Viewing format | PDF |
| Storage format | Word, PDF, odt (OpenOffice.org) |

Archiving

| | |
|-------------------|--|
| Self-archiving | No |
| Dissemination | Authorization by author and university |
| Removal of thesis | Allowed |

Institutional information

| | |
|-------------|---|
| Institution | Centre de documentation scientifique Doc'INSA |
| Website | http://docinsa.insa-lyon.fr/ |
| Contact | cither@insa-lyon.fr |

CYBERTHESES

General information

| | |
|------------------|---|
| URL | http://cyberdocs.univ-lyon2.fr/ |
| Content | ETDs from Cyberthèses members |
| Type | Collaborative |
| Language | Website in French and in English |
| Domains | Multidisciplinary |
| Input | Universities of Lyon-2, Montreal, Chili and several institutions from other countries |
| First release | 1998 |
| Number of theses | 883 (Lyon 2) 17130 ETDs from using Cyberthèses – 79 archives worldwide |
| First thesis | 1996 |

Technical information

| | |
|-------------------|--|
| Platform/software | Cyberdocs (platformGPL), open source software |
| Editorial chain | Conversion of structured word documents in XML with the TEI Lite DTD |
| Metadata format | Dublin Core, Electronic Theses and Dissertations (ETD-MS) |
| OAI-PMH | Yes |
| Viewing format | HTML, XHTML, PDF |
| Storage format | XML |
| Other | RSS 2.0 |

Archiving

| | |
|----------------|---|
| Self-archiving | No |
| Dissemination | Creative Commons licences, University agreement requested |

Institutional information

| | |
|------------------|--|
| Institutions | University of Montreal University of Lyon-2 |
| Websites | http://www.umontreal.ca/ http://www.univ-lyon2.fr/ |
| More information | CyberThèses, Agence Intergouvernementale de la Francophonie, FFI, Cyberdocs: structure and publish electronic documents [Online]. - http://sourcesup.cru.fr/cybertheses/ University of Lyon 2, Cyberthèses [Online]. - http://theses.univ-lyon2.fr/ |
| Contact | Service Général des Publications : edition.electronique@univ-lyon2.fr |

IRIS

General information

| | |
|------------------|---|
| URL | http://iris.univ-lille1.fr/ Opening soon |
| Content | Theses from Lille 1 |
| Type | Institutional, multi-type |
| Language | French |
| Domains | Sciences, technology and social sciences |
| Input | University of Lille-1 |
| First release | 2001 |
| Number of theses | About 500 |
| First thesis | 1936 |
| Others documents | Conference papers, digitized works in history of science |
| Total documents | 750 |

Technical information

| | |
|-------------------|-----------------------------------|
| Platform/software | Dspace 1.3.2 |
| Metadata format | Dublin Core and specific elements |
| Deposit format | PDF |
| Viewing format | PDF |
| Storage format | PDF |

Archiving

| | |
|-------------------|--|
| Self-archiving | No |
| Dissemination | Authorization by author and university |
| Removal of thesis | Allowed |

Institutional information

| | |
|-------------|---|
| Institution | University of Lille-1 |
| Website | http://www.univ-lille1.fr/bustl |
| Contact | iris@univ-lille1.fr |

INRIA and HAL-INRIA

General information

| | |
|------------------|---|
| URL | http://www.inria.fr/rrrt/index.fr.html |
| Content | Research reports, technical reports, thesis edited by INRIA |
| Type | Institutional, multi-type |
| Language | English, French |
| Domains | Communication and information science, computer science, automation |
| Input | INRIA |
| Number of theses | INRIA : about 600 (1985-2005) TEL : about 320 |
| First thesis | 1985 |

Technical information

| | |
|---------|-------------|
| Formats | See TEL-HAL |
|---------|-------------|

Archiving

| | |
|-------------------|----------------------------|
| Self-archiving | Self-archiving via TEL-HAL |
| Dissemination | See TEL-HAL |
| Removal of thesis | Allowed |

Institutional information

| | |
|------------------|--|
| Institution | INRIA |
| Website | http://www.inria.fr/index.fr.html |
| More information | INRIA, Annual report 2005 [Online]. - < http://www.inria.fr/inria/rapportannuel/ran.fr.html > |
| Contact | webmaster@inria.fr |

MATHDOC

General information

| | |
|------------------|---|
| URL | http://math-doc.ujf-grenoble.fr/Theses |
| Content | Online access to resources in mathematics |
| Type | Collaborative, domain-specific |
| Language | Website in English and French, documents in different languages |
| Domain | Mathematics |
| Input | Universities (institutes of mathematics), CCSD |
| First release | 1998 |
| Number of theses | 1000 (full text and abstracts) |
| First thesis | 1990 |
| Others documents | Preprints, journals, books, digital library of old documents |

Technical information

| | |
|-------------------|--|
| Platform/software | Harvest, OpenResolver (Open URL), open source software |
| Formats | See TEL-HAL |

Archiving

| | |
|-------------------|----------------------------|
| Self-archiving | Self-archiving via TEL-HAL |
| Dissemination | See TEL-HAL |
| Removal of thesis | Allowed |

Institutional information

| | |
|------------------|---|
| Institution | Cellule MathDoc CNRS University of Grenoble-1 |
| Website | http://www-mathdoc.ujf-grenoble.fr/ |
| More information | Cellule MathDoc, Rapport d'activité 2002-2005, comité de pilotage et d'évaluation:17 janvier 2006 [Online]. - < http://www.mathdoc.fr/Publications/Rapports/UMS5638_RA.pdf > INIST-CNRS, La Cellule MathDoc [Online]. - < http://www.inist.fr/openaccess/article.php3?id_article=30&var_recherche=mathdoc > |
| Contact | accueil@mathdoc.ujf-grenoble.fr |

PASTEL

General information

| | |
|------------------|---|
| URL | http://pastel.paristech.org/ |
| Content | Theses from ParisTech |
| Type | Institutional, collaborative |
| Language | French |
| Domains | Engineering sciences, business and management, information sciences |
| Input | Paris Institute of Technology includes 11 engineering schools |
| First release | 2002 |
| Number of theses | 968 |
| First thesis | 1987 |

Technical information

| | |
|-------------------|-----------------------------------|
| Platform/software | Eprint 2.2.1 |
| OAI-PMH | Yes |
| Metadata format | Dublin Core and specific elements |
| Deposit format | PDF, PostScript, MS Word... |
| Viewing format | PDF, PostScript ... |
| Storage format | PDF, Text (ASCII)... |

Archiving

| | |
|-------------------|--|
| Self-archiving | Yes |
| Dissemination | Authorization by author and university |
| Removal of thesis | Allowed |

Institutional information

| | |
|-------------|---|
| Institution | ParisTech |
| Website | http://www.paristech.org//index_en.php |
| Contact | admin@rilk.com |

TEL - HAL

General information

| | |
|------------------|--|
| URL | http://tel.ccsd.cnrs.fr |
| Content | Multidisciplinary theses server |
| Type | Collaborative |
| Language | Website in English, German and French, theses in different languages |
| Domains | Multidisciplinary but mostly physics, mathematics and engineering sciences |
| Institutions | French universities, universities from other countries |
| First release | 2001 |
| Number of theses | 6028 |
| First thesis | 1940 |

Technical information

| | |
|-------------------|--|
| Platform/software | HAL 2.0 |
| Editorial chain | No |
| OAI-PMH | Yes |
| Metadata format | DC and specific scheme |
| Deposit format | (La)TEX, RTF, Word, PostScript, EPS, JPEG, HTML, PDF |
| Viewing format | (La)TEX, Word, HTML, PostScript, PDF, XML |
| Storage format | (La)TEX, Word, XML |
| Other | RSS 2.0 |

Archiving

| | |
|-------------------|---|
| Self-archiving | Yes |
| Dissemination | Authorization by author, often also by university |
| Removal of thesis | Allowed |

Institutional information

| | |
|-------------|--|
| Institution | CCSD Cellule MathDoc |
| Website | http://www.ccsd.cnrs.fr http://www-mathdoc.ujf-grenoble.fr/ |
| Contact | tel.support@ccsd.cnrs.fr |