

**Processus de diffusion des mémoires de Master.
Littérature grise et accessibilité : un enjeu majeur dans
le domaine de la Recherche. Étude de cas : projet
MemorySID
Sylvain Vanacker**

► **To cite this version:**

Sylvain Vanacker. Processus de diffusion des mémoires de Master. Littérature grise et accessibilité : un enjeu majeur dans le domaine de la Recherche. Étude de cas : projet MemorySID. Sciences de l'information et de la communication. 2017. <dumas-01558312>

HAL Id: dumas-01558312

<https://dumas.ccsd.cnrs.fr/dumas-01558312>

Submitted on 7 Jul 2017

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.

Sylvain VANACKER
Master en Sciences de l'Information et de la Documentation
Première année

Mémoire de stage
Mission effectuée du 10 avril au 10 juillet 2017

à
L'Atelier National de Reproduction des Thèses
Villeneuve d'Ascq, Université Charles-de-Gaulle – Lille 3

**Processus de diffusion des mémoires de Master – Littérature Grise
et accessibilité, un enjeu majeur dans le domaine de la Recherche**

Étude de cas : Projet MemorySID

Sous la direction de :
M. J. SCHÖPFEL (Tuteur Professionnel)
M. E. KERGOSIEN (Tuteur Universitaire)

Soutenu le 30 Juin 2017 à l'UFR DECCID-SID
Université de Lille 3, Sciences Humaines et Sociales
BP 60 149, 59 650 Villeneuve d'Ascq Cedex

Année Universitaire 2016/2017



Remerciements

« Tout d'abord, je tiens à remercier Monsieur Joachim SCHÖPFEL pour m'avoir accordé sa confiance en me proposant d'effectuer mon stage au sein de l'Atelier National de Reproduction des Thèses, mais aussi pour ses nombreux conseils et son soutien. Je souhaite aussi remercier le personnel de l'Atelier pour son accueil ainsi que son écoute et sa disponibilité tout au long du stage. Je remercie, bien évidemment, Monsieur Eric KERGOSIEN pour m'avoir guidé lors de l'élaboration de ce mémoire mais aussi pour sa disponibilité.

Enfin, de manière plus personnelle, je remercie sincèrement ma famille qui m'a toujours soutenu au quotidien et au cours de mon cursus scolaire/universitaire. J'ai aussi une pensée particulière pour Marielle BACCARO puisque sans elle je n'aurais probablement pas entamé une poursuite d'étude dans un domaine qui me plaît qui plus est.

J'ai sûrement oublié certaines personnes mais je tiens sincèrement à remercier toutes celles et ceux qui ont contribué, de près ou de loin, à l'élaboration de ce mémoire. »

Table des matières

Introduction.....	1
I.L'ANRT, un service dynamique dédié à la diffusion de la recherche universitaire.....	3
1.Bref historique.....	3
2.Mission, services et prestations.....	3
a)La préservation et la diffusion.....	3
b)De multiples services et prestations.....	4
3.Personnel et locaux.....	4
4.Contenu, équipement et partenaires.....	5
a)Un équipement de pointe.....	5
b)De nombreux partenariats.....	5
II.Le projet MemorySID.....	7
1.Ambitions du projet.....	7
2.Ma mission.....	7
a)Vérification et correction des méta-données dans Nuxeo.....	7
b)Ré-numérisation des mémoires et rectification.....	9
c)Diffusion des mémoires dans l'ENT de Lille 3 et en libre accès.....	9
d)Tâches en parallèles.....	11
3.Problèmes rencontrés et choix entrepris.....	13
a)Problème de droits et d'auteur.....	13
b)Création d'un portail spécifique au Master SID.....	14
4.Méthodes de travail.....	14
a)Moyens de communication.....	14
b)Organisation personnelle.....	14
III.Littérature Grise, caractéristiques et État de l'art.....	16
1.Qu'est-ce que la Littérature Grise ?.....	16
a)Définitions et notions.....	16
b)La question de l'accessibilité.....	17
c)L'accès via les archives ouvertes et ses spécificités.....	18
2.État de l'art dans l'environnement des mémoires de Master.....	19
a)MemSIC et DUMAS : Les plates-formes de dépôt.....	19
b)OpenDOAR et ROAR : Activité de dépôt.....	20
c)ROARMAP : Politique de libre accès.....	21
d)Trois projets de stage d'étudiants pour une prise de conscience commune.....	21
e)Finalité de ces analyses.....	22
IV.Le cas du projet MemorySID : Processus de traitement des mémoires de Master.....	24
1.Bref récapitulatif et vue d'ensemble.....	24
2.L'accès aux mémoires, un enjeu majeur pour la Recherche ?.....	25
a)L'accès à une information de qualité et contexte d'utilisation.....	25
b)L'optimisation de l'organisation interne.....	26
c)Facilité la mise à disposition où comment rendre le travail du chercheur plus agréable.....	27
3.Diffusion et accessibilité, deux notions pour un même objectif.....	28
a)Un objectif commun.....	28
b)L'importance du référencement.....	28
Conclusion.....	30
Bibliographie.....	32
Webographie.....	32
Lexique.....	33
Sigles.....	34
Résumé.....	35

Introduction

Au cours de la première année en Sciences de l'Information et de la Documentation, il est demandé aux étudiants d'effectuer un stage d'au moins six semaines au sein d'une entreprise en corrélation avec les aspirations de la formation suivie. Après avoir fait quelques demandes auprès d'organismes en lien avec mon projet professionnel (touchant le domaine de la Rédaction Web) et qui n'ont pas abouti, c'est avec surprise que j'ai reçu une proposition de stage de la part de Monsieur SCHÖPFEL. Celle-ci concernait le projet MemorySID (projet, par ailleurs, pour lequel nous avons dû élaborer un cahier des charges fictif pour le dossier d'évaluation du cours de Numérisation). De manière générale, le stage avait pour ambition la numérisation, le stockage, la conservation et la diffusion des mémoires de stage d'étudiants en Master SID via l'ENT de Lille 3 mais aussi en libre accès par l'intermédiaire d'une archive ouverte.

Pour mener à bien ce projet déjà entamé par l'équipe de l'Atelier, j'ai travaillé au sein même de l'ANRT du 10 avril 2017 au 10 juillet 2017.

Pour présenter mes missions dans leurs grandes lignes, les objectifs principaux du stage étaient :

- La création, la gestion, le contrôle et la mise à jour des méta-données des mémoires de stage sur la plate-forme en ligne Nuxeo.
- La numérisation des mémoires (ou d'autres documents).
- La mise en œuvre de la publication des mémoires dans l'ENT de Lille 3.
- L'étude de la faisabilité de l'option d'une publication sur Internet.
- La diffusion des mémoires dans une archive ouverte de type HAL.

Au fil de la progression du projet, d'autres tâches m'ont été attribuées telles que, par exemple, la prise de connaissance des projets de numérisation de trois étudiants. Ou encore, la réception des Procès Verbaux des mémoires de stage d'étudiants en Master SID des années précédentes, la création d'une maquette du futur portail accueillant la collection des mémoires SID, et bien d'autres choses sur lesquelles je reviendrai au cours de ce mémoire.

Les compétences visées et à développer, spécifiées dans la proposition de stage, étaient :

- La prise en main d'un Système de Gestion Électronique de Documents (Nuxeo).
- La prise de connaissance des méthodes d'insertion des méta-données (Dublin Core).
- L'apprentissage des techniques d'indexation, de numérisation et de gestion terminologique, entre autres (XML).

Les principales aptitudes requises pour intégrer ce projet étaient de faire preuve d'un sens de l'organisation, d'être rigoureux, et de disposer d'une aisance relationnelle (Travail en équipe).

Au cours du stage, les étudiants doivent incorporer une problématique en lien avec la mission effectuée. Dans mon cas, il me semblait évident de faire de la Littérature Grise le fil conducteur de mon mémoire. Après m'être concerté avec J. SCHÖPFEL puis E. KERGOSIEN, j'ai donc poursuivi dans cette dynamique en orientant mon étude sur le cas du projet MemorySID. Le lien entre ma problématique et le sujet de mon stage fût des plus évocateurs.

Monsieur SCHÖPFEL et moi-même nous sommes intéressés à l'importance et l'impact que pouvait engendrer un accès aux mémoires de Master dans le secteur de la Recherche. Y avait-il un intérêt quelconque à établir l'ensemble des processus de traitement des mémoires dans l'optique d'en faciliter leur accessibilité et leur visibilité ? Un besoin informationnel était-il perceptible au sein du secteur de la Recherche concernant ce type de document ? Ces interrogations pouvaient alors aussi se poser à plus grande échelle, c'est-à-dire à l'ensemble des documents composant la Littérature Grise, dont les mémoires.

Avant de procéder à une réflexion, il est nécessaire d'agir de manière structurée en annonçant la construction du plan pour guider la lecture. Dans un premier temps, j'effectuerai une présentation de l'Atelier National de Reproduction des Thèses afin de situer l'environnement de travail dans lequel j'ai « baigné » tout au long du stage. Ensuite, je reviendrai en détail sur les spécificités du projet, les différentes tâches que j'ai menées au cours de l'avancée du projet MemorySID, ainsi que les différents problèmes auxquels il a fallu faire face. Puis, j'établirai un État de l'art et une analyse de l'espace dans lequel la Littérature Grise se met en place et se développe pour, enfin, aborder et étudier le cas du projet de stage afin de rendre compte d'une expérience de terrain et appuyer les constats émanant de mes recherches.

À noter : Pour des raisons de commodité, j'évoque les différentes tâches que j'ai effectuées lors de mon stage au passé malgré que ce dernier soit toujours en cours à l'heure où j'écris ces lignes. De plus, certaines phases de ma mission n'ont pas encore eu lieu, c'est pourquoi je préciserai cet état de fait dans les parties concernées constituant ce mémoire.

I. L'ANRT, un service dynamique dédié à la diffusion de la recherche universitaire

1. *Bref historique*

L'ANRT, soit l'Atelier National de Reproduction des Thèses, a été fondé en 1971 par le Ministère de l'Éducation Nationale. Il s'agit d'un organisme public français sous la tutelle du Département de l'Information Scientifique et Technique et Réseau Documentaire (DIST-RD). À ses début, l'Atelier effectuait la reproduction des thèses de doctorat. Par ailleurs, cette reproduction ne concernait que le secteur des « Lettres ». Le procédé d'impression utilisé était l'Offset permettant une souplesse et une capacité d'adaptation large selon une grande variété de produits. Dès 1984, sa méthode d'impression s'étend avec la reproduction sur microfiche permettant de stocker les documents numérisés sur un support de stockage beaucoup plus compact.

À noter que pendant longtemps il existait deux ateliers, l'un rattaché à l'Université de Lille 3 Charles-de-Gaulle et l'autre à l'Université de Grenoble. C'est au début de l'année 2011 que les deux organismes fusionnent et s'installent au sein du campus Pont de Bois de Villeneuve-d'Ascq (annexes 1 et 2).

2. *Mission, services et prestations*

a) *La préservation et la diffusion*

Sa mission principale est « *la préservation et la diffusion aux bibliothèques universitaires des thèses soutenues en France* » ainsi que « *la valorisation de la production scientifique de l'enseignement supérieur et de la recherche* ». De ce fait, l'ANRT « *reçoit les thèses imprimées, saisit les données bibliographiques, numérise les documents, produit une microfiche de haute qualité puis copie cette microfiche pour la diffusion auprès des Bibliothèques Universitaires et Services Communs de la Documentation* »¹. Elle reproduit, sous différents supports (imprimés, microfiches et électroniques) les thèses soutenues en France dans toutes les disciplines, et assure leur diffusion d'un point de vue institutionnelle mais aussi commerciale en France et à l'international (annexe 3).

Cette mission est possible grâce à la possession de deux catalogues par l'Atelier. Un catalogue général contenant l'ensemble des thèses reproduites depuis la création de l'ANRT (celles-ci sont uniquement disponibles pour les auteurs et les établissements publics. Par ailleurs, il faut savoir qu'il y a la possibilité de les imprimer pour une consultation en bibliothèques uniquement et à la demande

1 Site web institutionnel de l'ANRT - Lille 3, *Présentation de l'ANRT*, 2016

• <https://anrt.univ-lille3.fr/>

d'organismes institutionnels). Et un catalogue nommé « Thèses à la carte » composé d'environ 7800 titres commandables (ceux-ci étant disponibles en accord avec leurs auteurs).

b) De multiples services et prestations

L'organisme propose divers services et prestations (annexe 4) :

- Il peut fournir sur demande une version imprimée d'une thèse et/ou son fichier numérique.
- Il peut imprimer une thèse sous forme de livre à la demande d'un auteur (à des fins non-commerciales). Ce qui lui octroie donc la possibilité de proposer le service appelé « Thèses à la carte ».
- Il peut fournir une assistance aux doctorants dans la réalisation de leurs exemplaires de soutenance.
- Il peut imprimer des travaux scientifiques et prendre en charge des projets de numérisation de collections scientifiques afin de les diffuser sur des plate-formes ou des archives ouvertes (cette prestation ne figure pas dans sa mission nationale).

À cela s'ajoute l'implémentation d'une cellule de communication ayant pour objectif de répondre à toute demande personnalisée.

Pour accéder à ses services, il est possible de contacter le personnel sur place au sein des locaux de l'Atelier, par téléphone ou par mail afin de prendre rendez-vous. Il est de même possible de commander directement des thèses via son site internet, entre autres (annexe 5).

3. Personnel et locaux

L'équipe de l'ANRT est composée de quatre pôles spécifiques afin d'organiser et mettre en œuvre ses services (annexe 6):

- **Le Pôle Impression** – Impression (Noir&Blanc, couleur) et façonnage.
- **Le Pôle Numérisation** – Numérisation des thèses et des livres anciens et précieux.
- **Le Pôle Microfiche** – Codage, saisie des bordereaux, création des microfiches mères et duplication.
- **Le Pôle Diffusion** – Élaboration des contrats, facturation et communication.

D'autres éléments viennent compléter la structuration de l'Atelier afin d'organiser la gestion interne du service, tels que le Service expédition, la salle des serveurs informatiques, la Direction et les Ressources Humaines (annexes 7 et 8).

4. Contenu, équipement et partenaires

L'ANRT dispose donc d'un vaste catalogue contenant environ « 230 000 thèses de doctorat ». Le tout stocké dans ses locaux sous la forme de microfiches (dans des conditions de stockage adaptées). Elle bénéficie d'environ « 7600 licences non-exclusives »². Ce qui a permis la mise en place du service de « Thèses à la carte ».

a) Un équipement de pointe

Pour mener à bien sa mission, l'Atelier dispose d'un panel d'équipements adéquat (annexes 9 et 10) :

- Équipements de numérisation :
 - Trois numériseurs à livre ouvert (Suprascan Digibook 6002 RGB (i2S), Suprascan Quartz 400 BC25 (i2S), OS 15000 (Zeutschel)).
 - Cinq numériseurs de production (DR-X10C (Canon)).
 - Un numériseur de microfiches (OM 1500 (Zeutschel)).
- Équipements d'impression :
 - Deux chaînes de production de microfiches (Pour les Master (argentique) : Digi-fiche (Staude), Pour les copies (diazographique) : Datamaster II (Datagraphix)).
 - Deux chaînes d'impression (Press C6000 (Konica Minolta, couleurs), VarioPrint 6160 (Océ, NB)).
- Équipements divers pour façonnage (massicotage, reliure, ...).

b) De nombreux partenariats

À cela s'ajoute divers partenariats lui permettant d'affirmer son statut et de démontrer son savoir-faire (annexe 11) :

- Le Service Commun de la Documentation de Lille 3.
- L'Agence Nationale de la Recherche.
- Le Projet CHispa – e-man.
- L'Agence Bibliographique de l'Enseignement Supérieur.
- L'Université de Lille 1 Sciences et Technologies.
- L'IRIS – La Bibliothèque Numérique en Histoire des Sciences.
- Le Musée d'Histoire Naturelle de Lille.

En effet, l'Atelier National de Reproduction des Thèses est sollicité régulièrement par ces organismes mais aussi par des étudiants, doctorants, ... Les nombreux services qu'il offre lui permettent de prendre

² Site web professionnel de l'ANRT, l'ANRT, 2016

• <http://www.diffusiontheses.fr/content/4-anrt-lille-reproduction-theses>

part à divers projets tels que celui faisant l'objet de mon stage : MemorySID.

II. Le projet MemorySID

1. *Ambitions du projet*

L'objectif du projet (en partenariat avec le Département des Sciences de l'Information et de la Documentation, et le laboratoire de Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication) est multiple. Il se décompose en plusieurs étapes pour permettre une cohérence dans la démarche entreprise :

1. La dématérialisation de la gestion des mémoires de stage d'étudiants de Master.
2. La mise à disposition des mémoires dans Nuxeo via l'ENT.
3. Le développement d'un outil TDM pour faciliter la recherche d'un stage, en intégrant les méta-données des stages.
4. La diffusion des mémoires de Master en libre accès via une archive ouverte.
5. La création d'une bibliothèque numérique pour les formations SID.

Par ailleurs, l'on peut ajouter que la dématérialisation des mémoires contribuera à un gain de place au sein des locaux du Département SID. Un élément non-négligeable puisque ces documents s'accumulent de plus en plus au fil des années.

Par la suite, dès lors que le projet aura abouti, il pourra s'ouvrir à de plausibles finalités comme par exemple : Servir de modèle pour d'autres initiatives, aussi bien à Lille qu'ailleurs. Faire l'objet de publications et communications au niveau national et international en vue d'inspirer d'autres organismes. Les perspectives futures du projet seront abordées plus tard au cours de ce mémoire. Ses ambitions montrent l'intérêt et l'importance d'avoir une structuration rigoureuse dans le processus de traitement et de diffusion des mémoires (annexe 12).

2. *Ma mission*

Ma mission de stage s'insérait donc dans ce processus et concernait les phases 1, 2 et 4 soit la dématérialisation de la gestion des mémoires, la mise à disposition des mémoires dans Nuxeo via l'ENT de Lille 3 ainsi que la diffusion en libre accès par l'intermédiaire d'une archive ouverte.

a) *Vérification et correction des méta-données dans Nuxeo*

En effet, à mon arrivée dans les locaux de l'Atelier, l'ensemble des mémoires des étudiants (au nombre d'environ 740 et datant de 1996 à aujourd'hui) avait déjà été numérisé et déposé sur la plate-forme Nuxeo (annexe 13). Ma première tâche consistait alors à la finalisation de l'étape 1 du processus : la

vérification et la correction de la démarche de dématérialisation. Certaines méta-données avaient été complétées par les employés en charge de la numérisation et du dépôt des mémoires (Jérémy BERTHE et Jean-Philippe DELCOURT). Il m'était donc demandé de vérifier et compléter les méta-données de chaque mémoire le cas échéant. Monsieur SCHÖPFEL et moi-même nous étions mis d'accord sur la manière dont je devais remplir les champs. Nous avons donc mis en place une norme d'écriture afin que les méta-données soient identiques pour l'ensemble des mémoires. Par exemple, le champ « Source » devait toujours contenir le terme « Master » suivi du nom de la formation ainsi que de l'année de Master puis la date de soutenance « 2000-2001 ». Cela permettait d'obtenir une cohérence et une rigueur pour l'ensemble des méta-données. À noter que la liste des méta-données présente sur Nuxeo utilisait la norme du Dublin Core pour décrire chaque document (annexe 14).

Voici la liste des champs présents dans la plate-forme Nuxeo :

- Nature du document → Rapport de stage.
- Sujets → Sciences Humaines/Information.
- Droits → Confidentialité ou non.
- Source → Master / Année de formation / Date de soutenance.
- Couverture → Lieu où s'est effectué le stage.
- Date de création → Soit la date de dépôt (automatique).
- Date de dernière modification.
- Format → PDF.
- Langue → Français.
- Auteur → Nom et prénom de l'étudiant.
- Contributeur(s)
- Dernier contributeur
- Description du document → Résumé de quelques lignes.

Enfin, je devais ajouter une liste de tags permettant de décrire le contenu du mémoire à l'aide de mots-clés évocateurs.

Ma seconde tâche, s'inscrivant dans la continuité, était de m'assurer que chacun des mémoires avait été correctement numérisé. Par exemple, vérifier qu'il ne manquait aucune page ou que l'une d'entre-elles n'était pas numérisée à l'envers. Pour ce faire, je feuilletais une à une les pages des mémoires physiques ainsi que leur double numérique de manière synchrone. Par ailleurs, dans certains cas des mémoires étaient déposés en doublons sur Nuxeo. Il me fallait alors supprimer la version obsolète.

Ces deux missions communes m'étant attribuées se révélaient être une étape majeure dans l'application de mon stage. J'y ai accordé une grande partie de mon temps dans sa finalisation. Ce

processus de travail était assez éprouvant puisque je devais allier rigueur, précision ainsi qu'une forte concentration pour éviter de commettre des erreurs d'inattention telles que des coquilles lors de l'insertion de texte dans les champs ou l'omission de certains champs.

b) Ré-numérisation des mémoires et rectification

Étant donné que l'étape de numérisation des mémoires de stage avait déjà été traitée par J. BERTHE et J-P. DELCOURT, cette tâche n'a pas été incluse dans ma mission. Cependant, dès lors que j'apercevais une erreur de numérisation, je devais en faire part à mes deux collègues afin qu'ils puissent numériser de nouveau les mémoires concernés et les déposer sur Nuxeo. Je n'ai donc pas participé directement à cette démarche mais j'ai tout de même apporté ma pierre à l'édifice en indiquant la liste des mémoires à « ré-numériser ».

c) Diffusion des mémoires dans l'ENT de Lille 3 et en libre accès

La deuxième phase de ma mission se centrait sur la diffusion des mémoires de stage, aussi bien de manière interne qu'externe. Avant de procéder à cette étape, il nous a fallu nous demander s'il était possible de la mettre en œuvre en bonne et due forme. D'où l'étude de faisabilité d'une option de publication sur Internet. Évidemment, chaque mémoire était soumis à des droits de publication bien spécifiques. De ce fait, il demeurait impératif de nous renseigner quant aux droits associés à chacun des mémoires. Fort heureusement, le secrétariat du Département SID conservait l'ensemble des Procès Verbaux des soutenances de mémoire des étudiants (de 1996 à aujourd'hui). Il y était notifié la possibilité ou non de diffuser leurs mémoires et sous quelles conditions (intranet, en ligne, ...). Ce qui nous a permis de résoudre cette question de faisabilité puisque nous avons l'un des éléments essentiels à la mise en place du processus, à savoir l'autorisation ou non de publier certains mémoires. La voie nous était ouverte, il ne nous « restait plus qu'à » choisir la méthode à utiliser.

- **Diffusion dans l'ENT de Lille 3 :**

Tout d'abord, Monsieur SCHÖPFEL et moi-même devions trouver un endroit où publier les mémoires de stage dans l'ENT de Lille 3 (Environnement Numérique de Travail). Pour ce faire, nous avons parcouru l'arborescence du répertoire de Lille 3 sur Nuxeo (l'espace publication). Deux choix s'offraient à nous : La publication dans le répertoire « Bibliothèques » qui était vide, ou la publication dans le répertoire « Learning Center ». La première option m'a semblait être la meilleure puisque, le répertoire « Bibliothèques » étant vierge de toute publication, permettait de partir sur une base saine. De plus, il n'était pas nécessaire de créer un nouveau répertoire ou de modifier son contenu. Ce qui évitait un travail supplémentaire. La seule manipulation à effectuer aurait été l'import du répertoire contenant l'ensemble des mémoires de stage SID (situé dans l'espace de travail de Nuxeo) dans celui-ci. La seconde option était un peu plus délicate étant donné le contenu déjà présent dans le répertoire

« Learning Center ». Il aurait fallu modifier beaucoup plus de choses (le nom des répertoires, l'agencement général de l'arborescence, ...). Pour toutes les raisons évoquées dans l'option une, il m'a semblé évident et plus judicieux d'occulter cette seconde possibilité.

Nous avons donc pris rendez-vous avec Cécile MALLERET, bibliothécaire Responsable du Soutien à la Recherche (Service Commun de la Documentation), afin d'obtenir l'accord nous permettant de publier nos mémoires de Master dans le dossier « Bibliothèques ». De plus, nous avons pu donner de plus amples informations concernant l'objectif de notre projet. Une fois l'entretien terminé, il nous fallait attendre que le répertoire soit ajouté dans l'espace publication de Nuxeo par l'architecte du Système Informatique en charge de cette opération, Stéphane MAGNIES. Ensuite, il nous aurait été possible de publier notre collection dans l'ENT de Lille 3 via cet espace dédié.

Une fois ce répertoire ajouté, j'ai pu entamer le processus de publication des mémoires dans l'ENT de Lille 3 via Nuxeo. Pour ce faire, j'ai d'abord exporté une dizaine de mémoires pour tester l'opération et observer qu'il n'y avait aucun problème (perte de méta-données, perte du fichier PDF, ...). Fort heureusement, ce n'était pas le cas et j'ai pu publier l'ensemble de la collection SID (748 mémoires environs) de manière groupée dans l'espace publication. Puis j'ai procédé à la suppression des quelques mémoires ne devant pas être diffusés sur l'ENT, en aval. Il m'a fallu faire une recherche manuelle de ces mémoires sur l'espace Nuxeo en fonction de la liste que j'avais élaboré grâce aux PV.



Illustration 1: Le répertoire "Bibliothèques" contenant les mémoires SID publiés.

Exactement 37 mémoires ont été enlevés de l'espace publication puisque les étudiants concernés avaient interdit leurs diffusions en interne. L'espace publication contient donc 711 mémoires de stage d'étudiants en Master SID accessibles uniquement aux utilisateurs disposant d'un identifiant à l'ENT de Lille 3. Enfin, 748 mémoires de stage d'étudiants en Master SID sont conservés dans l'espace de travail de Nuxeo.

- **Diffusion en libre accès via une archive ouverte :**

Avant de procéder à une diffusion dans une archive ouverte, nous nous sommes interrogés quant à la

plate-forme à utiliser. Monsieur SCHÖPFEL s'orientait d'abord sur une diffusion via MemSIC. Mais il a rapidement constaté l'absence de popularité de cette plate-forme. De plus, les expériences passées l'ont conforté dans le choix d'utiliser une autre plate-forme. En effet, les années précédentes, des étudiants avaient fait la demande de publication de leurs mémoires. Étrangement, leur demande n'a jamais abouti et leurs mémoires n'ont jamais été contrôlés et publiés sur MemSIC.

Après avoir contacté Madame Agnès SOUCHON (Responsable du Centre de ressources de l'UFR Sciences humaines de l'université de Grenoble), nous avons décidé d'opter pour la plate-forme DUMAS (Dépôt Universitaire de Mémoires Après Soutenance) étant donné l'essor qu'elle connaît ainsi que son développement croissant. A. SOUCHON nous a proposé de créer un portail spécifique pour pouvoir publier notre collection composée de mémoires de stage d'étudiants en Master SID. Au préalable, je m'étais renseigné sur le nombre de dépôts total effectués sur DUMAS (annexes 15, 16 et 17), la manière dont il fallait procéder pour publier une collection, la structuration du portail, ou encore la méthode de publication mise en place par d'autres établissements déposant leur propre collection sur cette plate-forme (le portail du laboratoire GERiCO et le portail des universités de Lyon, pour être précis).

J. SCHÖPFEL a donc procédé à la complétion de la fiche de création d'un portail de collection DUMAS (annexe 18) pour que l'on puisse, ensuite, importer les mémoires de stage dans l'archive ouverte.

À ce jour, nous attendons la confirmation de la création de notre portail DUMAS, par le CCSD, pour pouvoir procéder à son agencement et à l'exportation de notre collection de mémoires.

d) Tâches en parallèles

Qui plus est, d'autres tâches sont venues se greffer dans le prolongement de ma mission principale.

- **Les différents types de droits :**

Au cours de l'insertion des méta-données sur Nuxeo, nous nous sommes interrogés quant aux dénominations choisies pour remplir le champs « Droits ». Pour obtenir plus d'informations et élaborer une liste de termes définitive et appropriée, j'ai pris contact avec Madame ALGARVIO (la référente du Master 1 SID) afin de récupérer l'ensemble des Procès Verbaux des soutenances ayant eu lieu de 1996 à 2016. Après les avoir consultés, j'ai pu définir trois catégories bien distinctes pour caractériser les droits de chaque mémoire et les trier.

Ceux-ci devaient être placés dans une seule de ces trois options :

- **ENT Lille 3** - Autorisation de diffusion sur l'ENT de Lille 3 uniquement.
- **En libre accès** - Autorisation de diffusion en libre accès sur Internet et l'ENT de Lille 3.

- **Confidentiel** - Interdiction de diffusion sur l'ENT de Lille 3 et en libre accès sur Internet.

- **Communication concernant la diffusion des mémoires sur l'ENT de Lille 3 :**

Après avoir déposé les mémoires de stage des étudiants de Master SID sur l'espace de publication de Nuxeo, j'ai procédé à la communication de l'opération. J'ai rédigé et publié un petit article puis un Tweet, via le site institutionnel de l'ANRT et son compte Twitter, pour spécifier la mise à disposition des mémoires SID auprès des utilisateurs de l'ENT de Lille 3. Aurélie HALIPRE, la Responsable Communication, m'a aidé dans cette tâche puisque je ne disposais pas des droits nécessaires pour accéder aux comptes (annexe 19).

- **Présentation de la collection et maquette sur DUMAS :**

De plus, étant donné que nous comptons diffuser les mémoires via un portail dédié sur DUMAS, il m'a été demandé de rédiger un texte de présentation de notre future collection mais aussi de structurer la mise en page de notre portail. L'objectif étant de rendre la page la plus attrayante et complète possible. Pour ce faire, j'ai constitué une petite liste des éléments à insérer dans la page d'accueil et j'ai élaboré une maquette pour visualiser le résultat attendu (annexe 20).

- **Prise de connaissance des projets d'étudiants :**

Enfin, Monsieur SCHÖPFEL m'avait demandé de contacter trois étudiants afin de prendre connaissance de leurs projets de stage. Chacun de ces projets de numérisation avait pour mots d'ordre la valorisation, l'indexation ou encore la diffusion d'un fonds documentaire :

- Anaïs TCHICAYA – Étudiante en M1 SID – Stagiaire à l'École Supérieure de Journalisme de Lille.
→ *"Pour ma part, j'ai fait de la numérisation de mémoire et d'une revue."*
- Élodie PLANCOT – Étudiante en DEUST 2 Métiers des bibliothèques et de la documentation – Stagiaire à l'École Nationale Supérieure des Arts et des Métiers de Lille.
→ *"Mon stage au sein de la bibliothèque de l'ENSAM (École Nationale Supérieure d'Arts et Métiers) à Lille, une école d'ingénieur principalement axée en automobile et énergies renouvelables, a pour objectif d'intégrer au fonds documentaire les travaux de fins d'études (des thèses professionnelles) de Mastères spécialisés. Ce travail s'est fait en plusieurs étapes : inventaire des travaux concernés; catalogage et indexation sur le Système Intégré de Gestion de Bibliothèque de la bibliothèque Flora; exemplarisation; et mise en valeur au sein de la bibliothèque."*
- Florian BOIZET – Étudiant en DEUST – Stagiaire à l'École Nationale Supérieure d'Architecture et de Paysage de Lille.

→ "La structure dans laquelle j'effectue mon stage réfléchissait à la possibilité de numériser des mémoires effectivement, et ma mission est de traiter ces mémoires pour les indexer et les rendre disponibles à la lecture pour les étudiants. Mais je ne les numérise pas, par manque de temps. Le nombre de documents à traiter est trop important pour commencer à les numériser, ils préfèrent attendre une occasion et que le retard ait un peu diminué. Ma mission de traitement concerne donc presque uniquement les mémoires papier, à l'exception des documents numériques directement fournis par les étudiants."

Évidemment, ces trois projets étaient en lien étroit avec le nôtre. Ce qui permettait de comparer et mettre en relation les ambitions visées de chacun.

Enfin, j'évoquerai les analyses et études faites concernant les portails de diffusion dans les parties dédiées à la Littérature Grise et aux mémoires d'étudiants. Cette tâche englobait une troisième phase majeure dans l'avancée de mon stage. Par ailleurs, la prise de connaissance de ces trois projets a contribué à cette dernière.

3. Problèmes rencontrés et choix entrepris

Au cours du projet, certains problèmes ont fait leurs apparitions. Bien entendu, la mise en œuvre d'une mission ne serait pas aussi « palpitante » sans la présence d'inévitables obstacles. Fort heureusement, Monsieur SCHÖPFEL et moi réussissions, dans une majorité des cas, à résoudre ces divers petits contre-temps.

a) Problème de droits et d'auteur...

Pour commencer, l'un des problèmes majeurs provenait de Nuxeo et, plus particulièrement, de la modification de certaines méta-données. En effet, le champs « Auteur » était rempli automatiquement par le nom et prénom de la personne déposant les mémoires. Dans notre cas, il s'agissait régulièrement de Jérémie BERTHE et Jean-Philippe DELCOURT (les deux employés de l'ANRT qui se sont occupés du dépôt des mémoires dans la plate-forme Nuxeo). Malheureusement, il était impossible de modifier ce champs rempli par défaut. Pour remédier à ce problème, nous avons parcouru de long en large la documentation technique. Nous avons contacté l'assistance technique de Nuxeo, ou encore un consultant en GED (Grégory CARLIN), ... J'ai consulté des forums en quête de solutions viables ou de « stratagèmes » techniques. Aucune de ces démarches n'a abouti. Cependant, lorsque je me suis créé un compte personnel sur Nuxeo, je me suis familiarisé avec les fonctionnalités proposées par la plate-forme. Je me suis aperçu qu'il était possible de créer ses propres méta-données et de les modifier à sa guise. Le problème étant que sur la version disponible via l'ENT de Lille 3, nous ne disposons pas des droits nécessaires pour effectuer cette manœuvre. La personne disposant des droits adéquats

n'était pas disponible et il nous était alors impossible d'entamer quoi que ce soit pour résoudre ce problème. Nous avons donc choisi une alternative en inscrivant le nom et le prénom de l'étudiant du mémoire de stage dans le champs « Droits ». De ce fait, nous n'avons pas spécifié les droits associés pour chaque mémoire.

En ce qui concerne l'étape de diffusion, et comme étant notifié précédemment, j'ai utilisé l'ensemble des Procès Verbaux récupérés au secrétariat afin de lister les mémoires pouvant être publiés ou non. Le travail de diffusion a été légèrement plus « lourd » puisque la méta-donnée « Droits » ne contenait pas l'information nécessaire. L'exportation des mémoires dans l'espace de publication de Nuxeo s'est faite de manière moins intuitive puisqu'il m'a fallu consulter la liste élaborée, régulièrement, afin de supprimer manuellement les mémoires confidentiels à l'aide de la barre de recherche de Nuxeo.

b) Création d'un portail spécifique au Master SID

Lors de notre entretien avec Madame Agnès SOUCHON, nous avons remarqué qu'un portail destiné à l'UFR Humanité avait déjà été créé sur DUMAS. De ce fait, il nous était impossible d'en créer un dédié à l'ensemble des formations en Sciences Humaines et Sociales et d'y inclure ses différentes composantes. Nous avons donc fait le choix de créer notre propre portail de collections pour le Département des Sciences de l'Information et de la Documentation puisqu'une concertation avec l'UFR Humanité n'était pas à l'ordre du jour et qu'aucun compromis n'a été trouvé dès lors.

4. Méthodes de travail

Bien entendu, nous nous devons d'avoir une méthodologie de travail construite et ordonnée pour mener à bien le projet tout au long de son application.

a) Moyens de communication

En ce qui concerne les moyens de communication, nous utilisons fréquemment la plate-forme Freedcamp (annexes 21 et 22) pour mettre à jour les tâches effectuées au sein du projet MemorySID. Lorsque nous avons des interrogations ou des informations à transmettre, nous le faisons par mail ou dans les locaux lorsque Monsieur SCHÖPFEL était présent. Nos échanges s'avéraient être des plus bénéfiques dans la coordination de nos « opérations ». Cela nous permettait de faire régulièrement des mises au point sur le travail accompli et sur ce qu'il restait à faire ou à mettre en place.

b) Organisation personnelle

De manière plus personnelle, j'organisais mon travail à l'aide de post-it, en écrivant les choses que je

devais faire (sur papier et sur un outil de traitement de texte), en consultant chaque jour les tâches qui m'étaient attribuées sur Freedcamp, etc...

Lors de l'insertion et la vérification des méta-données de chaque mémoire sur Nuxeo, je constituais des listes afin d'y noter :

- Les mémoires déposés en doublons à supprimer à la fin de cette tâche.
- Les mémoires à rectifier (problèmes de numérisation).
- Les mémoires disposant de beaucoup d'annotations (jugés inaptes à la diffusion).
- Les mémoires disposant de « Droits » particuliers.

J'ai donc supprimé la vingtaine de mémoires en doublons à la fin du processus de vérification. Puis j'ai consulté Jérémie BERTHE pour qu'il rectifie la dizaine de mémoires ayant des problèmes de numérisation et qu'il dépose les versions corrigées sur Nuxeo. Pour ce qui est des droits, je me suis plutôt servi des Procès Verbaux pour pouvoir préparer la diffusion des mémoires de stage. Enfin, au vu du nombre restreint de mémoires annotés (un poignée sur environ 740 mémoires), Monsieur SCHÖPFEL m'a conseillé de ne pas y prêter attention et de les inclure dans notre future collection.

Ma mission de vérification et de contrôle des méta-données s'est achevée au bout d'un mois environ. Suite à cela, Monsieur SCHÖPFEL m'a demandé d'effectuer des recherches à propos de la quantité de mémoires d'étudiants de Master publiés sur l'espace MemSIC. L'objectif étant d'étudier l'importance ou non de rendre disponible les mémoires de Master en libre accès. Cela permettait d'observer si leurs dépôts jouaient un rôle majeur dans le domaine de la Recherche. Chiffrer cette étude était d'une importance capitale pour appuyer ces recherches. De plus, j'ai récolté et comparé certains chiffres trouvés sur les sites OpenDOAR et ROAR (deux annuaires mondiaux recensant le nombre de dépôts académiques en libre accès) ainsi que décompté le nombre de dépôts de mémoires de Master uniquement sur le site ROARMAP. Je reviendrai en détail sur ces recherches un peu plus tard dans la suite de ce mémoire.

À cet instant, la phase de diffusion des mémoires en libre accès sur DUMAS n'a pas eu lieu. Pour anticiper cette future opération, j'ai gardé la liste des mémoires confidentiels que j'ai complétée afin de filtrer la future exportation. Le dépôt sur cet espace ne devrait donc pas poser de problèmes.

III. Littérature Grise, caractéristiques et État de l'art

Tout au long de la précédente partie composant mon mémoire de stage, j'ai développé la manière dont nous avons procédé pour indexer les mémoires de Master d'étudiants en Sciences de l'Information et de la Documentation via le projet MemorySID, mais aussi la démarche que nous avons utilisée pour les diffuser sur l'ENT et celle que nous comptons utiliser pour la publication sur DUMAS. Après quelques échanges avec Monsieur SCHÖPFEL, nous avons trouvé intéressant d'étudier l'impact et l'importance qu'engendre l'accès des mémoires de Master dans le domaine de la Recherche. À savoir, si son accessibilité demeure « utile » et efficiente dans ce domaine.

De ce fait, au cours de cette troisième partie, je me suis orienté dans une analyse plus générale en étudiant le secteur de la Littérature Grise pour ensuite recentrer cette analyse à plus petite échelle en m'axant sur le cas des mémoires de Master. Avant de s'interroger sur les processus de diffusion des mémoires de Master ainsi que sur l'intérêt de cette démarche, il est important d'évoquer le lien l'unissant à la Littérature Grise.

1. Qu'est-ce que la Littérature Grise ?

a) Définitions et notions

Ce que l'on appelle « Littérature Grise » se compose, notamment, de mémoires, de thèses ou encore de brevets.

Selon la définition de l'AFNOR (Association Française de Normalisation), la Littérature Grise est tout « *document dactylographié ou imprimé, souvent à caractère provisoire, reproduit et diffusé à un nombre d'exemplaires inférieur au millier, en dehors des circuits commerciaux de l'édition et de la diffusion* »³.

Lors de la troisième conférence internationale sur la Littérature Grise de 1997, une autre définition a été donnée. « *[Grey literature is] that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers* »⁴. Ce qui donne en français « *[La Littérature grise est] ce qui est produit par toutes les instances du gouvernement, de l'enseignement et la recherche publique, du commerce et de l'industrie, sous un format papier ou numérique, et qui n'est pas contrôlé par l'édition commerciale* ». Cette définition insiste sur l'universalité et l'omniprésence de l'information qui est contenue dans la Littérature Grise,

3 Wikipédia, *Définition AFNOR de la Littérature Grise*, 2016

• https://fr.wikipedia.org/wiki/Litt%C3%A9rature_grise

4 Joachim SCHÖPFEL, *Vers une nouvelle définition de la littérature grise*, 2012, Cahiers de la Documentation, 66 (3), p.14-24

• https://hal.archives-ouvertes.fr/sic_00794984/document

ainsi que sur sa difficulté à l'identifier et à y accéder par l'intermédiaire des plates-formes commerciales de l'édition classique. De ce fait, la Littérature Grise échappe quasiment à tout recensement malgré la présence de dépôt légal.

Puis, en 2004, lors de la 11^{ème} conférence internationale sur la Littérature Grise ayant eu lieu à New York, il a été précisé que la définition de 1997 incorporait les éditeurs. Cependant, la publication ne constitue pas l'activité principale du domaine de l'édition.

En 2012, lors du colloque *GL12, Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues*, la définition émise par le maître de conférence J. SCHÖPFEL montre bien que la définition donnée à la Littérature Grise tend à se préciser et, insiste sur la notion de droits associés à celle-ci : « *La littérature grise correspond à tout type de document produit par le gouvernement, l'administration, l'enseignement et la recherche, le commerce et l'industrie, en format papier ou numérique, protégé par les droits de propriété intellectuelle, de qualité suffisante pour être collecté et conservé par une bibliothèque ou une archive institutionnelle, et qui n'est pas contrôlé par l'édition commerciale* »⁵.

Nul doute que cette définition tente de se modeler au fil de l'évolution de l'environnement ambiant et plus particulièrement celui de l'édition. En effet, le web a totalement modifié ce marché. Il offre de nouveaux outils de production, de diffusion et d'évaluation de la Littérature Grise. Les comportements doivent s'adapter, aussi bien ceux des producteurs que des consommateurs. La Littérature Grise se meut et doit continuer à se mouvoir à l'intérieur de cette évolution numérique proposant de nouveaux outils (archivage, indexation, diffusion, ...). Elle ne peut rester « fixe », une adaptation de tous ces éléments et des acteurs est nécessaire.

b) La question de l'accessibilité

Déjà en 1984, cette notion d'accessibilité est une préoccupation première dans le secteur de la Recherche. Comme l'évoque Jacques KERIGUY à cette époque, « *la littérature grise est éphémère mais se révèle être une source de connaissance scientifique et technique sans cesse renouvelée* »⁶. Ce qui, par déduction, insiste bien sur la multiplication de la documentation dans ce domaine. Comme étant spécifié dans la partie « Définitions et notion » de la Littérature Grise, en 2004 l'aspect publication ne fait pas partie des fonctions principales de l'édition, ce qui accentue « l'éphémérité » du contenu de la Littérature Grise. En effet, en France elle n'est disponible qu'entre 5 et 10 ans. Cependant, elle est parfois conservée à long terme par le CINES (Centre Informatique National de l'Enseignement

5 Joachim SCHÖPFEL, *Vers une nouvelle définition de la littérature grise*, 2012, Cahiers de la Documentation, 66 (3), p.14-24

• https://hal.archives-ouvertes.fr/sic_00794984/document

6 Jacques KERIGUY, *L'accès à la littérature grise*, 1984, BBF n°2

• <http://bbf.enssib.fr/consulter/bbf-1984-02-0138-001>

Supérieur) afin de préserver tout le savoir dont elle dispose.

Aujourd'hui, la préoccupation première de la Littérature Grise porte toujours sur son accessibilité. En raison de son volume sans cesse croissant et de l'importance qualitative de l'information qu'elle véhicule, elle a incité de nombreux pays à mettre en œuvre des efforts particuliers pour en assurer la collecte, la description bibliographique, le signalement dans les bases de données et, bien évidemment, son accès. Le système de référencement est un axe primordial puisque la Littérature Grise est difficile à « retrouver ». C'est pourquoi d'autres alternatives sont proposées comme la mise à disposition d'archives ouvertes afin de remédier à ce problème. Le dépôt et la publication de l'ensemble de cette documentation doit alors se faire de manière plus simple et ordonnée.

À titre d'exemple, moi-même j'ai consulté des documents issus de la Littérature Grise pour mes recherches et pour élaborer ce mémoire. Pour y accéder, je ne suis pas de suite passé par des archives ouvertes, n'ayant pas connaissance de leur existence... C'est après avoir travaillé sur le sujet que j'ai découvert des plates-formes telles que MemSIC ou encore DUMAS. La présence de ce type de portails doit alors s'ancrer progressivement dans le paysage de la Littérature Grise pour devenir une constante dans la publication mais aussi dans la consultation. L'accessibilité doit passer par cette prise de conscience de l'existence de ces espaces.

c) L'accès via les archives ouvertes et ses spécificités

L'accès aux documents provenant de la Littérature Grise est rendue possible de par la présence et l'utilisation d'archives ouvertes. Celles-ci sont des plates-formes en ligne où sont disposés de nombreux documents en libre accès. Ici, l'on parlera d'accès via la voie verte (Green Road), soit l'auto-archivage des publications sur Internet (archives ouvertes telles que HAL, ou institutionnelles, réseaux sociaux et sites web), puisque le dépôt de documents s'effectue soit par le chercheur, soit par une personne habilitée à le faire, directement dans une archive ouverte. Chaque archive ouverte suit un protocole spécifique afin de conditionner l'échange de données lors du transfert des méta-données (OAI-PMH : The Open Archives Initiatives Protocol for Metadata Harvesting)⁷. L'objectif étant de « *faciliter le travail des moteurs de recherche spécialisés pour qu'ils offrent une réponse adéquate à une requête précise en interrogeant l'ensemble des serveurs des archives ouvertes* »⁸.

De manière plus générale, l'on retiendra deux grands « espaces » d'hébergement et d'accès à la Littérature Grise :

⁷ Site internet de l'Institut de l'information scientifique et technique, *Libre accès à l'information scientifique et technique*, 2016

• <http://openaccess.inist.fr/?+-Protocole-OAI-PMH->

⁸ Jacqueline NIVARD, *Les archives ouvertes*, La lettre de l'EHESS n°34

• <http://lettre.ehess.fr/988>

- **La Littérature Grise européenne :** Accessible gratuitement via la base OpenGrey (entre autres). Il s'agit d'un « *Système pour l'Information en Littérature Grise (en Europe) donnant un accès libre à 700 000 références bibliographiques de littérature grise papier produites en Europe et qui facilite l'accès aux documents via l'export et leur localisation. La littérature grise comprend des rapports techniques ou de recherche, des thèses de doctorat, des actes de congrès, des publications officielles, ... OpenGrey couvre les champs des sciences pure et appliquée, la technologie, les sciences biomédicales, l'économie, ainsi que les sciences humaines et sociales* »⁹.
- **La Littérature Grise « mondiale » :** Disponible via l'archive ouverte pluridisciplinaire HAL (Hyper Articles en Lignes) qui a été créée par le Centre pour la Communication Scientifique Directe (CCSD) du Centre National de la Recherche Scientifique (CNRS) en 2001. « *Elle 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* »¹⁰.

Les documents composant la Littérature Grise sont accessibles grâce au développement et l'usage des archives ouvertes dans le paysage informatique. Il devient alors plus aisé d'accéder à leur contenu en naviguant au sein des collections qui y sont regroupées.

2. **État de l'art dans l'environnement des mémoires de Master**

Au cours du projet MemorySID, nous nous sommes intéressés spécifiquement au cas des mémoires de Master. Comme étant vu dans la partie précédente, les mémoires composent cette littérature. Ceux-ci contiennent pléthore d'informations pouvant se révéler utile. Il nous a semblé intéressant d'étudier leurs « présences » au sein des plates-formes de diffusion pour observer s'il leur est accordé une importance particulière. De plus, cette interrogation corrélait avec le cadre du stage. C'est pourquoi, j'ai entamé un État de l'art concernant les dépôts et les consultations effectuées au sein des plates-formes d'archives ouvertes.

a) **MemSIC et DUMAS : Les plates-formes de dépôt**

La première étape a été de comptabiliser la quantité de mémoires d'étudiants de Master publiés, consultés et téléchargés sur l'espace MemSIC (annexes 23 et 24). Cette plate-forme est uniquement dédiée à l'archivage de mémoires de Master en Sciences de l'Information et de la Communication. 72

9 Site internet d'OpenGrey, *Présentation OpenGrey*, 2016

• <http://www.opengrey.eu/>

10 Site internet de HAL, *Présentation de l'archive ouverte*, 2016

• <https://hal.archives-ouvertes.fr/>

mémoires y ont donc été déposés. En 2010, un peu plus de 12000 consultations ont eu lieu. Une forte baisse s'est produite jusqu'en 2012 avec un peu moins de 2000 consultations. En 2016, une hausse est perceptible pour atteindre environ 6000 consultations dans l'année. À noter que les consultations peuvent s'effectuer à n'importe quelle période. De plus, énormément de mémoires ayant été soutenus en 2010 ont été publiés. L'on peut en déduire une certaine constante de la consultation tout au long des dépôts. Ce qui met en avant un certain intérêt éprouvé par les usagers concernant les mémoires de Master en SHS. Il en va de même pour le nombre de téléchargements effectués par les visiteurs puisqu'en 2010 un pic est atteint avec un peu plus de 160000 téléchargements.

Ensuite, j'ai consulté l'évolution des dépôts de mémoires sur DUMAS et HAL (retour aux annexes 15, 16 et 17). Les résultats observés sont sans appel, ces espaces de diffusion connaissent une augmentation considérable du nombre de dépôts de 2008 à 2016, soit un pic de 3720 mémoires archivés en 2016 pour DUMAS et une évolution de 50 000 dépôts effectués de 2005 à 2016 sur HAL. Nul doute que les chiffres de 2017 seront supérieurs.

On peut surtout conclure de cette analyse qu'une vague de dépôts de mémoires de Master s'est mise en place en 2010 sur l'espace MemSIC. Malheureusement, une forte baisse s'est constatée les années suivantes quant au dépôt de ce type de document. Cela a donc fortement confirmé notre choix de nous orienter vers une diffusion de notre collection de mémoires de stage sur la plate-forme DUMAS étant donné le manque d'intérêt suscité par MemSIC (et les précédents problèmes de dépôt connus par les étudiants, constatés par J. SCHÖPFEL).

b) OpenDOAR et ROAR : Activité de dépôt

La deuxième étape a été de récolter des chiffres sur deux annuaires mondiaux recensant le nombre de dépôts académiques en libre accès (OpenDOAR et ROAR). J'ai collecté diverses données chiffrées sous forme de diagrammes (à barres et circulaires) afin d'appuyer mes recherches. Les informations obtenues ont dévoilé par exemple, le nombre de mémoires publiés en France et dans le monde, la plate-forme de dépôt la plus utilisée, le type de mémoire fréquemment déposé ou encore la liste des disciplines concernées par les publications (annexes 25, 26, 27, 28, 29, 30).

Il m'a fallu trier et sélectionner les données me semblant être les plus pertinentes et en corrélations avec l'environnement étudié. Par ailleurs, les thèses et mémoires figuraient dans le même filtre de recherche. Il m'était donc impossible d'isoler ma demande uniquement sur le cas des mémoires. Cependant, les résultats obtenus m'ont permis de percevoir quelques constats. En effet, en France, un peu plus de la moitié des dépôts sont effectués sur l'espace HAL (58,8%) (annexe 31). Ce qui démontre bien l'engouement suscité par cette plate-forme et a affirmé notre choix d'utiliser DUMAS (base d'archive ouverte hébergée par HAL). Ajoutons que plus de la moitié des publications, en France,

concerne les mémoires et thèses (en deuxième position, la première place étant attribuée aux articles de journaux) (annexe 32). Enfin, les dépôts ayant pour discipline les Sciences Humaines et Sociales (Sciences Sociales Générales et Sciences de l'Information et de la Documentation) représentent une petite partie de la globalité des dépôts comptabilisés. La majorité étant détenue par les disciplines « économiques » et « environnementales » (annexe 33).

Concernant mes recherches via l'annuaire ROAR, il m'a été difficile de discerner les mémoires parmi les différentes dénominations proposées par le filtrage du moteur de recherche. Après avoir procédé par élimination, j'ai retenu le terme « e-Theses » étant donné qu'il englobait la publication de mémoires. Une liste des portails français était disponible et s'accompagnait de graphiques dévoilant « l'activité des dépôts » au cours des années. J'ai pu noter que, dans une majorité des cas, l'activité des dépôts est en constante évolution. La prise d'activité est variable selon les portails observés. Cependant, celle-ci débute généralement en 2004, 2006 et 2010 pour s'accroître jusqu'en 2016. (annexes 34 et 35). Ce qui démontre, une fois de plus, une prise de conscience dans la publication et l'archivage de thèses et de mémoires.

c) ROARMAP : Politique de libre accès

Ma dernière observation s'est orientée sur la Politique de libre accès mise en place en France. Malgré qu'elle ne se concentrait pas sur les mémoires de Master précisément, cela me permettait d'avoir une vision d'ensemble sur l'évolution de la politique menée jusqu'à aujourd'hui (annexe 36).

Le diagramme analysé mettait en avant le nombre de politiques adoptées en France de 2005 à 2017 (annexe 37 et 38). L'on peut noter que le nombre d'organisations de recherche concernées par le libre accès a augmenté (passant de 2 à 16). Par ailleurs, le nombre de politiques appliquées a aussi subi une hausse (évoluant d'environ 4 à 23). Ce qui confirme l'intérêt porté par les organismes de recherche pour la publication et la diffusion de documents en libre accès via des portails dédiés de type archives ouvertes.

d) Trois projets de stage d'étudiants pour une prise de conscience commune

Lors de ma prise de connaissance des projets des trois étudiants : Anaïs TCHICAYA, Élodie PLANCOT et Florian BOIZET, j'ai constaté qu'un intérêt était accordé à la valorisation, l'indexation et la diffusion du fonds documentaire faisant l'objet de leur stage. Ce qui confirme une prise de conscience de la part des universités. Celles-ci trouvent intéressant et important de valoriser le travail des étudiants en anticipant la plausible publication de leurs collections de mémoires, de thèses, ... Anaïs a procédé à la numérisation des mémoires de l'École Supérieure de Journalisme de Lille en vue de les conserver

sous un format numérique, soit la première phase dans le processus de traitement de documents. Élodie, quant à elle, a mis en place toute une démarche de valorisation de travaux d'étudiants (mémoires, thèses, ...) issus de l'École Nationale Supérieure des Arts et des Métiers de Lille. Ensuite, elle a indexé et catalogué la collection au sein de la bibliothèque Flora pour la rendre accessible aux étudiants en bonne et due forme. Enfin, Florian a uniquement procédé au traitement des travaux d'étudiants en École Nationale Supérieure d'Architecture et de Paysage de Lille, soit l'indexation, en vue d'une diffusion. De manière globale, leurs démarches sont ancrées dans l'anticipation d'une phase de diffusion prochaine. Chacun des trois stagiaires a accordé de l'importance au traitement de travaux d'étudiants, soit en numérisant, indexant ou cataloguant une collection pour qu'elle puisse être « conforme » lors de sa publication future. Nul doute que les établissements/formations concernés par ces trois stages ont perçu l'importance de diffuser et rendre accessibles leurs collections composées de mémoires d'étudiants entre autres. Tous ont procédé à l'étape de numérisation, de vérification, de correction et d'indexation de leurs corpus dans le but de les préparer à une future diffusion (interne ou externe).

e) Finalité de ces analyses

Au cours de mes analyses, je me suis aperçu que le qualificatif « Mémoire » n'était quasiment jamais utilisé par les portails consultés. Il était fréquent que celui-ci soit assimilé au terme « Thèse » et inclus dans son spectre. Il m'a alors été difficile de mettre en place cette étude en me concentrant uniquement sur les mémoires de Master. J'ai donc procédé à ce choix d'inclure les mémoires au sein de l'environnement des thèses dans ma démarche de recherche. Ce qui me semblait être le plus cohérent et judicieux des choix à adopter.

Mon observation via les archives ouvertes MemSIC et DUMAS (HAL) a permis de déceler une forte diminution des dépôts de mémoires d'étudiants en SHS dans l'espace MemSIC. Aujourd'hui, ce nombre certifie une forte baisse, voire un arrêt des dépôts sur cette plate-forme. Par ailleurs, l'espace MemSIC se concentre uniquement sur les mémoires d'étudiants en Information et Communication. Donc il restreint son activité et il lui est difficile de faire face à d'autres espaces beaucoup plus ouverts. A contrario, l'archive ouverte DUMAS connaît une nette progression de son chiffre de publication. Le nombre d'établissements inscrits, de collections disponibles et de mémoires archivés en son sein témoigne de sa popularité qui ne cesse de croître. Cela montre bien l'importance qu'accorde les établissements à la publication et l'archivage des mémoires d'étudiants à l'intérieur d'espaces en libre accès. À cela s'ajoute un nombre de consultations des mémoires d'étudiants tout de même conséquent observable sur MemSIC malgré le désintérêt marqué des déposants. Bien entendu, cette observation nous a, par la même occasion, conforté dans notre décision. Les diagrammes étudiés sur l'annuaire OpenDOAR confirment cet engouement provoqué par la plate-forme HAL (hébergeant DUMAS) avec un nombre de dépôt en France représentant plus de la moitié du total des publications effectuées. Par

ailleurs, celles-ci concernent en majorité les mémoires et les thèses. Ce qui indique une fois de plus un intérêt particulier accordé à la publication de mémoires. La consultation de l'annuaire ROAR démontre aussi la prise de conscience des établissements à publier les thèses et mémoires sur des plateformes en ligne du fait de l'évolution des dépôts observée. Enfin, et de manière plus globale, le site ROARMAP permet de constater l'importance de la mise en place des politiques de libre accès dans le paysage des organismes de recherche. Ce qui tend à confirmer une prise de conscience générale en ce qui concerne le dépôt et l'archivage de documents emprunts d'informations porteuses de connaissances.

Il y a donc bien un intérêt perceptible au niveau de la publication des mémoires. Certes, certaines données ne couvrent pas le domaine des Master précisément, mais elles nous confortent dans l'idée qu'aucun document issu des différentes instances ne doit être laissé de côté. Chacun ayant un apport de connaissances à partager. Les mémoires d'étudiants de Master font partis de cet environnement documentaire et, d'après cet État de l'art, il semble nécessaire de continuer à leur faire jouer un rôle au sein de cet espace puisqu'il bénéficie d'un attrait particulier.

IV. Le cas du projet MemorySID : Processus de traitement des mémoires de Master

Suite aux analyses effectuées et aux déductions apportées, cette quatrième partie aura pour objectif de mettre en relation notre démarche de diffusion des mémoires avec ces résultats afin de déterminer les liens pouvant exister entre ces deux notions (démarche et déductions). Ce projet de stage a pour particularité de servir d'exemple concret dans le but d'étudier les méthodes mises en place lors du processus de traitement des mémoires de Master d'étudiants et, plus généralement, celui de la Littérature Grise.

1. *Bref récapitulatif et vue d'ensemble*

Même si tous les aspects de l'élaboration du projet ont été détaillés dans les parties précédentes. Je me permet d'initier un bref récapitulatif de ces différentes étapes pour avoir une vue d'ensemble du travail effectué et remettre en contexte le cas MemorySID.

Au cours du projet, les points primordiaux à travailler résidaient dans la vérification des méta-données du corpus de mémoires afin qu'ils disposent tous du même format de description. Une fois un format normé appliqué, il était possible de publier les mémoires dans de bonnes conditions. Puis, le choix de la plate-forme de publication était essentiel afin de rendre accessible la collection de la meilleure des façons qui soit. La totalité de ce processus a dû se faire étape par étape pour ne rien omettre car un retour en arrière n'était pas permis. Dès lors que la vérification des méta-données et l'indexation des mémoires ont été complétées, nous devons passer à la deuxième phase du projet : la diffusion. Nous disposons d'une collection d'environ 740 mémoires de Master. Comme étant spécifié à de multiples reprises, nous nous sommes interrogés sur la manière dont nous allions procéder pour publier notre collection sur l'intranet de Lille 3 et en libre accès sur internet.

La publication en interne pour les étudiants et les enseignants ne posait pas réellement d'inconvénients. Le choix de la plate-forme d'archivage était évidente puisque tous les mémoires y étaient stockés. Il s'agissait de Nuxeo. Cette méthode de dépôt permettait à n'importe quel étudiant et enseignant inscrit à Lille 3 de parcourir l'intranet dédié (l'ENT) pour accéder aux mémoires des étudiants de Master SID. Il nous a fallu trouver l'emplacement adéquat dans l'arborescence de l'espace de publication de Nuxeo pour faciliter l'accès aux utilisateurs mais aussi optimiser l'utilisation de cet espace. La publication en libre accès, quant à elle, posait la question du choix de l'archive ouverte à utiliser pour y stocker nos mémoires. La plate-forme DUMAS a donc été choisie et ce choix concordait bien avec les analyses élaborées dans la troisième phase du projet.

2. **L'accès aux mémoires, un enjeu majeur pour la Recherche ?**

Après avoir accompli la vérification des méta-données et dès lors qu'il a fallu se lancer dans la phase de diffusion de notre collection de mémoires (aussi bien en interne qu'en externe), je me suis directement lancé dans l'analyse demandée par J. SCHÖPFEL. Cette étude a permis de démontrer qu'il y avait bien un intérêt croissant au sujet de la mise à disposition des documents issus de la Littérature Grise (y compris les thèses et mémoires d'étudiants). L'intérêt de déposer les mémoires d'étudiants de Master SID sur l'ENT de Lille 3 et en libre accès s'inscrivait dans cette volonté de rendre accessible le lot d'informations qu'ils possédaient, mais aussi de contribuer à un gain de place au sein des locaux du Département SID.

a) **L'accès à une information de qualité et contexte d'utilisation**

De manière générale et selon une étude datant de 2010, « *la littérature grise représente jusqu'à 30 % de la production scientifique. Les analyses scientométriques montrent son impact au niveau des citations, et elle fait partie des collections, portails et archives ouvertes* »¹¹. Les documents issus de la Littérature Grise sont donc pour le moins nombreux et il n'est pas étonnant de percevoir que les informations y étant contenues soit fréquemment citées.

L'accès à une information de qualité demeurait un élément essentiel dans notre démarche. C'est pourquoi nous avons recensé les mémoires étant annotés par les juges des soutenances. Le but était de trier notre collection pour diffuser, par la suite, uniquement les mémoires qualifiés de « bonne qualité » si je puis dire. Cependant, au vu de l'infime quantité de mémoires annotés, nous avons tout de même décidé de les ajouter à la collection de nos futures publications. Lors de mon travail de vérification des mémoires de Master, je n'ai pu m'empêcher de lire la plupart d'entre-eux pour avoir un petit aperçu des thèmes et sujets abordés. De plus, cela m'a aussi permis de me renseigner sur la structuration et la mise en page adoptées fréquemment par les étudiants. Leurs contenus, aussi bien les analyses que les recherches ou méthodologies de travail, étaient tous plus variés les uns que les autres. Ce qui générait une multitude d'informations et de connaissances au sein de cette collection. Pour ma part, il me semblait impensable de passer à côté de cette « masse » de savoir étant donné la diversité de l'information offerte et la qualité récurrente de cette dernière.

Il est donc impensable d'occulter la présence des mémoires de Master dans le paysage de la Recherche puisqu'ils recèlent d'informations « utilisables » et permettent d'alimenter et appuyer des recherches. Il en va de même pour la Littérature Grise. Ces informations peuvent être utilisées dans divers contextes comme celui de la veille. En effet, les sujets des mémoires de Master varient

11 Joachim SCHÖPFEL, *Vers une nouvelle définition de la littérature grise*, 2012, Cahiers de la Documentation, 66 (3), p.14-24

- https://hal.archives-ouvertes.fr/sic_00794984/document

considérablement selon les étudiants mais aussi selon les années. Ce qui permet d'observer l'évolution qui a lieu au sein de l'environnement de l'Information et de la Documentation. Un sujet traité à telle période peut évoluer et être interprété différemment comparé à une période antérieure. La récolte de ce type d'informations peut donc devenir des plus bénéfiques dans le secteur des Sciences Humaines et Sociales en constante évolution. Comme évoqué un peu plus haut, j'ai moi-même utilisé ce type d'informations émanant de la Littérature Grise et, plus particulièrement, des mémoires de Master d'étudiants en Sciences de l'Information et de la Documentation pour élaborer mon mémoire de stage. Impossible de passer à côté de ce genre de ressources au vu de la multiplicité du contenu informationnel dont ils regorgent en terme de recherches, d'analyses, de déductions ou encore de constats. Ce qui peut-être bénéfique pour appuyer nos propres études ou simplement pour prendre connaissance de la méthodologie à appliquer pour structurer son mémoire en bonne et due forme. Dans le cas de mon stage, il m'a été possible de me renseigner sur les prémices de la numérisation de la Littérature Grise via un mémoire de Master (anciennement DESS) datant de 1999, pour le comparer aux processus actuels¹².

Les documents composant la Littérature Grise engendrent donc un intérêt particulier dans la Recherche du fait de leur contenu informationnel florissant. Il est alors inconcevable de passer outre la notion d'accessibilité. Il s'agit ici d'un puits de connaissances en perpétuelle croissance. D'où l'importance de mettre en place des plates-formes adéquates afin d'optimiser leur accès.

b) L'optimisation de l'organisation interne

Bien évidemment, un travail interne doit être effectué en amont de la part des organismes ayant pour objectif de rendre accessibles leurs documents. Sans quoi, le processus de diffusion pourrait s'avérer biaisé. Toute démarche, quelle qu'elle soit, doit être mûrement réfléchi et disposer de bonnes fondations pour pouvoir être lancée. Tout ne dépend pas uniquement de la réception des documents par l'espace de diffusion concerné.

Nous l'avons bien perçu lors de notre évolution au fil des étapes du projet MemorySID. Il a été vital de mettre en place un dispositif cohérent, d'observer ce qui se faisait ailleurs et d'analyser l'environnement ambiant pour construire notre processus. Certes, la mise en application a demandé et demande beaucoup de temps du fait qu'elle soit la première tentative dans notre cas. Mais il est indéniable de constater que la méthodologie utilisée tout au long du traitement du projet porte ses fruits. Par exemple, la phase de numérisation et de dématérialisation des mémoires d'étudiants a obligatoirement dû passer par la case vérification pour s'assurer de la conformité du travail fait. Beaucoup de temps a été consacré à cette tâche. Certes, l'erreur est humaine mais il est préférable de minimiser la casse une fois la publication des mémoires terminée. Il a aussi fallu étudier le mode de fonctionnement d'autres

12 Marie-France CLAEREBOUT, *Numérisation de la Littérature Grise : Mise au point d'une chaîne documentaire*, Lille, 1999-2000.

établissements ayant déjà procédé à un processus similaire pour prendre connaissance de leur méthodologie utilisée (construction de leur collection, mise en page du portail, dénomination des mémoires, liste des méta-données d'indexation, ...). Dans notre cas, nous avons consulté les collections des universités de Lyon mais aussi ceux du portail du laboratoire GERiiCO. Enfin, l'analyse et l'étude de l'environnement dans lequel « baigne » notre projet ont été essentiels pour s'aventurer en terrain connu.

Globalement, il faut user de ces deux éléments que sont la construction d'une base de travail efficace et l'observation de l'environnement extérieur pour pouvoir entamer un processus de publication solide.

c) Facilité la mise à disposition où comment rendre le travail du chercheur plus agréable

Suite à l'étude de 2010, il a été observé que de nombreuses sources provenant de la Littérature Grise étaient citées par les auteurs de travaux de recherche. C'est pourquoi, l'accent doit être mis sur l'amélioration des espaces d'archivages pour en faciliter l'accès et, par la même occasion, décupler leur nombre de consultation. Les plates-formes accueillant les documents tels que les mémoires de stage sont donc les seconds acteurs favorisant une accessibilité en bonne et due forme. Toute recherche scientifique devrait s'orienter dans leur direction. De ce fait, il est nécessaire de faciliter leur accès. Plus les documents liés à la Littérature Grise seront accessibles, plus l'accès à une information de qualité et en rapport avec le sujet recherché sera rapide et efficace. Ce qui induira un travail de recherche de la part du chercheur des plus agréables, des plus intuitifs et des plus « faciles ».

Au cours du projet, Nuxeo a été le support interne nous permettant d'aller au bout de cet aspect. Ce dernier étant fréquemment utilisé par les enseignants et bientôt disponible pour les étudiants. L'utilisation de DUMAS comme portail de notre collection était tout aussi évidente au vu de sa popularité croissante. Dans un premier temps, il nous a été nécessaire de construire une arborescence logique sur Nuxeo pour faciliter la navigation (soit le répertoire « Bibliothèques »). Ensuite, le corps enseignant de Lille 3 ainsi que les étudiants auraient accès au contenu de ce répertoire intuitivement grâce à l'application dédiée disponible via leur espace personnel sur l'ENT de Lille 3. En ce qui concerne la mise en libre accès par l'intermédiaire de l'archive ouverte DUMAS, j'ai mis au point l'arborescence du futur portail accueillant notre collection. J'ai donc fait en sorte d'optimiser l'espace de la page d'accueil afin que les mémoires soient accessibles le plus simplement possible. Le tout dans le but de rendre le travail du chercheur moins contraignant. Pour ma part, au début de mes recherches, j'ai trouvé l'information issues des mémoires ailleurs que sur les plates-formes dédiées. Je me suis alors aperçu qu'il était impératif de se focaliser sur les espaces de recherche de la Littérature Grise. En effet, l'élaboration d'une structuration efficace et d'une visualisation épurée permettait de valoriser l'accès et l'utilisation de ce type de plate-forme, cependant il fallait déjà connaître l'existence de ces

espaces pour pouvoir y accéder. J'ai constaté qu'une démarche de sensibilisation du public à l'existence de ce type de portail serait des plus bénéfiques afin qu'il puisse savoir où chercher ce genre d'informations et comment y procéder. L'aspect « communication » est aussi un point central lors du processus de diffusion de ces documents. Il est important de ne pas le négliger pour maximiser leur consultation et montrer aux chercheurs qu'il existe des espaces appropriés et complets. C'est pourquoi, Monsieur SCHÖPFEL avait anticipé cette notion en comptant communiquer à ce sujet par l'intermédiaire du projet MemorySID. Ce qui permettrait de mettre en lumière la présence de ces espaces dans l'environnement de la Recherche (chercheurs, enseignants, étudiants, ...). Par ailleurs, nous avons fait part de la disponibilité des mémoires sur l'ENT de Lille 3 via un petit article et un Tweet sur le compte de l'ANRT, en espérant élargir un peu plus la communication sur ce sujet.

Une mise à disposition plus cohérente et plus transparente passe alors par une simplification du mode d'accès des plates-formes d'archivage de la Littérature Grise, une valorisation des collections et une prise de conscience de l'existence de ce type de plate-forme au sein de l'environnement de la Recherche.

3. *Diffusion et accessibilité, deux notions pour un même objectif*

a) Un objectif commun

La diffusion de documents issus de la Littérature Grise tels que les mémoires de Master passe irrémédiablement par une cohérence dans la démarche entreprise au sein de l'organisme concerné. Ce processus doit être suffisamment réfléchi à l'aide d'un contrôle sérieux de la collection potentiellement diffusée et d'une prise de connaissance de l'environnement « d'accueil ». Le contrôle permettant de proposer une information de qualité, et la prise de connaissance favorisant le transfert des documents. Une fois concrétisé, la mise à disposition de la documentation doit se faire au sein de plates-formes adéquates et disposant des éléments nécessaires à un archivage effectué dans de bonnes conditions. L'accès sera des plus appropriés afin de faciliter le travail de recherche par les chercheurs, que ce soit dans une optique de veille ou toute autre pratique requérant la collecte d'informations.

Sans une démarche de diffusion cohérente et un accès adapté, la consultation de documents ne peut se faire dans de bonnes conditions. Ces deux éléments doivent donc fonctionner de paire pour : Faciliter la recherche d'informations, la rendre intuitive et proposer une information de qualité.

b) L'importance du référencement

L'on a vu au cours du processus de diffusion que l'accessibilité de ce type de documentation passe obligatoirement par des espaces habilités à le contenir, soit les archives ouvertes. Les mémoires sont

donc référencés par l'intermédiaire de ces plates-formes. Ce qui permet d'en faciliter l'accès et d'éviter aux collections d'être noyées dans une masse de documents provenant de divers établissements, formations, ... Les archives ouvertes telles que DUMAS proposent différents portails étant spécifiques à chaque établissement. Ensuite, chacun d'eux doit s'organiser pour structurer correctement son portail et éviter les doublons ou toute autre incohérence. Dans notre cas, par exemple, l'UFR Humanité avait déjà créé son propre portail. Il nous était alors impossible de créer un nouveau portail en Sciences Humaines et Sociales pour pouvoir ensuite ajouter notre collection concernant le Master en Sciences de l'Information et de la Documentation. De plus, la collection déjà créée n'était pas très active... Nous aurions préféré construire un portail SHS dans lequel y aurait été regroupés les différentes UFR. Puis, chaque UFR aurait disposé de sa propre collection. Cette absence de concertation nous a mené à créer notre propre portail de notre côté...

Un bon système de référencement permet à ce type de document d'être retrouvé plus facilement et éviter des pertes ou des illogismes. Ce travail de structuration doit être fait en amont pour permettre un référencement logique et intuitif. Il est important de se concerter au sein même des établissements pour améliorer cette accessibilité et la diffusion des futurs mémoires de Master. Par ailleurs, la réunification des trois universités de Lille étant effective en 2018 a été l'un des éléments que nous avons laissé en suspens pour éviter de prendre du retard dans l'accomplissement de notre projet.

Conclusion

Malgré que je ne sois pas encore arrivé au terme de ma période de stage, les missions auxquelles j'ai participé jusqu'à présent m'ont permis d'engranger des connaissances au sein d'un projet d'indexation et de diffusion de documents. De plus, cette expérience au sein d'une petite entreprise m'a donné l'opportunité de percevoir en profondeur le travail collectif et coopératif puisque, en dehors de ma mission principale, il m'a été possible d'observer le rôle de chacun des membres du personnel de l'ANRT avec précision. Chacun ayant une fonction précise et majeure dans l'élaboration de divers tâches de dématérialisation. Les conditions de travail ont été idéale pour me renseigner sur les différentes étapes de ce processus. L'apport de connaissances était donc double du fait de mon implication et du travail fourni au sein du projet MemorySID mais aussi de part ma présence dans les locaux de l'ANRT idéale pour me renseigner sur le fonctionnement interne de l'atelier.

Mes tâches de vérification et de correction des méta-données des mémoires, aussi répétitives fussent-elles, ont été l'occasion d'acquérir de plus amples compétences sur le sujet de l'indexation de documents, et plus particulièrement de documents faisant partie du champs de la Littérature Grise. J'ai eu la chance de pouvoir solliciter mes collègues afin de collaborer et échanger tout au long de ma mission. L'étude concernant la mise en place d'une démarche de diffusion de mémoires et l'analyse de l'environnement de publication m'ont été bénéfiques en me permettant de me renseigner sur le sujet mais aussi de me remettre en question au cours ma mission. Cette dernière a contribué à l'étude de ma problématique. Problématique qui a émergé de manière évidente au fil de mon travail. Ce qui a permis d'étendre mes interrogations et analyses à l'échelle de la Littérature Grise tout en étant ancrées dans le périmètre des mémoires de Master d'étudiants.

L'étude de cas du projet MemorySID a été nécessaire et idéale dans la compréhension des processus de traitement des documents composant la Littérature Grise. En effet, elle m'a permis de réaliser que ce type de processus n'était et n'est pas aussi simple qu'il ne le prétend. Celui-ci requiert une très bonne organisation interne de la part des organismes voulant publier leurs collections de mémoires d'étudiants, entre autres. Mais il dépend aussi des plates-formes accueillant ce genre de documentation puisqu'il leur est nécessaire de proposer un « système d'utilisation » efficace pour pouvoir accueillir les collections dans de bonnes conditions et donner aux chercheurs un confort dans leur processus de recherche d'information. De ce fait, trois étapes sont vitales dans l'aboutissement d'un tel projet de diffusion : Le travail interne contribuant à la conformité des documents et à la sélection d'une information de qualité. La mise à disposition d'espaces de stockage adaptés à l'accueil de ce type de documentation. Et, la diffusion de la collection dans de bonnes conditions nécessitant

une cohésion entre les divers organismes. Ces trois étapes contribuent irrémédiablement à l'accessibilité de documents composant la Littérature Grise et facilitent, par la même occasion, le travail de recherche.

La Littérature Grise ne cesse d'évoluer et il est difficile d'en donner une définition précise. Cependant, elle doit continuer à s'adapter à l'évolution de l'environnement numérique qui l'entoure et des nouveaux outils mis à sa disposition. Tout cela, dans une ambition perpétuelle de valorisation de son contenu. Il est important d'accorder une place à chaque composante du processus de traitement de la Littérature Grise puisque, chacune d'elles contribue à cette valorisation. Cela lui permet d'entretenir l'intérêt porté à son égard et d'offrir aux chercheurs (étudiants, enseignants, ...) un environnement de recherche de qualité et diversifié au possible. Il ne faut pas oublier que chaque mémoire d'étudiants, quelque soit son contenu, est une source de connaissances même si infime soit-elle.

Enfin, et si toutes les conditions sont favorables, ce dernier mois au sein de l'ANRT me permettra de finaliser ma mission : L'agencement de la structuration du portail SID sur DUMAS ainsi que la diffusion de notre collection à l'intérieur de celui-ci. Monsieur SCHÖPFEL compte, par la suite, transférer les anciens mémoires déposés sur MemSIC dans notre portail DUMAS. Si le laps de temps restant me le permet, je pourrais peut-être participer à cette opération. Entre-temps, j'utiliserai précieusement ces derniers jours passés à l'ANRT pour m'intéresser et contribuer à d'autres tâches comme par exemple l'élaboration d'un cahier des charges concernant la mise en place d'un nouveau site web pour l'ANRT.

Perspectives futures du projet :

La principale ambition du projet MemorySID était de servir d'inspiration pour d'autres UFR ou organismes produisant des mémoires de stage d'étudiants de Master. Il pourrait donc servir de modèle pour d'autres initiatives. Dès lors qu'il aura abouti, il fera l'objet de publications et de communications au niveau national et international pour étendre son impact à plus grande échelle. De plus, le projet pourrait contribuer à rendre le processus de dématérialisation, d'indexation et de diffusion automatique et récurrent à la suite des soutenances de Master ayant lieu à chaque fin d'année universitaire. Ce processus pourrait, pourquoi pas, devenir une norme dans le domaine. Enfin J. SCHÖPFEL envisage la création d'une bibliothèque numérique, en complément, pour les formations en Sciences de l'Information et de la Documentation pour le dépôt de futurs mémoires en Master SID. Cette bibliothèque serait alors la plate-forme de référence pour le dépôt et la consultation de mémoires de Master SID d'étudiants. Le projet bénéficie de bases solides pour impacter les processus futurs de dématérialisation et de diffusion de mémoires de stage d'étudiants de Master.

Bibliographie

Cours :

- J. SCHÖPFEL, *Numérisation*, 2016-2017, Master 1 Sciences de l'Information et de la Documentation.

Mémoire de stage :

- Marie-France CLAEREBOUT, *Numérisation de la Littérature Grise : Mise au point d'une chaîne documentaire*, Lille, 1999-2000.

Webographie

Sites internet :

- ANRT (professionnel), *l'ANRT* : <http://www.diffusiontheses.fr/content/4-anrt-lille-reproduction-theses>
- ANRT (institutionnel) - Lille 3, *Présentation de l'ANRT* : <https://anrt.univ-lille3.fr/>
- ANRT (institutionnel) - Lille 3, *Équipement* : <https://anrt.univ-lille3.fr/nos-equipements>
- ANRT (institutionnel) - Lille 3, *Partenaires de l'ANRT* : <https://anrt.univ-lille3.fr/nos-partenaires>
- Centre National de Ressources Textuelles et Lexicales, *Microfiche* : <http://www.cnrtl.fr/definition/microfiche>
- HAL, *archive ouverte* : <https://hal.archives-ouvertes.fr/>
- Institut de l'information scientifique et technique, *Libre accès à l'information scientifique et technique* : <http://openaccess.inist.fr/?+-Protocole-OAI-PMH->
- OpenGrey, *Présentation OpenGrey* : <http://www.opengrey.eu/>
- UrfistInfo, *DUMAS* : <https://urfistinfo.hypotheses.org/43>
- Wikipédia, *AFNOR - Littérature Grise* : https://fr.wikipedia.org/wiki/Litt%C3%A9rature_grise
- Wikipédia, *Atelier National de Reproduction des Thèses* : https://fr.wikipedia.org/wiki/Atelier_national_de_reproduction_des_th%C3%A8ses
- Wikipédia, *Offset (imprimerie)* : [https://fr.wikipedia.org/wiki/Offset_\(imprimerie\)](https://fr.wikipedia.org/wiki/Offset_(imprimerie))

Articles internet :

- Jacqueline NIVARD, *Les archives ouvertes*, La lettre de l'EHESS n°34 : <http://lettre.ehess.fr/988>
- Extrait du rapport annuel du Conseil supérieur des bibliothèques, *Le traitement de la Littérature Grise*, 1993, p.1. : <http://www.enssib.fr/bibliotheque-numerique/documents/1093-le-traitement-de-la-litterature-grise.pdf>
- Jacques KERIGUY, *L'accès à la littérature grise*, 1984, BBF n°2 : <http://bbf.enssib.fr/consulter/bbf-1984-02-0138-001>
- Joachim SCHÖPFEL, *Vers une nouvelle définition de la littérature grise*, 2012, Cahiers de la Documentation, 66 (3), p.14-24 : https://hal.archives-ouvertes.fr/sic_00794984/document
- The Grey Journal, *Research on Grey Literature in Europe*, 2010, Volume 6 n°3, (ISSN 1574-1796) : http://www.textrelease.com/images/TGJ_V6N3.pdf

Lexique

- **Archive ouverte** : Plate-forme en ligne où sont déposés de nombreux documents issus de la recherche scientifique et de l'enseignement et dont l'accès n'est pas restreint.
- **Dématérialisation** : Procédé permettant le « transformation » d'un document physique (papier) en un document numérique (fichier).
- **Dublin Core** : Format normé utilisé dans la description de ressources physiques ou numériques. Il est composé de 15 champs.
- **Façonnage** : Procédé suivant l'impression qui permet de mettre en forme un document physique (sous forme de livres, ...).
- **Format PDF** : C'est un langage de description de pages préservant la mise en forme d'un document.
- **Impression** : Procédé utilisé pour affiché du contenu sur un document physique (papier).
- **Massicotage** : Opération de découpage du papier.
- **Microfiche** : Support de stockage analogique compact.
- Microfiche argentique : Il s'agit du support de stockage mère.
- Microfiche diazoïque : Il s'agit d'un duplicata de la microfiche mère.
- **Numérisation** : Procédé qui permet de stocker le contenu (données) d'un document physique sous forme de fichier numérique.
- **Nuxeo** : Espace de Gestion Électronique de Documents.
- **Offset** : Procédé d'impression majeur s'adaptant à différents supports (papier, carton, ...) et s'appliquant à divers publications (presse, publicité, livre, brochure, catalogue, ...).
- **Outil TDM** : Outil de traitement informatique permettant d'extraire de la connaissance au sein de documents.
- Reliure : Opération finale d'impression permettant de rassembler l'ensemble des feuilles imprimées pour la mise en forme (livre, prospectus, ...).
- **Système GED** : Procédé informatisé permettant d'organiser et gérer des documents électroniques dans une organisation.
- **Tags** : Ensemble de mots-clés décrivant un documents.
- **« Thèse à la carte »** : Ce service proposé à l'ANRT permet d'obtenir un exemplaire d'une thèse sous format livre.

Sigles

- **AFNOR** : Association Française de Normalisation.
- **ANRT** : Atelier National de Reproduction des Thèses.
- **CCSD** : Centre pour la Communication Scientifique Directe.
- **CINES** : Centre Informatique National de l'Enseignement Supérieur.
- **CNRS** : Centre National de la Recherche Scientifique.
- **DESS** : Diplôme d'Études Supérieures Spécialisés.
- **DIST-RD** : Département de l'Information Scientifique et Technique et Réseau Documentaire.
- **DSI** : Direction des systèmes d'information.
- **DUMAS** : Dépôt de Mémoires Universitaires Après Soutenances.
- **ENT** : Espace Numérique de Travail.
- **GED** : Gestion Électronique de Documents.
- **GERiCO** : Groupe d'Études et de Recherche Interdisciplinaire en Information et Communication.
- **HAL** : Hyper Articles en Ligne.
- **PDF** : Portable Document Format.
- **ROAR** : Registry of Open Access Repositories.
- **SHS** : Sciences Humaines et Sociales.
- **SIC** : Sciences de l'Information et de la Communication.
- **SID** : Sciences de l'Information et de la Documentation.
- **TDM** : Text Data Mining.
- **UFR** : Unité de Formation et de Recherche.
- **XML** : Extensible Markup Language.

Résumé

Ce mémoire a été élaboré à la suite d'un stage ayant eu lieu à l'Atelier National de Reproduction de Thèses de Villeneuve d'Ascq. Il s'agit d'une étude de cas concernant le Projet MemorySID et s'intéressant aux processus de traitement et de diffusion de mémoires de stage d'étudiants de Master en Sciences de l'Information et de la Documentation. L'environnement de travail y est décrit ainsi que les missions confiées et la méthodologie mise en place. Ce mémoire de stage évoque les phases d'indexation et de vérification des méta-données de mémoires sur la plate-forme Nuxeo ainsi que les étapes de diffusion dans l'ENT de Lille 3 et en Libre Accès dans l'archive ouverte DUMAS. Enfin, une analyse étudiant l'impact et l'importance qu'engendre l'accès aux mémoires de Master et à la Littérature Grise dans le domaine de la Recherche complète ce mémoire et permet de s'interroger sur l'intérêt de cette démarche.

Mots-clés : Mémoire de stage, master, Nuxeo, DUMAS, GED, Gestion Électronique de Documents, ENT, Environnement Numérique de Travail, Archive Ouverte, indexation, diffusion, méta-données, Littérature Grise, Libre Accès, Sciences de l'Information et de la Documentation.

Abstract

This report had been developed following an internship that took place in the « Atelier National de Reproduction de Thèses » in Villeneuve d'Ascq. It is a case study about the « MemorySID » Project. It seeks to deal with Treatment and dissemination processes of Master student's internship report (Information and Documentation Science). It specifies the working environment, missions, and methodology used. This internship report mentions indexation and verification of the report's meta-data in Nuxeo. It specifies the dissemination's stages on the « ENT de Lille 3 » and in Open Access on the open archive DUMAS. Finally, it is completed by an analysis about the impact and the importance of the internship report and Grey Literature's accessibility in the Research field.

Keywords : Internship report, master, Nuxeo, DUMAS, EDM, Electronic Document Management, DWE, Digital Work Environment, Open Archive, indexation, dissemination, meta-data, Grey Literature, Open Access, Information and Documentation Science.