Abstract: Focused on the collection of theses and dissertations, the aim of the following study is to look to the scientific production of ULisboa and analyse the international visibility of it through the statistical report of these two institutional repositories for the last 10 years (2010-2019). For this purpose, the total number of thesis and dissertations of the ULisboa, archived in both repositories, were collected, for the period 2010-2019 (10 years of implementation of the institutional repositories). Usage statistics (views and downloads) were obtained as well as the home countries of these accesses. A descriptive analysis was performed. Results showed that the number of thesis and dissertation archived in the repositories has been growing, as well as their use (views and downloads) is very significant in both repositories.

Keywords: Open science; open access; institutional repositories; university; Portugal

1. Introduction

The Institutional Repositories (IR) of universities are initiatives that integrate the concept of Open Science, enhancing the free circulation of academic and scientific production. The current University of Lisbon (ULisboa) results from the fusion of the former University of Lisbon and Technical University of Lisbon, in 2013. Currently ULisboa comprises 18 schools (Faculties and Institutes) and 57 Research Units. Since the fusion, University of Lisbon still have two independent IR based on DSpace – the Repositório.UL
Focused on the collection of theses and dissertations, the aim of the following study is to look at the scientific production of ULisboa as a whole and analyse the international visibility (ADBS, 2005) of it through the statistical report of these two IR for the last 10 years (2010-2019).

For this purpose, the total number of thesis and dissertations of the schools of ULisboa archived in these repositories nowadays, is collected. The chronologic scope is the period from 2010 to 2019 (10 years of implementation of the IR). Usage statistics (views and downloads) are obtained as well as the home countries of these accesses. All the collected data is analysed.

The self-archive policies of the University of Lisbon express as major goals, to gather and organize systematically the set of intellectual, academic, and scientific production; to disseminate, provide access and more visibility to research developed; to improve monitoring, evaluation and management of research and teaching activities and to promote the enhancement and preservation of the intellectual and cultural heritage of this University.

To comply with these policies and to reinforce ULisboa’s position as a producer of scientific knowledge and avoid data dispersion, the merge of the two repositories becomes essential.

2. Literature review and background

An institutional repository is an online digital archive that organizes, preserves, and provides access to an institution's educational, academic and research output. For more than a decade, academic libraries started creating these repositories, which have become a fundamental element for dissemination and communication in academia and in the context of educational institutions, consolidating the understanding of their concept and application (Silva & Corujo, 2018).

The repositories have been as allies to promote and safeguard the scientific production developed by its faculty and students, being a requirement of funding agencies to give visibility to funded research (Coneglian, Vidotti, Martínez, da Costa, & Segundo, 2019) and a means of affirming national public policies around open access (Weitzel, 2019). Meanwhile, thematic repositories were created, those dedicated to specific themes; institutional, with the purpose of organizing the collection of an institution; research data, focused on storing the raw material of scientific research and, more specifically, the aggregating repositories by type of documents, such as those of Theses and Dissertations, which will usually absorb the contents to IR.

IR maintained by librarians, by gathering and making information available in open access, respond to the citizen who seeks to learn, provide advanced knowledge to professionals, support political decision-making, and enable researchers to monitor cutting-edge scientific production through academic and scientific studies and works found there, including theses and dissertations,
ultimately contributing to a global movement for the internationalization of knowledge.

At the same time, “the creation of a digital repository, that centralizes the materials produced in the institution, which organizes and allows their search and information retrieval and that enables the authors themselves to feed the database of the digital repository, is an asset in the academic and scientific context because it promotes, preserves and disseminates the scientific production of a university” (Lopes, Lopes, & Campos, 2010: 9). Open Access, aligned with Open Science (Portugal, MCTES, 2016) provides free and public access for any user to read, download, copy, print, distribute, search or link to full text of articles, respecting the legal rights of authorship, it can bring researchers from distant countries closer to international cooperation. This integration will certainly result in benefits for all nations. Regardless of the social and cultural reality, most countries around the world suffer from the lack of visibility of the knowledge generated by their researchers (Melo, Sampaio, & Pires, 2008). At the same time “the credibility of institutions devoted to research is demonstrated by its publications and availability of its scientific production” (Lopes, Lopes, & Campos, 2010: 8). It is of fundamental importance that, through digital repositories, Higher Education Institutions gather and make available the set of scientific publications, contributing to the growing impact of the research developed in the institution, increasing its visibility and that of those who work there, and guaranteeing the preservation of Portuguese science heritage and memory (Carvalho, 2018).

As of December 2013, and according to OpenDOAR - Directory of Open Access Repositories (https://v2.sherpa.ac.uk/opendoar) data, there were approximately 2,100 IR worldwide (Palmer, 2014). In May 2018, the number of IR increased to just over 3,000. The emergence of specific repositories for theses and dissertations demonstrates how important these documents are for the dissemination and reuse of scientific and technical knowledge. Some of the most internationally recognized ETD (electronic theses and dissertations) repositories are: OATD (Wani, 2019), and NDLTD (Yotis, 2008). The OATD aims to be the best possible resource for finding open access graduate theses and dissertations published around the world. Metadata (information about the theses) comes from over 1100 colleges, universities, and research institutions. Currently, OATD indexes 5,039,477 theses and dissertations (https://oatd.org). NDLTD, the Networked Digital Library of Theses and Dissertations (http://www.ndltd.org), provides access for electronic theses and dissertations (ETDs). According to Patra & Das (2019), there are about 5,572,750 ETDs deposited in the NDLTD’s global open repository. Among the contributed countries, the top ten repository countries are UK, Taiwan, France, Germany, India, Portugal, Japan, Finland, China, and Brazil regarding uploaded resources.

However, the availability of theses and dissertations in repositories does not seem to be sufficient to guarantee widespread access to their contents. Moskaleva & Kettler (2019:160) refer that “even openly available content can
be difficult to discover and retrieve for interested readers, depending on the language, quality of metadata, and indexing”. The authors conclude that the availability of English-language metadata becomes crucial to increase the visibility of dissertations, underling that visibility and discoverability are key points to promote the scientific outputs of universities. They also state that “readership and citations help to promote the institution’s profile and result in improved ranking positions and, ultimately, in enhanced attention by potential new students as well as funders” (Moskaleva & Kettler, 2019: 163).

Some studies (Baro & Otiode, 2014; Malapela, Chisita, Hadebe, 2019; Stanton, & Liew, 2011) report that the scientific outputs and visibility of scholarly research produced by the higher education institutions, specifically focus on the ETDs, explaining that an IR creates an enabling environment for scholarly publishing and makes the research productivity of a particular institution more visible globally. There are also perceived benefits of enhanced exposure and potential for sharing worldwide research (Ahmed, Alreyaee, & Rahman, 2014).

Schöpfel (2013: 1) recommends that “IR services for the dissemination of ETDs should be flexible, with a capacity for rapid adaptation; the software should be user-friendly and reliable, perhaps also open to other service providers and/or integrated in another service environment, such as extended (distance) learning”. Novak & Day (2018) in their research about the conditions for IR success, underlie that it is necessary to effectively position it within the context of the libraries’ collections, research support, and scholarly communication services. The authors state that a major component of this process is re-examining the fundamental aims of the IR and aligning it to the libraries and the campus strategic goals. Additionally, Casella (2010), presenting an asset of items and criteria for IR evaluation, states that the variety, the richness, and the completeness of the collections deposited in the repository and the number of value-added services developed for authors are strategic components of a successful repository.

On the other hand, Serrano (2018) supports the idea that IR must be an important part in the formula for valuing the production and data of scientific and technical research, and for the evaluation of science in general, although it is important that we will be considered in broader methods of scientific production, beyond traditional metrics.

The Portuguese Foundation for Science and Technology (Portugal, MCTES, 2020) states that this country is recognized in Europe for its innovative policy on Open Access, as well as by the Scientific Open Access Repositories of Portugal - RCAAP (https://www.rcaap.pt). This national aggregator of IR has an essential role to increase the visibility of Portuguese universities as well as the rapid expansion of institutional repositories nationwide.

Carvalho and Lopes (2018: 183-184) explain the emergence and relevance of this content aggregator: “the year 2013 was still marked by two events of remarkable importance in the national panorama of Open Access.

On the one hand, the RCAAP network has become the instrument enshrined in the law to comply with the legal deposit of doctoral theses, master’s dissertations, and other scientific works. Indeed, point 1 of article 50 of Decree...
Law 115/2013 of 7 August 2013, states that “The doctoral theses, the works provided for in paragraphs a) and b) of paragraph 2 of article 31 and dissertations by master's degrees are subject to the mandatory deposit of a digital copy in a repository that is part of the network of the Scientific Repository of Open Access in Portugal, operated by the Foundation for Science and Technology. This legislative initiative aims at the treatment and preservation of doctoral theses, dissertations from master's degree and other scientific works, as well as the dissemination, under Open Access regime, of production that is not subject to restrictions or embargoes in the network of repositories and is of great importance to increase the critical mass and dynamics of RCAAP”.

The current University of Lisbon (ULisboa) resulted from the merger of the previous University of Lisbon and Technical University of Lisbon, exactly in 2013 and has two repositories in DSpace - Repositório.UL and UTL Repository - which aim to gather, organize, disseminate, and preserve the academic scientific production of the University of Lisbon. They include collections of theses and dissertations from the entire University of Lisbon, the focus of this work. Both repositories are integrated into RCAAP which, in addition to national repositories, also aggregates a set of Brazilian repositories through the OASIS.br Portal (http://oasisbr.ibict.br).

3. Objectives
This study aims to demonstrate how institutional repositories contribute to the internationalization of scientific knowledge. Focused on the contribution of theses and dissertations to the globalization of science, perceived by the usage statistics of these collections.

4. Methods
The collections of theses and dissertations archived in the two DSpace repositories of the University of Lisbon were identified. The total number of theses and dissertations, from all communities (schools, institutes, and university rectory) was collected for the period of 2010-2019. Usage statistics related to the number of views (made through repositories) and downloads (made through repositories and other platforms, e.g., search engines) and the home countries of these accesses were gathered. Data collection took place on April 18, 2020.

A descriptive analysis was performed, concerning the schools listed in Table 1, showing the scope of scientific areas covered (Life and Health Sciences, Science and Technology, and Social Sciences and Humanities).

<table>
<thead>
<tr>
<th>ULISBOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initials</td>
</tr>
<tr>
<td>FA</td>
</tr>
</tbody>
</table>
5. Results

In the period 2010-2019, a total of 30863 theses and dissertations were archived in ULisboa DSpace Repositories. Figure 1 shows the distribution of theses and dissertations by schools and institutes.
Different dimensions of schools and institutes, in the number of students and courses, has a direct relationship with the number of theses and dissertations defended and deposited. However, it is important to understand these numbers and their global impact, as a whole or by scientific areas, which is substantiated by the total number of views and downloads.

Figure 2 indicate the distribution by scientific area of the schools and institutes communities with existing theses and dissertation in both repositories.
The number of theses and dissertations by scientific area seems to follow the same distribution of the number of schools and institutions by scientific areas (Figure 3).

The same trend occurs when analysing the number of downloads and views by scientific area (Figure 4). In figure 4, it is possible to compare both distribution curve of the number of theses and dissertations, and downloads and views by scientific area.
Regarding usage statistics, in 10 years, downloads reached 19,975,293 and views reached 9,006,808 (Table 2). The major number of downloads (comparing with views) is justified by the diversity of available platforms that allow users to search and download without going through repositories.

Table 2. Total of downloads and views

<table>
<thead>
<tr>
<th>Year</th>
<th>Downloads</th>
<th>Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>367328</td>
<td>200171</td>
</tr>
<tr>
<td>2011</td>
<td>1027561</td>
<td>461329</td>
</tr>
<tr>
<td>2012</td>
<td>1669261</td>
<td>691573</td>
</tr>
<tr>
<td>2013</td>
<td>2459786</td>
<td>1082665</td>
</tr>
<tr>
<td>2014</td>
<td>3847076</td>
<td>1838277</td>
</tr>
<tr>
<td>2015</td>
<td>3213723</td>
<td>1749775</td>
</tr>
<tr>
<td>2016</td>
<td>1467390</td>
<td>666651</td>
</tr>
<tr>
<td>2017</td>
<td>1456393</td>
<td>704566</td>
</tr>
</tbody>
</table>
For a better visualization of these results, Figure 5 is presented, which shows the total of downloads and views for theses and dissertations in the period 2010-2019.

![Downloads and Views (2010-2019)](image)

**Figure 5. Total of downloads and views for theses and dissertations in the period 2010-2019** (Source: Repositório.UL and UTL Repository, on April 18, 2020)

The total of downloads and views for theses and dissertations in the period 2010-2019 demonstrates consistent increase until 2014. However, in 2015 there was a slight drop. This does not exactly reflect the decrease in interest in consulting these documents, but rather the irrecoverable loss of statistical data due to a system update.

Later, in 2016, there was also a sharp drop, which this time is justified by the implementation of a robotic access control that made the usage data more reliable. Thus, the data from the most recent years appear to be more consistent, revealing regular access and a new gradual and solid increase in views and downloads of theses and dissertations.

It is also important to consider the countries of origin of the searches, to understand the scope and reach out of the dissemination of this information. In the results analyses were identified 228 countries or regions as origin of downloads and views.

The top 10 countries represent 93.6% and 94.3% of total downloads and views,
respectively (Table 3 and 4). Portugal, Brazil, and United States, with a total of 79% of downloads and 78% of views, are in the Top 3 of downloads and views ranking, followed by China.

Table 3. Top 10 - Downloads by Country

<table>
<thead>
<tr>
<th>Countries</th>
<th>Nr.</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portugal</td>
<td>7370935</td>
<td>35,72%</td>
</tr>
<tr>
<td>Brazil</td>
<td>5561921</td>
<td>26,95%</td>
</tr>
<tr>
<td>United States</td>
<td>3381053</td>
<td>16,38%</td>
</tr>
<tr>
<td>China</td>
<td>648509</td>
<td>3,14%</td>
</tr>
<tr>
<td>Angola</td>
<td>425121</td>
<td>2,06%</td>
</tr>
<tr>
<td>Mozambique</td>
<td>392252</td>
<td>1,90%</td>
</tr>
<tr>
<td>Germany</td>
<td>285078</td>
<td>1,38%</td>
</tr>
<tr>
<td>Mexico</td>
<td>275056</td>
<td>1,33%</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>263010</td>
<td>1,27%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>128489</td>
<td>0,62%</td>
</tr>
<tr>
<td>N/A</td>
<td>587238</td>
<td>2,85%</td>
</tr>
</tbody>
</table>

(Source: Repositório.UL and UTL Repository, on April 18, 2020)

Note that Angola and Mozambique, Portuguese-speaking African countries, appear in the Top 10 both in terms of downloads and views. Finally, N/A correspond to home countries not identified (data not available, satellite providers and anonymous proxy).

Table 4. Top 10 - Views by Country (2010-2020)

<table>
<thead>
<tr>
<th>Countries</th>
<th>Nr.</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portugal</td>
<td>3619142</td>
<td>37,59%</td>
</tr>
<tr>
<td>Brazil</td>
<td>2057495</td>
<td>21,37%</td>
</tr>
<tr>
<td>United States</td>
<td>1860612</td>
<td>19,32%</td>
</tr>
<tr>
<td>China</td>
<td>332201</td>
<td>3,45%</td>
</tr>
<tr>
<td>Poland</td>
<td>322802</td>
<td>3,35%</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>229403</td>
<td>2,38%</td>
</tr>
<tr>
<td>France</td>
<td>131808</td>
<td>1,37%</td>
</tr>
</tbody>
</table>
These results are consistent with the literature review, where a correlation was shown, demonstrating the influence of the language of the documents deposited with their access, represented here by the Portuguese language and by countries that speak the same language. It should also be noted that the existence of theses and dissertations in English seems to favor views and downloads by non-Portuguese speaking countries all over the world.

6. Discussion

Data reveal the importance of Repositories in the visibility of these collections, over time, and the potential impact on the national and international scientific community. Its aggregation in the RCAAP network is a key factor. The connection between RCAAP network and OASIS.br project as well as the Portuguese language justify the high percentage of downloads and views from Brazil and from Portuguese-speaking African countries, mainly Angola and Mozambique, which are among the top 10 countries of origin of downloads and views. As mentioned in the literature review, despite cultural differences, repositories can bring together distant countries, promoting international cooperation and bringing researchers closer together, benefiting nations. And this is more visible when it comes to the same language, in this case, Portuguese.

Considering usage data by year, there is a slight decrease in 2015 that reflects the loss of statistical data due to a system upgrade. In 2016 there is a sharp decrease that is justified by the implementation of a robotic access control that made the usage data more reliable.

The existence of two Dspace repositories, with decentralized management, brings some limitations, namely the impossibility of getting data from the UTL repository during the development of this study. In the future, after merging the repositories, a detailed analysis of some parameters will be possible, such as the full-text availability of theses and dissertations or the total number of items archived annually.

In the future could be interesting to analyse if there is a different distribution of the origin of accesses, by scientific area.

7. Conclusions

Results show us that the number of thesis and dissertation archived in the repositories has been growing, as well as their use (views and downloads) is
very significant in both repositories. Considering that statistical data of the repositories are very consistent and that simultaneously it ensures the security regarding digital preservation, both the institution and the community (teachers, researchers, and students) benefit from the archive and dissemination of their academic publications through the repositories.

This increased use, demonstrated by the statistical data, should be considered as a valid argument to justify a bigger investment for the upgrade of the two current platforms that should be merged and transformed in the new IR for the ULisboa. That way, we could avoid duplications of archived documents, usage statistics and author profiles. Users would also benefit by having a unique access point to the scientific outputs of the entire university.

Studies like this one are useful to verify the scientific activity carried out in the institutions and their impact at national and international levels. It is also possible to obtain some indications and suggestions for optimized procedures, to improve future actions, monitoring the evolution resulting from possible interventions at the IR level and as a way of giving visibility to this platform.

The institution, teachers and students can benefit from the dissemination of their theses and dissertations in the repositories, and the community in general collects tangible benefits from access to the knowledge produced. Fulfilling the objectives expressed in the deposit policies, the universities where ULisboa is included can be important partners in the internationalization of science, namely through the availability of theses and dissertations in its IR, since, by disseminating scientific knowledge properly communicated to society, access to knowledge is provided, improving the quality of information available on the web.

In the case presented, the geographic dimension of the information reach is evident. The relevance of this study is based on the need to value and encourage Open Access in the country and in the world, through a more sustained knowledge of IR. This objective contributes to promote the development of science and to stimulate other studies that can collaborate for the promotion of national policies for repositories, including more effective actions to map the results and impacts on science of IR. Studies like this contribute to the visibility and promotion of scientific knowledge, with its dissemination on a large scale driving informed, participative, and capable citizens to decide based on quality scientific information. In the future more studies are needed to demonstrate how the use of institutional repositories of University of Lisbon contributes to enhance science dissemination, research and researchers’ visibility and promotes the access to knowledge, improving the quality of the information available on the web.

Acknowledgements
This work is funded by national funds through the Foundation for Science and Technology, under the project UIDB/00057/2020. This work is also financed by national funds from FCT – Foundation for Science and Technology under UID/CED/04107/2019 - Unidade de Investigação e Desenvolvimento em Educação e Formação, Instituto de
Educação, Universidade de Lisboa.

References


