



HAL
open science

Master's theses and open scholarship: a case study

Joachim Schöpfel, Sylvain Vanacker, Eric Kergosien, Bernard Jacquemin

► **To cite this version:**

Joachim Schöpfel, Sylvain Vanacker, Eric Kergosien, Bernard Jacquemin. Master's theses and open scholarship: a case study. *Digital Library Perspectives*, 2018, 34 (4), pp.276-287. 10.1108/DLP-07-2018-0021 . hal-01911939v1

HAL Id: hal-01911939

<https://hal.univ-lille.fr/hal-01911939v1>

Submitted on 12 Nov 2018 (v1), last revised 21 Nov 2018 (v2)

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.

Automatic Identification of Research Fields in Scientific Papers

Joachim Schöpfel, Sylvain Vanacker, Eric Kergosien, and Bernard Jacquemin

Univ. Lille, EA 4073 – GERiiCO, F-59000 Lille, France
prenom.nom@univ-lille.fr

Abstract

Purpose – This paper aims to show how Master’s theses can contribute to open scholarship and give reasons why this should be done.

Design/methodology/approach – The paper provides an overview of published studies and, based on the experience at the University of Lille (France), describes some essential aspects for the processing and valorization of these documents in the academic cloud, as a contribution of open scholarship.

Findings – Because of their number and diversity, collections of Masters’ theses in open repositories could be an excellent showcase for the universities’ Master programs and research. They could also offer interesting and large samples for content analysis, citation analysis and text and data mining (TDM). However, some issues need attention, above all intellectual property, quality and preservation. Quality is crucial, and the paper describes how the Lille project proceeds to assure sufficient quality and right clearance, and why the project shifted from students’ self-archiving to a digital library collection in the academic cloud, run by faculty and information professionals. The paper presents also some usage statistics to illustrate the potential, global impact of such a collection.

Practical implications – The paper provides helpful and empirical evidence and insight for those who want to develop the dissemination of Master’s theses via open repositories.

Originality/value – In the context of open scholarship, only few studies deal with Master’s theses, and this paper is the only recent reference that brings together a review of other papers and a case study with empirical evidence.

Keywords – Digital libraries, Open repositories, Open access, Electronic theses and dissertations, Master’s theses, Open scholarship, Cloud

Paper type – Case study

1 Introduction

Open scholarship has become the new paradigm of scientific communication. The term refers to “teaching and research practices that espouse openness (i.e.) a collection of emergent scholarly practices that espouse openness and sharing”, including open access, open publishing and open education (Veletsianos et al., 2012, p. 167). An ever-increasing number of books, journals and repositories provide free and unrestricted access to scientific content from all disciplines. Two third of the open institutional repositories contain more or less important collections of electronic theses and dissertations (ETDs), mostly PhD theses, considered as the most useful kinds of invisible scholarship and the most invisible kinds of useful scholarship (Suber, 2012).

The situation is quite different for Master’s theses. Even if their number is several times higher¹, they are much less well represented in open repositories than PhD theses. Less than 1 per cent of all repositories clearly indicate that they contain Master’s theses, and freely available Master’s theses represent less than 0.5 per cent of the scientific items retrieved by the Bielefeld Academic Search Engine. However, Master’s theses can be an excellent showcase for the academic Master programs and research. They are original work of the mind, and at least for the best of them, they contribute to scientific and professional knowledge. The following paper provides an overview on the nature, publishing an interest of Master’s theses and, based on our project at the University of Lille (France), describes some essential aspects for the processing and valorization of these documents in the academic cloud, as a contribution to open scholarship.

2 Master’s theses – an overview

The Master’s degree dates back to the origin of European universities. Today, it is usually a second-cycle academic degree requiring previous study at the Bachelor’s level (or equivalent). It is awarded by universities or colleges upon completion of a one or two years program demonstrating advanced disciplinary knowledge or professional skills. Master’s degree names and the structure and duration of Master programs vary according to the country and even the university; diversity is the rule.

In general, the granting of a Master’s degree requires some kind of thesis or dissertation on a research or a professional project (internship), where the students have to demonstrate their ability of independent and rigorous thinking and of complex problem solving, along with field-specific methodology and knowledge. A recent report issued by the US National Academies of Sciences, Engineering, and Medicine, 2017 highlights the potential scientific quality of Master’s theses based on substantial undergraduate research experiences defined as “an inquiry or investigation conducted by an undergraduate student that makes an original intellectual or creative contribution to the discipline”².

¹In France for instance, Master students represent 35% of all university students, compared to 4% PhD students, in other words, the relation is 10 to 1. Statistics from 2017, covering all disciplines including medicine, produced by the French Ministry for Higher Education and Research www.enseignement-sup-recherche.gouv.fr/cid126315/reperes-et-references-statistiques-edition-aout-2017.html

²Council of Undergraduate Research <https://www.cur.org/>

Alongside research capacities and originality, quality is assessed in terms of innovation and discovery, refinement of experimental design, new research techniques, teamwork and writing skills. “As Master’s theses aim to provide evidence of the skills learned, both practical and theoretical skills are exemplified in theses” Seliger, 2015, p. 131. Their quality is impacted by the Master program and in particular, by the importance of research. Master’s theses are a “complex genre”, considering both the set of problems related to its configuration (structure, language, norms of reference) and the factors that constrain its production (methodological procedures, student/supervisor relationship, time management etc.) (Carvalho et al., 2017).

As mentioned above, few repositories contain Master’s theses. Out of the 3,519 repositories of the OpenDOAR directory, 57 per cent contain ETDs but only 0.5 per cent clearly indicate that this means also Master’s theses. The international Registry of Open Access Repository Mandates and Policies (ROARMAP) with 941 open access policies keeps a record of 37 per cent mandates and instructions regarding ETDs but nothing about Master’s theses. Many institutions do not consider undergraduate work as research output worth open scholarship. Those which do have different options. Three vectors of open access dissemination via repositories can be distinguished.

Institutional repositories: Some institutions disseminate their undergraduate students’ theses and dissertations in their own institutional repository, as part of its intellectual output and along with articles, communications, working papers etc. For instance, in the University of Amsterdam, *Scripties* repository Master’s theses represent 70 per cent of the nearly 50,000 undergraduate works³.

General aggregating repositories: These multi-institutional or consortial repositories aggregate files and metadata from different document types and institutions. The Latin America portal *LA Referencia* for instance provides access to nearly 1.5 million publications, one third being Master’s theses, mostly from Brazil⁴.

ETD repositories: As a specific variant of the multi-institutional repositories mentioned above, they only contain ETDs from different institutions, like the *OhioLINK Electronic Thesis and Dissertation Center* with over 90,000 undergraduate, Master’s and doctoral theses and dissertations from 32 universities and colleges⁵ and the Scandinavian *DiVA* portal for academic publications and student papers produced at 44 universities, colleges of higher education, research institutes and museums, with more than 250,000 student papers from BA, Master and other diplomas of which 72 per cent are in open access⁶.

Because of their number, diversity and quality, Master’s theses prove their usefulness for studies on the evolution and distribution of research subjects and methodologies. A synthesis of published research with samples of Master’s theses reveals three different approaches (for more details and further reading, see annex).

Content analysis: Nearly 20 recent papers from Iran, Slovenia, Brazil, China, Zambia, Turkey and other countries share an exploratory-descriptive approach and aim at assessing research paradigms, the diversity and trends of research topics, writing skills and information literacy, the students’ background and the choice of research

³*Scripties* www.scriptiesonline.uba.uva.nl

⁴*LA Referencia* www.lareferencia.info

⁵*OhioLINK ETD Center* <https://etd.ohiolink.edu/>

⁶*DiVA* www.diva-portal.org

methods like qualitative or action research. They also share the basic assumption that Master's theses generally display sufficient quality and relevance for content analysis and scientometrics.

TDM: Some studies, especially from China, apply TDM techniques to Master's theses, such as quantitative analysis of the frequency of key words, co-word matrixes, co-occurrence network analysis, factorial or cluster analysis, strategy coordinate analysis and multidimensional scaling analysis. Their objective is the assessment of research topics ("hot issues"), methodologies or funding trends through the automatic exploitation of Master's theses.

Citation analysis: Citation analysis constitutes a third approach to Master's theses. These studies are based on the references produced by the Master students and assess the use of scientific literature in terms of document types, preferred journals, geographic origin of cited sources, and so on.

Along with the potential for scientometrics and TDM, there are other reasons for open access to Master's theses, such as enhancing accessibility, increasing the impact in scientific journals and conferences (Gillet et al., 2013), preventing plagiarism, strengthening visibility for Master's institutions (Xia et al., 2010), contributing to the curriculum quality and fostering the potential use of outcomes especially from less known programs and disciplines (Martens, 2011).

Quality is essential for the interest and potential of Master's theses. Not all Master's theses' reflect substantial undergraduate research experiences or professional projects. Poor quality of abstracts can be a limit to TDM. Evaluation can be helpful to control the quality and to select the best papers. Even if the criteria are neither homogeneous nor applied in the same way everywhere, with a large diversity of written guidelines and evaluation grids, of the supervisors' role and of oral exams to adjust the final mark (Bøhn et al., 2011), the evaluation generally includes some similar features, such as background knowledge, description of methods, data and findings, discussion of the findings and the use of references. Sometimes, the Master assessment includes research skills and the relation with the supervising research team. And even if their quality may vary and should be controlled, their diversity and richness make them interesting and useful resources for citation analyses, for content mining and scientometric assessment. But this potential requires open accessibility and reusability. The Lille project *MemorySID* provides some insight how to proceed.

3 The Lille project *MemorySID*

The following section describes the project *MemorySID* at the University of Lille (France), with nearly 70,000 students the largest French University and one the largest French-speaking universities worldwide. The project started in 2008, to contribute to open scholarship through open access publishing of Master's theses in library and information sciences, via a national infrastructure in the public academic cloud. With the years, the project rationale shifted from self-archiving to a two-layer approach with two digital libraries, one in intranet and the other in the cloud.

3.1 Phase one (2008-2015): self-archiving

Ten years ago, from 2008 on, the Department of Information and Documentation Sciences (SID) at the University of Lille (France) started encouraging the Master students to deposit their theses on a national repository for Master's theses library and information sciences called *mémSIC*⁷. The objective was to publish a selection of good and non-confidential works in open access and to guarantee long-term findability, accessibility and preservation through a national OA infrastructure, to valorize the students' work and to increase the Master programs' visibility (Chauvin et al., 2010; Mann, 2010). The three guiding principles for this first project phase were:

- validation by the jury (faculty);
- double authorization by the student and by the project tutor (internship); and
- self-archiving by the students.

The academic cloud was preferred to a local, institutional solution to optimize visibility and usage. The success was mixed. Between 2008 and 2015, 70 theses were deposited on *mémSIC*, with scores equivalent to the US academic A and B grades (for French readers: 14/20 or more). The impact in terms of usage was more than satisfying, with statistics going up to several thousand downloads per paper (see below). However, the main problem was the lack of motivation for self-archiving and, in some cases, the lack of quality of metadata and files. Another problem was that *mémSIC* does not allow the creation of an institutional collection or portal, thus reducing the impact for the Master program.

3.2 Phase two (from 2016 on): digital library

To resolve these issues, a follow-up project was launched in 2016 called *MemorySID* Vanacker, 2017. For this second phase, the priorities were different:

- retrodigitization of all print Master's theses still in the academic library and department holdings;
- integration of all digitized Master's theses into the campus-wide Nuxeo document management system (DMS), accessible for students and scholars via their intranet workspace; and
- creation of a new collection on the national repository *DUMAS*, with a selection of the best theses ("showcase").

In other words, we privileged a two-layered approach with two digital libraries, one as a folder in the academic intranet and the other as a portal in the cloud, on the national ETD repository *DUMAS* (Figure 1). *DUMAS* is hosted by the French Center for Direct Scientific Communication⁸ which operates also the French national

⁷Mémoires de master en sciences de l'information et de la communication <https://memsic.ccsd.cnrs.fr>

⁸CCSD <https://www.ccsd.cnrs.fr/>

open repository HAL, and it contains more than 19,000 Master’s theses. Only library and faculty can deposit files and create metadata. There is no self-archiving. Each deposit needs validation by the repository’s administrator. Insofar, *DUMAS* is a kind of crossover between an open repository and a digital library.

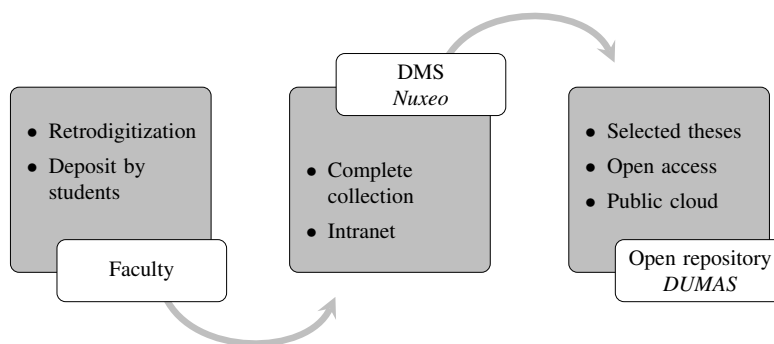


Figure 1: Two-layered approach of *MemorySID*

After removing duplicates, 765 print theses were digitized, 583 from the library holdings (1996-2004) and 227 from the SID department (2005-2016). The text files (PDF) were uploaded in the DMS and indexed in a Dublin Core (DC) compliant metadata format. To the 765 print works from 1996-2016, 9 native digital Master’s theses were added in 2017. From these 774 Master’s theses, 715 are searchable and available for all students and scholars on the campus, via the academic virtual workspace, while the others remain in the “dark” part of the intranet because of their confidential content. The digital library (folder) is administered by the faculty (SID Department), while the content of the DMS is maintained and backed-up by the university IT Department.

From the 715 Master’s theses in the public DMS, 87 per cent have a sufficient score (grade A or B) and are authorized for online dissemination, according to the administrative records (i.e. defense minutes). So far (July 2018), 523 theses from the DMS have been deposited on the *DUMAS* repository and are online. Additionally, the metadata from 64 self-archived theses on *mémSIC* (first project phase) were transferred to *DUMAS*. This means that at present, the new Lille LIS Master’s theses collection on the *DUMAS* repository provides unrestricted access to 587 theses from 1996 to 2017.

3.3 Usage of the *DUMAS* collection

The following usage statistics are from July 19, 2018. They assess the download figures for all 587 Master’s theses available on *DUMAS*, for the period from January 2015 to June 2018.

Cumulative figures: up to now, the 587 items have been downloaded 296,074 times (Figure 2). The median download statistics for the last 12 months are 3,855 per month, and they are continuously increasing.

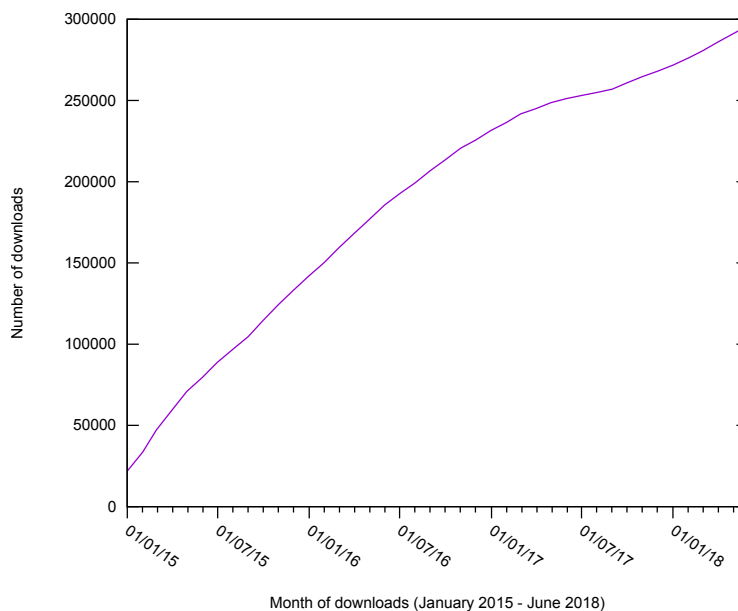


Figure 2: Cumulative download statistics (N = 587)

Usage statistics per item: the analysis of the usage statistics reveals that all deposited Master's theses, without exception, have been accessed and downloaded (Figure 3).

Figure 3 shows that the differences of usage among the Master's theses are very important, ranging from 10 to more than 60,000 downloads, with a median of 43. At least two factors may explain these differences:

1. The date of deposit and publishing: the items with more than 10,000 downloads have all been deposited between 2009 and 2011, while the items with low download figures are all from 2017 and 2018.
2. The topics: probably, theses with "hot issues" will be more downloaded. But the real impact of the content on usage statistics is difficult to assess and surely affected by other variables like the quality of indexing and referencing.

Information of interesting and recent deposits is published via social media, e.g. Twitter, Facebook and LinkedIn; metadata, links and the full text can be discovered on sites like Scoop.it, CiteULike or portals for students. However, without Digital Object Identifier minting and a more sophisticated altmetrics tool it is impossible to assess the real impact of the deposits inside and outside of the academic communities.

Where do the successful download requests come from? The *DUMAS* repository identifies the country of nearly 80 per cent downloads.

In total, 30 per cent successful download requests come from France, another 13 per cent are from the three French-speaking Maghreb countries Morocco, Algeria and

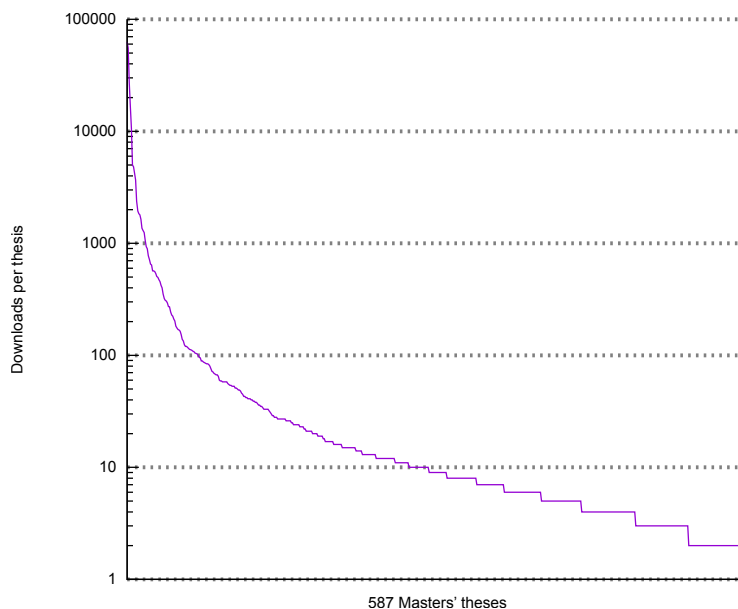


Figure 3: Download statistics per thesis (N = 587, logarithmic scale)

Tunisia. The USA represents 3 per cent, Germany and Canada 2 per cent each, the UK, South Africa and China 1 per cent each. The member states of the so-called “Francophonie” represent 60 per cent of the download requests, i.e. 75 per cent of all geographically identified downloads. However, the collection’s outreach is global and not limited to the French-speaking world: the complete download list contains 186 countries from all continents and regions, including a long list of countries and states like Mongolia, Trinidad and Tobago, Surinam and the Vatican.

4 Discussion

MemorySID is a faculty-based project in social sciences and humanities, launched and undertaken by the SID department staff to increase the outreach and impact of the Lille LIS Master program. From the beginning, it raised questions and problems that had to be fixed by the faculty, in partnership with the academic library, the IT department and the host of the national open repository. Four issues appear crucial for the future development of open and unrestricted dissemination of Master’s theses.

4.1 IP clearance

Intellectual property (IP) is an important issue. Open access to Master’s theses via open repositories means uploading and publishing, and both need authorization from the

rights holder and, at least in France where dissertations are considered as administrative proof for the diploma, also from the institution.

Once the student has left the Master program, this IP clearance is difficult to obtain and the risk is high that time produces an increasing number of “orphan theses”. The best moment is at the defense (*viva*), when the student presents the results of her or his research work and internship. This is the opportunity for the student and the jury to sign the authorization for open access publishing, as part of the minutes of the thesis defense. The signed minutes are archived by the department and serve as proof for the dissemination via the DMS and the open repository.

But this is not enough. Some theses contain sensitive information or personal data while others include material protected by third party rights (photos, maps, etc.). Also it may be appropriate to request a third authorization from the professional tutor of the internship. If the student cannot guarantee full rights clearance for included material, it will be prudent either to drop this material or to exclude the thesis from open access.

At present, of all theses preserved in the DMS, four out of five are freely shared on the internet, while 12 per cent are restricted to on-campus access via the intranet and 8 per cent are in the DMS but not available at all (Figure 4).

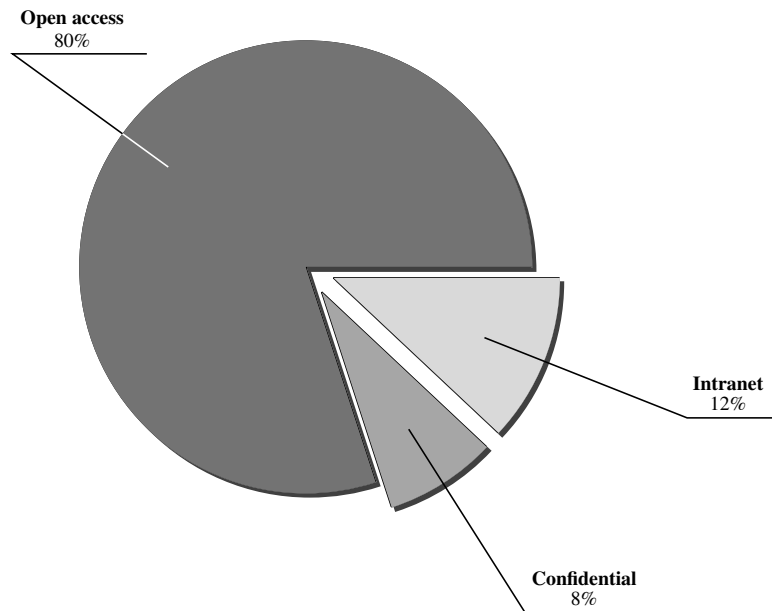


Figure 4: Status of Master's theses in the DMS (N = 774)

The key factor for the IP-related process is the faculty support, i.e. the personal and professional commitment to open scholarship which implies two “upstream actions”:

1. A clear and explicit communication about open repositories, open scholarship and sharing.

2. Guidelines about how to write a Master's thesis, how to use protected third-party material, how to handle confidential and sensitive information, personal data (privacy issues), etc.

Sometimes, fears are expressed that unrestricted dissemination of Master's theses may increase the risk of plagiarism. Indeed, this "collateral damage" cannot be excluded but on the other hand, open access to Master's theses will improve the performance of anti-plagiarism software.

4.2 Quality, metadata and identifiers

The second issue is quality. While all Master's theses should be preserved in the intranet DMS, only the best ones should be disseminated in open access, for two reasons:

1. Insofar, as the collection provides an excellent showcase and high visibility for the Master program, the dissemination of poor papers should be avoided.
2. To foster the potential for TDM, particular attention should be paid to the quality of metadata.

The quality issue has a paradoxical effect: if the deposit is limited to the best ("outstanding") Master's theses, only a small part will be available, reducing their diversity and representativeness. On the other hand, a selective "institutional quality label" will increase their value for the human reader and the machine.

At Lille, the threshold for online dissemination is set at 14 marks out of 20, which is similar to the US A and B grades. This score is the result of collegiate evaluation by at least two scholars, applying the usual criteria, e.g. quality of the analysis, independent thinking, presentation, structure, language and style and references. Nevertheless, the score is only one criteria and the final decision is taken at the defense (*viva*), by the jury.

The format and quality of metadata remains a challenge. Standard metadata improve findability and interoperability. Both systems, the DMS (Nuxeo) as well as *DUMAS*, make use of the Dublin Core (DC) elements. To limit inconsistent indexing, some elements were refined, such as rights (level of availability), source (Master program, level, date of the *viva*), description (author's abstract) and filename.

The *DUMAS* metadata schema follows the Text Encoding Initiative format of the French national HAL repository with some specific elements for Master's theses, e.g. the name of the Master program with its subdivisions, specialties and options. The problem is that the DMS and *DUMAS* follow different schemas, with different refinements and values, and that their metadata silos are not connected. This means the creation of two different metadata sets for the same document, one for the DMS and the other for the repository. The lack of interconnection and interoperability is a source of inconsistency and errors, causes double workload and requires more post-deposit curating than necessary. A three-level routine reduces the metadata heterogeneity and error level:

1. Inconsistencies and anomalies are rejected by the *DUMAS* administrator during the validation.

2. The index of authors and directors' names is periodically controlled by the local administrator.
3. The tag cloud is checked from time to time for synonymy, variants etc.

A last comment on unique identifiers: *DUMAS* systematically attributes a specific HAL identifier to each new deposit. To track impact on social media (altmetrics), it may be interesting to consider additionally digital object identifier minting.

4.3 Text and data mining

TDM tools and methods are already applied to Master's theses, especially to assess research topics, methodologies and funding trends. In Lille, a research team started to mine PhD and Master's theses in exploratory research projects, to identify specific topics in different disciplines like Agronomics, Cultural Studies or Information Sciences, and geographical information related to the concept of territory (Kergosien et al., 2018).

These studies confirm that the diversity and volume of Master's theses bear interesting potential for TDM and can produce relevant new knowledge about research topics, methods, people, etc. As the output of TDM is conditioned by the quality of the corpus, this means that special attention should be paid to the quality of titles, abstracts, key words and the structuration, annotation, indexing and open formats, whenever possible.

In an uncertain legal environment of TDM like in France, it would be appropriate and helpful to publish Master's theses under a liberal open license (CC-BY or CC-BY-SA) to foster their reuse and exploitation; however, this remains an exception up to now.

4.4 Preservation

Master's theses are part of what is called gray literature, not easy to identify, not always well controlled by catalogs and databases and often at risk of disappearance (Schöpfel; Farace, 2010). In France, Master's theses must be conserved for at least five years, as a proof of graduation. *MemorySID* adopts a three-level strategy to guarantee long-term preservation at least for the best Master's theses:

- Remaining print copies are preserved by the faculty administration during the obligatory five-year period, in the facilities of the SID Department.
- All digital copies are preserved in the university DMS and maintained by the IT Department for at least 5-10 years, the usual lifespan of documents in intranet.
- As for the selection of best and non-confidential Master's theses, they are preserved in the *DUMAS* open repository. *DUMAS* is part of the French national open scholarship infrastructure HAL which is backed up by the French National Computing Center for Higher Education⁹ specialized in long-term preservation of digital library resources, research data etc.

⁹CINES <https://www.cines.fr/en/>

Thus, the *DUMAS* solution includes outsourcing of the long-term availability of Master's theses, in a dark archive in the public academic cloud which can be used as a failsafe during disaster recovery.

5 Conclusion

The environment of open scholarship provides excellent opportunities for a new approach to the management of Master's theses. Based on the Lille project and other studies, the following matrix sums up some key factors for the contribution of Master's theses to open scholarship:

- **Strength:** their diversity, richness and representativeness make Master's theses interesting for reading, content mining, citation analysis, etc. which is an important feature of the emerging landscape of open science. Also, they contribute to the visibility and promotion of the institution and the Master program.
- **Weakness:** the main problem is related to their unequal quality. Sensitive content (personal data, corporate information etc.) is another problem.
- **Opportunity:** the European and French open science policy creates a favorable environment for the dissemination of Master's theses on open repositories; this is emphasized by the institutions' interest to promote and valorize their Master programs.
- **Threat:** similar to PhD dissertations (Schöpfel; Prost, 2013), the major threat is access restriction because of intellectual property, i.e. no authorization by students. Restrictive policies of institutional repositories ("scientists only") and lack of interest for undergraduate work are other barriers.

A collection of Master's theses in the public academic cloud is a unique showcase for the academic excellence of Master programs, of teaching and research. The download statistics show that its impact and visibility are real and significant. It can also produce a rich and representative corpus for TDM and scientometrics. However, this potential is conditioned by a couple of key variables, such as a selective approach, long-term preservation and free dissemination on open repositories. It is only when these conditions are met, that Master's theses fully contribute to open scholarship.

References

BØHN, Henrik; HASSELGREEN, Angela, 2011. *The assessment of master's theses: findings of a survey, spring 2011*. Oslo. Available also from: <http://ekstern.alt.hist.no/%E2%88%BCeivindnt/nfe/Assessment-master11.pdf>. National Academic Council for English Studies.

- CARVALHO, José A.; PEREIRA, Luísa; LARANJEIRA, Rómína, 2017. Writing a master dissertation – students’ difficulties and coping strategies. In: *Writing a master dissertation – students’ difficulties and coping strategies. 9th Conference of the European Association for Teaching Academic Writing - EATAW 2017*. Available also from: <http://hdl.handle.net/1822/46838>.
- CHAUVIN, Sophie; GALLETZOT, Gabriel; SCHÖPFEL, Joachim, 2010. Les mémoires de Master dans les archives ouvertes. In: *Les mémoires de Master dans les archives ouvertes. CIUEN 2010, Strasbourg, 14-16 juin 2010*, pp. 293–310. Available also from: http://archivesic.ccsd.cnrs.fr/sic_00497390/fr/.
- GILLET, Dominique; BÉRARD, Alain; FAIVRE, Olivier; SARRAZY, Bernard, 2013. Measuring the international spreading of the knowledge produced by French dental master theses. *European Journal of Dental Education*. Vol. 17, no. 1, pp. 26–29. ISSN 13965883. Available from DOI: 10.1111/eje.12003.
- KERGOSIEN, Éric et al., 2018. Automatic Identification of Research Fields in Scientific Papers. In: *Automatic Identification of Research Fields in Scientific Papers. Proceedings of the Eleventh conference on International Language Resources and Evaluation (LREC’18)* [online]. Miyazaki, Japan: ELRA, pp. 1902–1907 [visited on 2018-05-17]. Available from: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/1067.pdf>.
- MANN, Vivien, 2010. *La valorisation des mémoires d’étudiants sur une archive ouverte institutionnelle : un fonds et un contexte spécifiques*. Available also from: https://dumas.ccsd.cnrs.fr/MEM-UNIV-LILLESID/mem_00502624. Master’s Thesis. Université de Lille Sciences humaines et sociales.
- MARTENS, Bob, 2011. Enhanced Accessibility of Master Theses Related to Real Estate. In: *Enhanced Accessibility of Master Theses Related to Real Estate. 18th Annual European Real Estate Society Conference. ERES: Conference. Eindhoven, The Netherlands, 2011*. European Real Estate Society (ERES). Available from DOI: RePEc:arz:wpaper:eres2011_80.
- NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE, 2017. *Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities*. Washington DC: The National Academies Press. Available also from: <https://doi.org/10.17226/24622>.
- SCHÖPFEL, Joachim; FARACE, Dominic J., 2010. Grey Literature. In: BATES, Marcia J.; MAACK, Mary N. (eds.). *Encyclopedia of Library and Information Sciences, Third Edition*. London: CRC Press, pp. 2029–2039. Available from DOI: 10.1081/e-elis3-120043732.
- SCHÖPFEL, Joachim; PROST, Hélène, 2013. Degrees of Secrecy in an Open Environment. The Case of Electronic Theses and Dissertations. *ESSACHESS - Journal for Communication Studies*. Vol. 6, no. 2. Available also from: <http://www.essachess.com/index.php/jcs/article/view/214>.

- SELIGER, Marja, 2015. Research Meets Practice in Master's Theses. In: VANDE ZANDE, Robin; BOHEMIA, Erik; DIGRANES, Ingvild (eds.). *Learn x Design. Proceedings of the 3rd International Conference for Design Education Researchers. 28–30 June 2015, Chicago, Il, USA. Volume 1*. Spotswood, NJ: DESIGN-ED Coalition, pp. 131–141. Available also from: https://www.researchgate.net/profile/Erik_Bohemia/publication/279182704_Proceedings_of_the_3rd_International_Conference_for_Design_Education_Researchers_volume_1/links/558cedf808aee43bf6ae4b5f.pdf#page=155.
- SUBER, Peter, 2012. *Open access*. Cambridge, Mass: MIT Press. Essential Knowledge series. ISBN 978-0-262-51763-8. Available also from: <http://mitpress.mit.edu/books/open-access>.
- VANACKER, Sylvain, 2017. *Processus de diffusion des mémoires de Master. Littérature grise et accessibilité : un enjeu majeur dans le domaine de la recherche*. Available also from: <https://dumas.ccsd.cnrs.fr/dumas-01558312>. Master's Thesis. Université de Lille Sciences humaines et sociales.
- VELETSIANOS, George; KIMMONS, Royce, 2012. Assumptions and challenges of open scholarship. *The International Review of Research in Open and Distributed Learning*. Vol. 13, no. 4, pp. 166–189. ISSN 1492-3831. Available from DOI: 10.19173/irrodl.v13i4.1313.
- XIA, Jingfeng; OPPERMAN, David B., 2010. Current Trends in Institutional Repositories for Institutions Offering Master's and Baccalaureate Degrees. *Serials Review*. Vol. 36, no. 1, pp. 10–18. ISSN 00987913. Available from DOI: 10.1016/j.serrev.2009.10.003.

Further reading

- ÇELIKGÜN, Bahtiyar, 2016. Publication rates of audiology master and doctoral theses in peer-reviewed journals. *The Turkish Journal of Ear Nose and Throat*. Vol. 26, no. 5, pp. 276–282. ISSN 13007475. Available from DOI: 10.5606/kbbihtisas.2016.03780.
- DUDZIAK, Reiner; MOHR, Dirk, 2014. Bachelor and Master Theses in mechatronics engineering study courses - Experiences and perspectives. In: *Bachelor and Master Theses in mechatronics engineering study courses - Experiences and perspectives. 15th International Workshop on Research and Education in Mechatronics (REM), September 2014, El Gouna, Red Sea, Egypt*. El Gouna, Red Sea, Egypt: IEEE, pp. 1–8. ISBN 978-1-4799-3029-6. Available from DOI: 10.1109/rem.2014.6920441.
- ELBAN, Mehmet, 2017. Investigation of the History Education Researches in Turkey in Terms of Some Variables (Master's Theses and Dissertations Sample). *Universal Journal of Educational Research*. Vol. 5, no. 9, pp. 1529–1534. ISSN 2332-3205. Available from DOI: 10.13189/ujer.2017.050910.

- JUZNIC, Primož, 2010. Grey Literature produced and published by Universities: A Case for ETDs. In: FARACE, Dominic; SCHÖPFEL, Joachim (eds.). *Grey Literature in Library and Information Studies*. München: De Gruyter Saur, pp. 39–51.
- LI, Xing-li; WU, Wei-hua; ZHANG, Jia-sheng, 2013. Citation Quantitative Analysis on the Outstanding Master Dissertation – For the Construction of Key Disciplines Collections Strategies. *Journal of Hebei United University (Social Science Edition)*. No. 2. Available also from: http://en.cnki.com.cn/Article_en/CJFDTOTAL-HLXB201302014.htm.
- NORDRUM, Lene, 2014. A comparison of data commentary in chemical engineering research papers and master theses – exploring the intersemiosis of written and visual material for applied purposes. In: *A comparison of data commentary in chemical engineering research papers and master theses – exploring the intersemiosis of written and visual material for applied purposes*. ICAME 35, Nottingham, United Kingdom. Available also from: <https://lucris.lub.lu.se/ws/files/6049083/5038483.pdf>.
- PEPONAKIS, Manolis, 2013. Libraries' Metadata as Data in the Era of the Semantic Web: Modeling a Repository of Master Theses and PhD Dissertations for the Web of Data. *Journal of Library Metadata*. Vol. 13, no. 4, pp. 330–348. Available from DOI: 10.1080/19386389.2013.846618.
- SALAWU, Abiodun; OYERO, Olusola; MOYO, Mathew; MOYO, Rachel, 2016. A survey of research foci and paradigms in media and communication Master's dissertations and doctoral theses in South Africa. *Communicatio*. Vol. 42, no. 1, pp. 136–154. Available from DOI: 10.1080/02500167.2016.1140667.
- SINHA, Manoj K.; SINGHA, Sudipta, 2016. Citation Analysis of Master of Library and Information Science Dissertations Submitted to Department of Library and Information Science during 2012-2013. *International Journal of Library, Information, Network and Knowledge*. Vol. 1, no. 2, pp. 6.1–17. ISSN 2455-5207. Available also from: <http://ssarsc.org/volumes/ijlink/volume1/paper-6-volume1.pdf>.
- TANINMIŞ, Gamze E.; ORHAN, Şebnem Y., 2010. Master theses and doctoral dissertations conducted between 1986 and 2009 in music education. *Procedia - Social and Behavioral Sciences*. Vol. 2, no. 2, pp. 5051–5058. ISSN 18770428. Available from DOI: 10.1016/j.sbspro.2010.03.820.
- TEIXEIRA JUNIOR, Paulo Roberto; RIOS, Monica Piccione Gomes, 2017. Dez anos de SINAES: um mapeamento de teses e dissertações defendidas no período 2004 - 2014. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* [online]. Vol. 22, no. 3, pp. 793–816 [visited on 2018-10-24]. ISSN 1414-4077. Available from DOI: 10.1590/s1414-40772017000300012.
- XU, Xiao-Li; FENG, Qing, 2011. Discussion on the Strategies for the Construction of Key Discipline Collection Based on Citation Analysis of Master Dissertation – Taking the Discipline of Textile Engineering as an Example. *Sci-Tech Information Development & Economy*. No. 29. Available also from: http://en.cnki.com.cn/Article_en/CJFDTOTAL-KJQB201129014.htm.