

Examining Accesses by Country and Language

Ana Maria Beltran Pavani

Ana Carla Epitácio Mazzeto

Internal Research Reports

Number 16 | February 2011

Examining Accesses by Country and Language

Ana Maria Beltran Pavani

Ana Carla Epitácio Mazzeto

CREDITS

Publisher:

MAXWELL / LAMBDA-DEE

Sistema Maxwell / Laboratório de Automação de Museus, Bibliotecas Digitais e Arquivos

<http://www.maxwell.vrac.puc-rio.br/>

Organizers:

Alexandre Street de Aguiar

Delberis Araújo Lima

Cover:

Ana Cristina Costa Ribeiro

Examining Accesses by Country and Language

Ana M B Pavani and Ana C E Mazzeto

Departamento de Engenharia Elétrica

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

apavani@lambda.ele.puc-rio.br, amazeto@lambda.ele.puc-rio.br

Abstract

Portuguese is one of the most spoken languages in the world. Brazil is the largest Portuguese speaking population in the world and has the largest ETD collection written in this language. This article analyzes international accesses to PUC-Rio's ETD collection from June 2004 to April 2010. It is specially concerned with accesses from both Portuguese and Spanish speaking countries. Data from system access logs are examined and preliminary results are presented. This work shows the importance of language in the numbers of accesses in general. Future works will focus this relation with specific areas of knowledge.

Keywords: Accesses; ETDs in Portuguese; Portuguese speakers; Spanish speakers

1. Introduction

PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro (<http://www.puc-rio.br/>) is a university in Rio de Janeiro, Brazil. It has had graduate programs since the 1960s; they are in Humanities, Social Sciences and Science & Technology.

Its ETD program started in 2000 and ETDs became mandatory on August 01, 2002. The number of ETDs grew from 7, in December 2000, to over 5,100 in April 2010. Though the first ETD was published in May 2000, access logs have only been saved since June 2004. Access information is available for 71 full months.

BDTD – Biblioteca Digital de Teses e Dissertações (<http://bdtb.ibict.br/>) is the national consortium of Brazilian ETDs. Currently, 92 institutions contribute to it and the union catalog of metadata records holds over 130K records. Most ETDs are in Portuguese. Brazilian ETD records are available from catalogs worldwide since both cooperating institutions and BDTD have complied with the OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting since December 2002. PUC-Rio is one of the three founding members of BDTD.

Brazil is the only Portuguese speaking country in Latin America, has the largest Portuguese speaking population in the world and has the largest ETD collection in this language.

PUC-Rio's ETDs are accessed by users from all over the world and the numbers of monthly accesses are counted by the tens of thousands.

This paper presents the results of the preliminary analysis of the data contained in the access logs. It is not a complete work since examination of data yielded new ideas of threads to follow. This first step is concerned with Portuguese and Spanish speaking users or persons who can understand either language well enough to consult scholarly information.

The Maxwell System (<http://www.maxwell.lambda.ele.puc-rio.br/>) is an institutional repository that hosts and makes available the ETD collection. It has a set of applications that extract and group access logs to yield information of accesses to all digital contents. The authors decided to examine the logs to try to understand the importance of the ETD collection as a reference to speakers of the two languages.

Section 2 addresses a profile of Portuguese and Spanish speaking countries, while section 3 presents the decisions authors made to deal with the differences among countries. Section 4 describes some characteristics of the ETD collection that are necessary to understand the access numbers. Section 5 presents and interprets the numeric results. Section 6 is devoted to listing other institutions whose ETD systems were visited in an effort to find analogous results. Section 7 briefly mentions the next planned steps for the continuation of this work.

2. Portuguese and Spanish in the World

Portuguese is the seventh most spoken language in the world and the third Western language in the ranking [01]. Spanish correspondingly occupies the second and the first places, according to the same source. When Internet users are considered, Spanish and Portuguese are, respectively, the third and the sixth languages [02].

Portuguese is the official or one of the official languages of Angola, Brazil, Cape Verde, East Timor, Equatorial Guinea, Guinea-Bissau, Macau, Mozambique and Portugal. The analysis in this article ignored East Timor because the number of Portuguese speakers is very low; this

language had been banned during the Indonesian rule. The website of the State Library of Victoria, Australia, cites that the number of Portuguese speakers is estimated in less than 5% of the population [03] of a little over 1,100,00.

Spanish is the official or one of the official languages of Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Spain, Uruguay and Venezuela.

Some remarks about the languages and the countries can be made at this point:

- There are speakers of both languages spread all over the world, included the United States; this country has a very large Spanish speaking group.
- Equatorial Guinea has both languages as official languages, for this reason it was counted in both groups.
- The largest number of Portuguese speaking countries is in Africa, while Spanish is concentrated in Latin America.
- Brazil is the largest Portuguese speaking country with a population over 193 million and no second official language.
- The sizes, populations and living conditions of the countries in the two groups are very diverse.

An interesting aspect of Portuguese and Spanish is that they are quite similar languages; educated speakers of either language can manage, at least, to read the other. If scholarly information is considered, it is even easier. This is the reason this work addresses accesses from the two groups.

In the following sections, countries will be referred to as pt- and es-speaking.

3. Dealing with Countries Differences

This work considers data collected during 71 full months – from Jun 2004 to April 2010. In this time frame, users from 200 countries accessed ETDs. A way to deal with countries had to be defined.

The first decision was that Brazil and the United States would not be considered. Brazil is the home country of the collection and the United States has a very large population with speakers of many languages, besides many international students. For this reason, all accesses from these two countries were removed from the data. Two groups of countries were defined:

- "international group" – all countries except Brazil and the United States.
- "pt+es group" – all countries that have pt and/or es as an official language, except Brazil.

The objective of this work was to understand the significance of accesses from the "pt+es group". For this reason, two types of analysis were focused:

- The behavior of the "pt+es group" within the "international group".
- The behavior of pt- and es-speaking countries within the "pt+es group".

The countries in the "pt+es group" are very distinct in size, population, educational level and standard of living. For this reason, it was not expected that they would access ETDs in a similar way. It was necessary to try to identify expectations.

Factors that Influence the Numbers of Accesses to ETDs

The first factor that seems important is the size of the country in terms of population; no doubt this must be considered. The level of education of the population is a second important factor since ETDs are a highly specialized type of literature. A third factor that has to be considered is access to the Internet.

In order to assign numbers to expectations of accesses it was necessary to find a way to compute the numbers taking into consideration the three factors. Seeking data was not an easy task – levels of education were very difficult to find since the available data (for all countries) referred to literacy rates. Besides that, the other two factors had to be measured at the same time for all countries under consideration so that comparisons were valid.

Human Development Index

The United Nations Development Program (<http://www.undp.org/>) created the HDI – Human Development Index in 1990. It “introduced a new way of measuring human development by combining indicators of the life expectancy, education attainment and income into a composite human development index (HDI)” [04].

This index is available for most countries of the world. The latest update was published in October 2009 and covered data up to 2007.

The HDI was chosen to be used side by side with the country population because it contains information on living conditions (life expectancy and income) as well as education, it is measured the same way for all countries and the numbers had been generated at the same time (2009).

Human Development Index and Population

Once the HDI had been chosen, it became necessary to take into consideration the population of each country. Countries with the same HDI have distinct numbers of potential ETD readers depending on the sizes of their populations.

For this reason, an index was defined: **$I = \text{HDI} \times \text{population}$**

Table 1 shows countries, continents, HDIs, populations and indices I for pt-speaking countries. Table 2 displays the same but for es-speaking countries. Data were collected from the web portal of IBGE – Instituto Brasileiro de Geografia e Estatística [05] in March 2010.

Country	Continent	HDI	Population	I
Angola	Africa	0.561	18,497,632	10,377,171.55
Cape Verde	Africa	0.708	505,606	357,969.05
Equatorial Guinea	Africa	0.719	676,273	486,240.29
Guinea-Bissau	Africa	0.396	1,610,746	637,855.42
Macau	Asia	0.909	538,000	489,042.00
Mozambique	Africa	0.402	22,894,294	9,203,506.19
Portugal	Europe	0.909	10,707,130	9,732,781.17
Sao Tome e Principe	Africa	0.651	162,755	105,953.51
Total			55,592,436	31,390,519
Average		0.657	6,949,055	3,923,815

Table 1 – Countries, continents, HDIs, populations and indices I for pt-speaking countries

Country	Continent	HDI	Population	I
Argentina	Latin America	0.867	40,276,376	34,903,507.44
Bolivia	Latin America	0.729	9,862,860	7,190,024.94
Chile	Latin America	0.878	16,970,265	14,899,892.67
Colombia	Latin America	0.807	40,659,709	32,812,385.16
Costa Rica	Latin America	0.854	4,578,945	3,910,419.03
Cuba	Latin America	0.863	11,204,180	9,669,207.34
Dominican Rep	Latin America	0.777	10,090,151	7,840,047.33
Ecuador	Latin America	0.806	13,625,069	10,981,805.61
El Salvador	Latin America	0.747	6,163,050	4,603,798.35
Equatorial Guinea	Africa	0.719	676,273	486,240.29
Guatemala	Latin America	0.704	14,026,947	9,874,970.69
Honduras	Latin America	0.732	7,465,998	5,465,110.54
Mexico	Latin America	0.854	109,610,036	93,606,970.74
Nicaragua	Latin America	0.699	5,742,800	4,014,217.20
Panama	Latin America	0.840	3,453,898	2,901,274.32
Paraguay	Latin America	0.761	6,348,917	4,831,525.84
Peru	Latin America	0.806	29,164,883	23,506,895.70
Puerto Rico	Latin America	0.942	3,900,000	3,673,800.00
Spain	Europe	0.955	44,903,659	42,882,994.35
Uruguay	Latin America	0.865	3,360,854	2,907,138.71
Venezuela	Latin America	0.844	28,853,366	24,352,240.90
Total			410,938,236	345,314,467
Average		0.812	19,568,487	16,443,546

Table 2 – Countries, continents, HDIs, populations and indices I for es-speaking countries

The two tables show some significant differences between the two sets of countries. They are:

- There are 21 countries in the es- and 7 in the pt-speaking groups. Equatorial Guinea was considered in both, as stated before.
- The average HDI in table 2 is 23.6% higher than in the first.
- The population in table 2 is 7.4 times that of table 1.
- The total I index in table 2 is 11 times that of table 1.

The numbers show that there are potentially more accesses from the es-speaking group than from the other. Access logs showed that the real usage was the opposite.

4. Understanding the ETD Collection

The ETD collection of PUC-Rio has some characteristics that must be presented before accesses are considered. Currently, it holds over 5K ETDs of “different generations”.

The first generation is of ETDs defended between Apr 2000 and Aug 2002 – each one was published in one single file except for being too big (at the time bigger than 1.2 MB) or having some restricted content. In these cases, they were partitioned to allow saving the file on a diskette or temporarily protecting part of the work while the remaining would be immediately public. Except for this partial temporary restriction, they were also of public access since ETDs were voluntary and only authors who wanted to make their works public would participate in the ETD program..

The second generation is of ETDs defended after Aug 2002 – they are published in many files, one for the initial parts (title, copyright, tables of contents, etc), one for each chapter and one for references and appendices.

The third generation is of ETDs that came from retrospective digitization. They were made available with the same policy of the first generation, but they differ because they may require more partitions for being image pdf files. At the same time, the threshold for partition has been changed from 1.2 to 5 MB. Most of them are restricted because authors were impossible to contact and seek authorization for public access.

In the time frame of this analysis, the partitions numbers are:

- In June 2004, the average number of partitions per ETD was 5.9.
- In April 2010, the average number of partitions per ETD was 7.2.
- The average number of averages of partitions per ETD, from Jun 2004 to Mar 2010, is 6.9.

This is important because, since numbers of accesses are counted in tens of thousands, it is possible to associate 6,900 accesses to 1,000 ETDs being completely accessed.

5. Numbers of Accesses

The compilation of data and the subsequent analysis were based on two distinct axes. The first was time and the second countries.

A month was the unit to count numbers of ETDs in the collection, numbers of countries accessing the ETDs, average numbers of partitions and numbers of accesses (total & groups). Therefore, the organization of data had 71 sets, one for each month.

A group of countries was the unit to count accesses. The first group was all countries, the second was the "international group" and the third was the "pt+es group".

The variables computed for each month were:

- Number of ETDs in the collection.
- Average number of partitions.
- Number of countries that accessed ETDs.
- Number of pt-speaking countries that accessed ETDs.
- Percentage of pt-speaking countries $[(\text{number of pt-speaking countries} / 8) * 100]$ that accessed ETDs.
- Number of es-speaking countries that accessed ETDs.
- Percentage of es-speaking countries $[(\text{number of es-speaking countries} / 21) * 100]$ that accessed ETDs.
- Numbers of accesses by country.
- Total number of accesses.
- Number of accesses from countries in the "international group".
- Number of accesses from countries in the "pt+es group".
- Percentage of accesses from the "pt+es group" compared to the accesses from the "international group" $[(\text{number of accesses from the "pt+es group"} / \text{number of accesses from the "international group"}) * 100]$
- Number of accesses from the pt-speaking countries.
- Percentage of accesses from pt-speaking countries compared to the accesses from the "international group" $[(\text{number of accesses from pt-speaking countries} / \text{number of accesses from the "international group"}) * 100]$

- Number of accesses from the es-speaking countries.
- Percentage of accesses from es-speaking countries compared to the accesses from the "international group" $[(\text{number of accesses from es-speaking countries} / \text{number of accesses from the "international group"}) * 100]$

Since there were 71 data sets, the analysis was performed using simple concepts of descriptive statistics. A more accurate analysis is planned in the future.

The data sets have presented a time-varying behavior since, as time went by, the collection grew bigger and the number of national and international union catalogs also increased, making the collection more exposed to the world.

The Behavior of "pt+es group" as Compared to the "international group"

The behavior of the "pt+es group" in the "international group" was the first to be examined. To understand the results, it is important to take into consideration the following numbers:

- Total number of countries in the "international group" that have accessed ETDs over the 71 months under consideration – 198
- Maximum number of countries in the "international group" that accessed ETDs in a single month – 143
- Maximum number of countries in the "pt+es group" that accessed ETDs in a month – 28 (the number of countries in the group is 29)

Figure 1 shows the histogram of the numbers of countries in the "international group" per month. The highest frequencies are for the 101-125 and 126-150 intervals (countries per month). They total 38, meaning that more than 50% of the months had accesses from more than 100 countries.

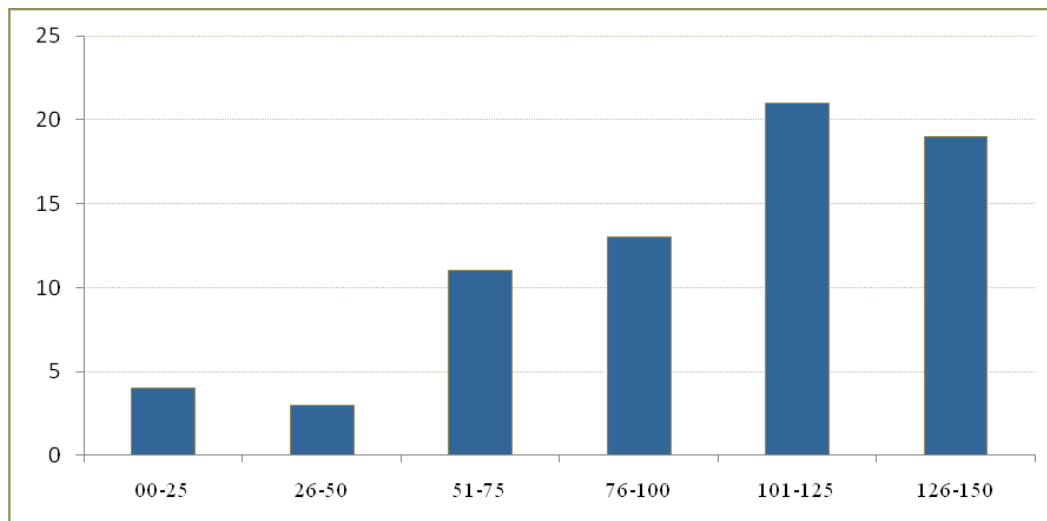


Figure 1 – Histogram of numbers of countries in the "international group" per month

For each of the 71 months, percentages were computed of the number of countries belonging to the "pt+es group" in the "international group" of the month. The histogram of the percentages is shown in figure 2. It is important to remember that the two sets have 28 countries. For this reason, the highest interval being that of 20.x to 40 is not surprising since most months had more than 100 countries accessing ETDs.

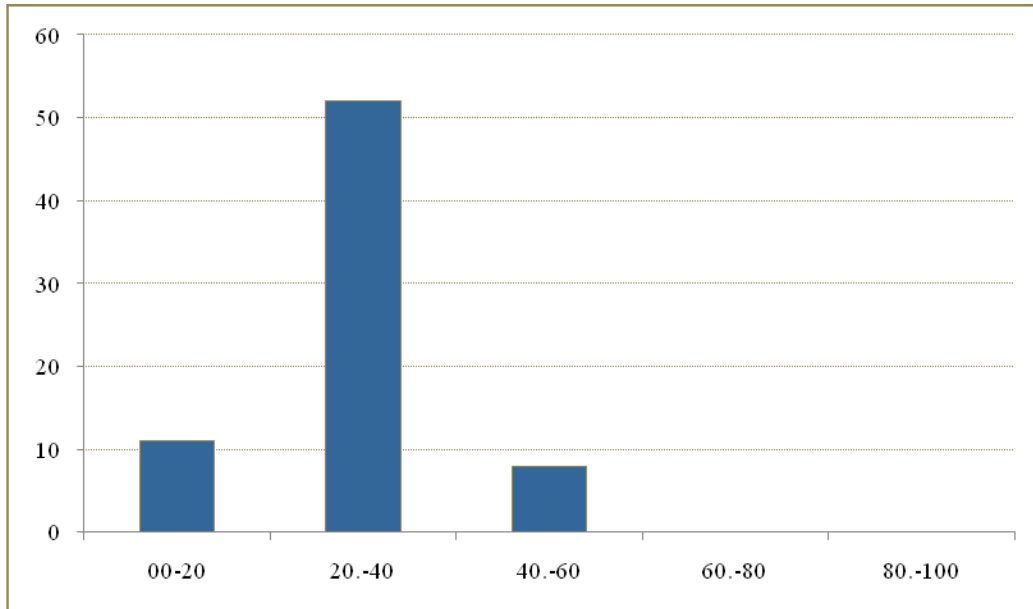


Figure 2 – Histogram of percentages of the "pt+es group" in the "international group"

After comparing numbers of countries, the next comparison focused on the numbers of accesses. For each month, the percentage of accesses from the counties in the "pt+es group" in the accesses of the countries in the "international group" were computed. Figure 3 shows the histogram of the percentages.

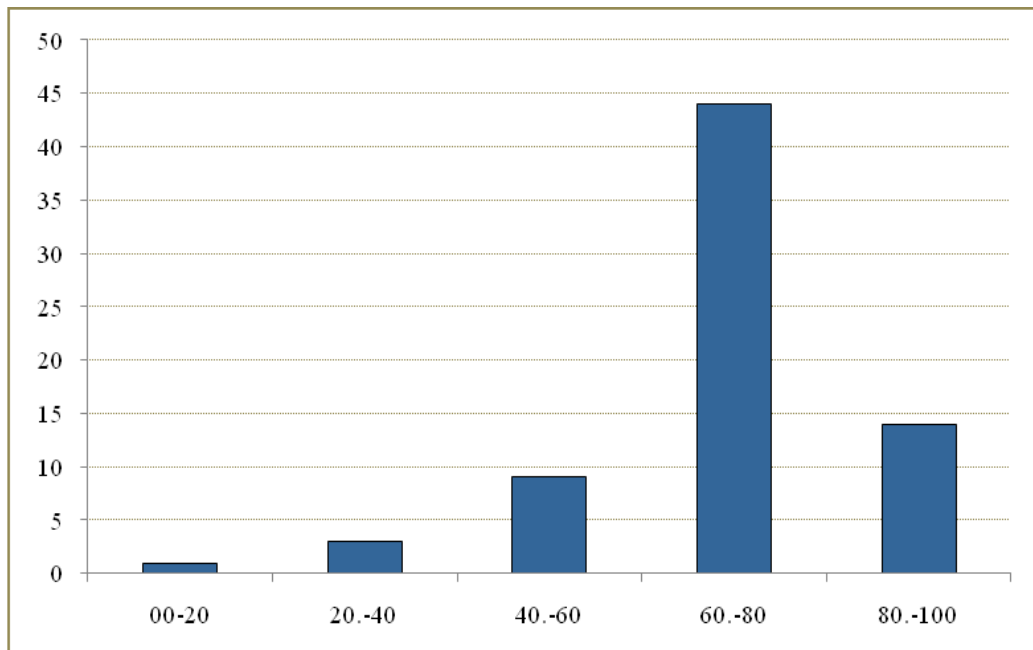


Figure 3 – Histogram of percentages of accesses from countries in the "pt+es group" in the accesses of countries in the "international group"

Examining figures 2 and 3 it is clear that the accesses from countries in the "pt+es group" represent higher percentages than the numbers of countries do. In the 71 months under consideration, the numbers were:

- Number of accesses from countries in the "international group" – 351,470
- Number of accesses from countries in the "pt+es group" – 269,641
- Percentage of accesses of the "pt+es group" – 76.80

It is important to remember that in the same time frame the average number of partitions is 6.9. Taking averages, the numbers approximately are 51K and 39K ETDs accessed, respectively, the two groups.

The percentage shows that the "pt+es group" significantly contributes to the numbers of accesses to PUC-Rio's ETDs.

The Behavior of Each Language in the "pt+es group"

Tables 1 and 2 showed that there are different expectations of accesses if numbers of countries, HDI and population are the only variables to influence the result. Index I is 11 times for Spanish speaking countries that of the Portuguese group.

This subsection presents the numbers computed for the each of the two subgroups. The first analysis was performed on the percentages of countries in each subgroup that accessed ETDs; they were computed for each month. Figure 3 and 4 show, respectively, the histograms for pt- and es-speaking countries.

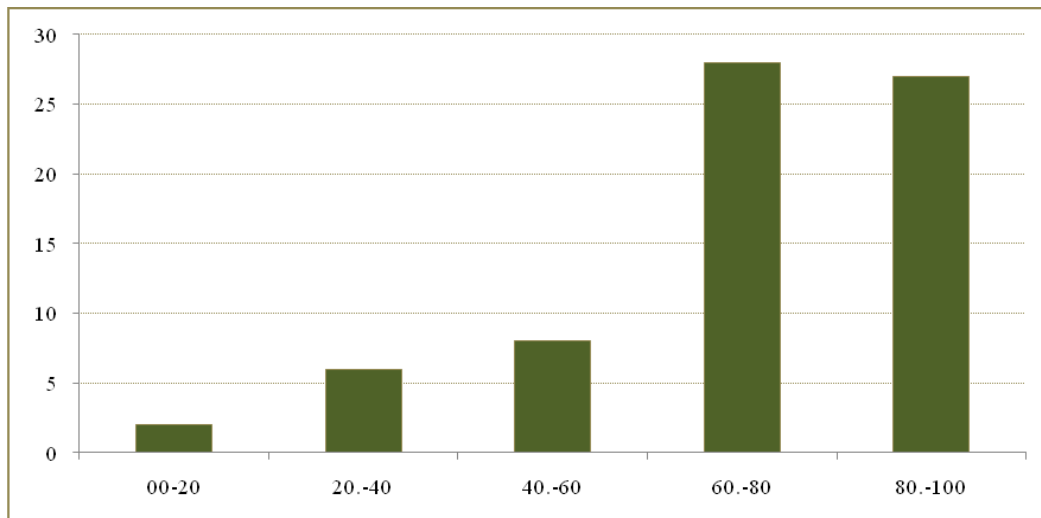


Figure 4 – Histogram of percentages of pt-speaking countries that accessed ETDs

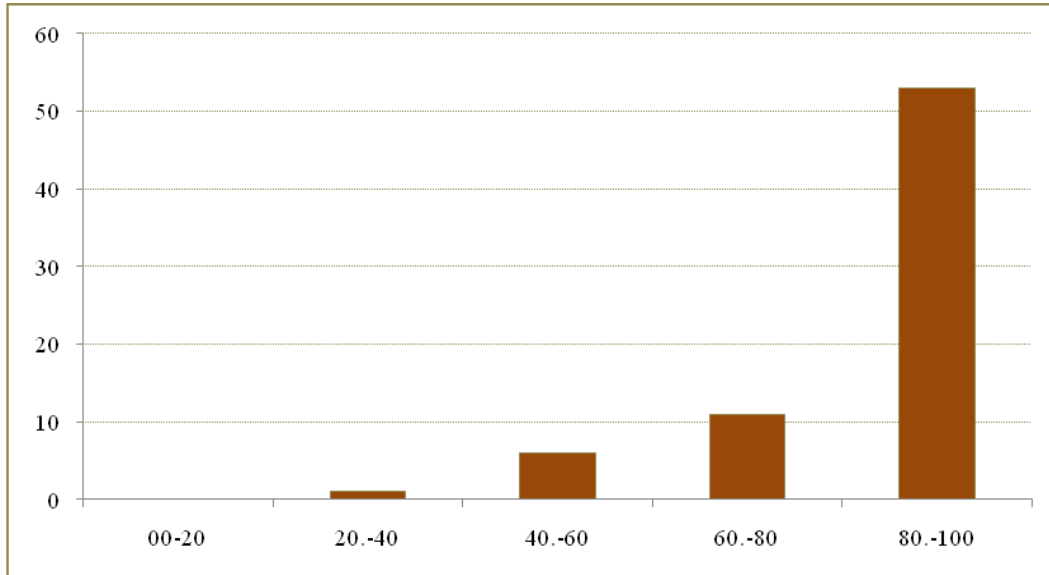


Figure 5 – Histogram of percentages of es-speaking countries that accessed ETDs

Examining figures 4 and 5 it is possible to conclude that about 60% of the countries in the two subgroups accessed ETDs most of the time (more than 50 months in 71).

The second analysis was on the numbers of accesses of each subgroup in terms of their percentages of the total of the "pt+es group". For each month the percentages were computed – one for the pt-speaking and the other for the es-speaking subgroups. Figures 6 and 7 show, respectively, the histograms of the percentages of pt-speaking and es-speaking countries.

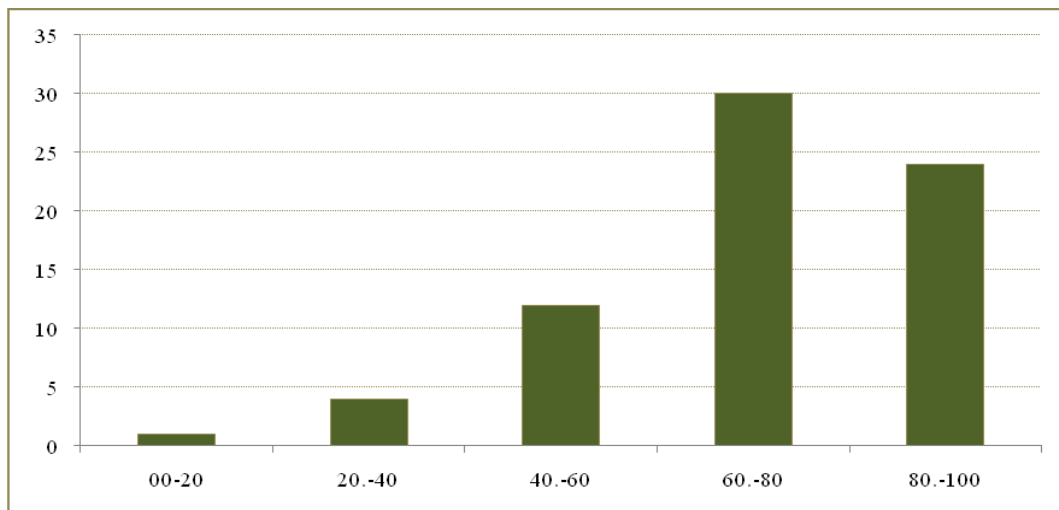


Figure 6 – Histogram of percentages of accesses from pt-speaking countries in the accesses of the "pt+es group"

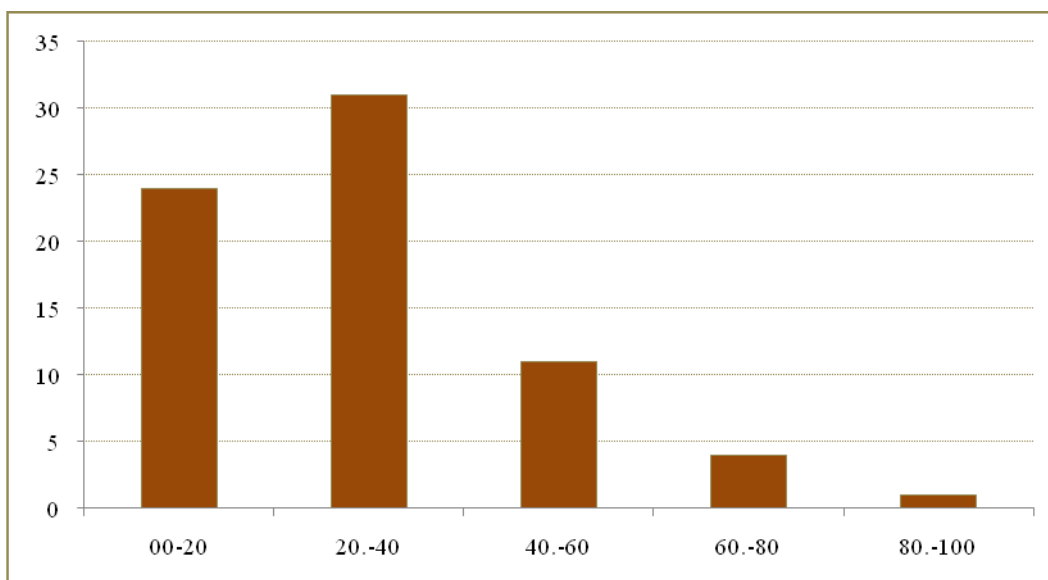


Figure 7 – Histogram of percentages of accesses from es-speaking countries in the accesses of the "pt+es group"

The histograms in figures 6 and 7 show that pt-speaking countries had the highest percentages of accesses in most of the months. Since the numbers of accesses vary from month to month, the total numbers were calculated. They are:

- Number of accesses from pt-speaking countries – 233,234
- Number of accesses from es-speaking countries – 36,707

This is a very interesting result that confirms data in the last two histograms and it shows that the number of accesses from pt-speaking countries is 6.35 the other number. It is opposed to the expectation based only in index I, that would suggest that the number of accesses from es-speaking countries should be 11 times that of the pt-speaking group. A possible interpretation is the language factor.

It is important that Portugal be examined separately. Data on accesses from Portugal in the "pt+es group" and in the "international group" were computed in terms of percentages. Considering averages of percentages, the results are:

- Average of percentages of accesses from Portugal in the "pt+es group" – 67.90
- Average of percentages of accesses from Portugal in the "international group" – 47.52

Figures 8 and 9 show the histograms of percentages of accesses from Portugal, respectively, in the "pt+es group" and the "international group".

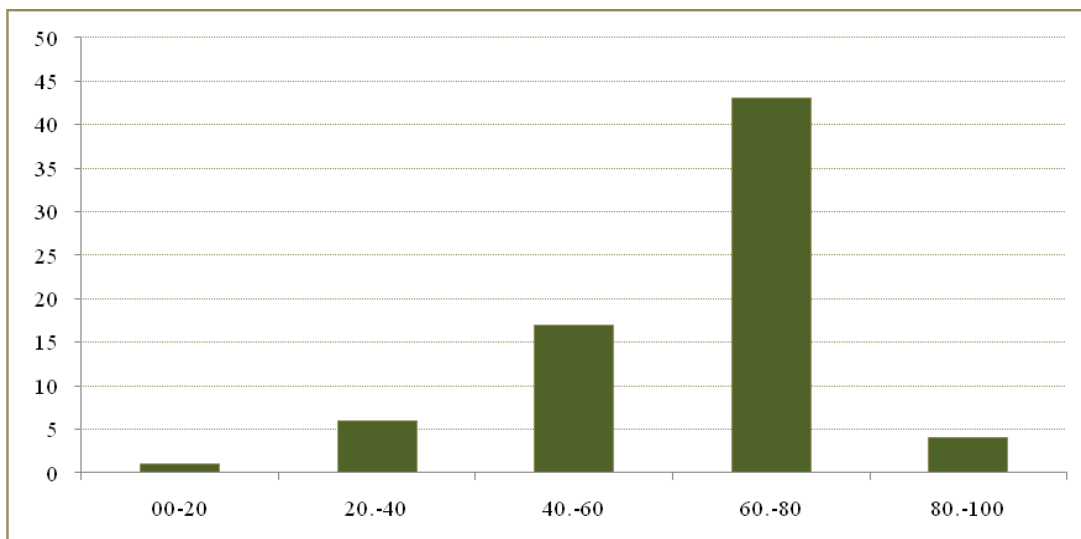


Figure 8 – Histogram of percentages of accesses from Portugal in the accesses of the "pt+es group"

Figure 8 shows that accesses from Portugal were between 60 and 80% of accesses from countries in the "pt+es group" in over 40 months in the 71 that were observed. This result confirms the average of 67.90%.

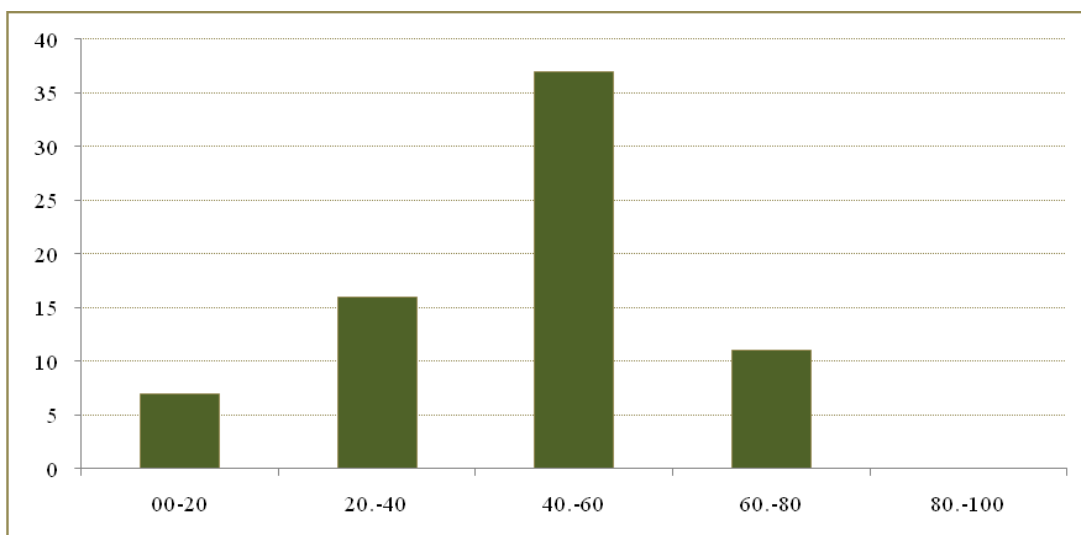


Figure 9 – Histogram of percentages of accesses from Portugal in the accesses of the "international group"

Both the averages and the histograms show that Portugal is very important in terms of accesses. This is probably due to the fact that Brazilian ETDs make up the largest collection in the Portuguese language available from international union catalogs.

Some Conclusions on Numbers of Accesses

The previous subsections of this section yielded results that lead to some conclusions on the accesses of PUC-Rio's ETDs. They can be stated:

- Countries in the "pt+es group" are responsible for 76.80% of the accesses of the "international group";
- About 60% of the countries in the "pt+es group" have accessed ETDs for more than 50 out of 71 months;
- Pt-speaking countries are responsible for 86.40% of the accesses of the "pt+es group" and this can only be explained by the common language of the countries and of the collection, since index I for the es-speaking subgroup is 11 times the other;
- Pt-speaking countries are responsible for 66.36% of all accesses from the "international group" though they number 8 in a set of 198;
- Portugal is the most important single country both in the "pt+es group" and in the "international group" in terms of contributions to the access numbers.

Both the averages and the histograms show that Portugal is very important in terms of accesses. This is probably due to the fact that Brazilian ETDs make up the largest collection in the Portuguese language available from international union catalogs.

The Brazilian ETD collection is OAI-PMH compliant and provides metadata records to international union catalogs; the number of records exceeds 132K. It is the largest in this language.

Though there is no formal proof, it seems reasonable to suppose that the language is a strong cause of this result. At the same time the authors are aware that even in the "pt+es group" there are many persons who can read other languages and so ETDs in Portuguese are not the only source of information graduate students, faculty and researchers.

6. Other Institutions

Brazil has 92 institutions with ETD programs that are members of BDTD and have their metadata harvested. The authors decided to seek information on accesses to ETDs of the largest BDTD contributors to try to check the results.

The number of institutions with over 2.5K ETDs in the collections is 10; PUC-Rio is the sixth in the rank. The digital libraries / institutional repositories of the other nine were visited in an attempt to find access data. The results of the visits were:

- UNICAMP – Universidade Estadual de Campinas (<http://www.unicamp.br/>)

UNICAMP has the largest ETD collection in Brazil – over 31.6K ETDs. The digital library system (Nou Rau) offers numbers of accesses by ETD, graduate program and year. There are no results related to the origins of the accesses, therefore it was not possible to compare with the numbers in this paper.

- USP – Universidade de São Paulo (<http://www.usp.br/>)

USP has the second largest ETD collection in Brazil – over 22.3K ETDs. The only information available concerns the set of the 10 most accessed theses and the set of the 10 most accessed dissertations. No numbers to compare.

- UFRGS – Universidade Federal do Rio Grande do Sul (<http://www.ufrgs.br/>)

UFRGS has over 12.1K ETDs. The digital library system (DSpace) offers access information by graduate program. There are no results concerning the origins of the accesses, therefore it was not possible to compare with the numbers in this paper.

- PUC-SP – Pontifícia Universidade Católica de São Paulo (<http://www.pucs.br/>)

PUC-SP hosts the ETD collection on TEDE – Teses e Dissertações Eletrônicas, an open and free software product that was developed and made available by IBICT. This product does not offer access information. PUC-SP has more than 7.5K ETDs in the collection.

- UFPE – Universidade Federal de Pernambuco (<http://www.ufpe.br/>)

UFPE also uses TEDE and for this reason has no info on accesses. The number of ETDs in the collection is over 6.6K.

- UnB – Universidade de Brasília (<http://www.unb.br/>)

UnB also uses TEDE and for this reason has no info on accesses. The number of ETDs in the collection is higher over 4.4K.

- UFMG – Universidade Federal de Minas Gerais (<http://www.ufmg.br/>)

UFMG hosts ETDs on OPUS that is based on DSpace (<http://www.dspace.org/>). The collection has over 3.2K ETDs. A visit to OPUS did not indicate the existence of any type of access information.

- UFSCar – Universidade Federal de São Carlos (<http://www.ufscar.br/>)

UFSCar also uses TEDE and for this reason has no info on accesses. The number of ETDs in the collection is over 2.8K.

- UFSM – Universidade Federal de Santa Maria (<http://www.ufsm.br/>)

UFSM also uses TEDE and for this reason has no info on accesses. The number of ETDs in the collection is over 2.5K.

The results do not allow any type of comparison between the accesses to PUC-Rio's ETDs and those in other Brazilian institutions. It is important to mention that these 10 universities are OAI-PMH data providers and their metadata records are available from BDTD and all other union catalogs that harvest from it; NDLTD union catalogs included.

An alternative for comparison would be other collections (with ETDs in Portuguese) not in Brazil. They should be OAI-PMH data providers and with their metadata available from NDLTD union catalogs, so that they could compare to the ETDs analyzed in this work.

A visit to OCLC XTCat – NDLTD Union Catalog (<http://alcme.oclc.org/ndltd/>) showed that the only other collection comes from Portugal. It is the DiTeD – Depósito de Teses e Dissertações (<http://dited.bn.pt/>), a repository of BN – Biblioteca Nacional de Portugal (<http://www.bn.pt/>). The number of metadata records available from the Union Catalog is 222, while the number of records of full-text works is 87, if the information on the DiTeD website is considered. There is no information on accesses, but if there were it would not be significant due to the small number of ETDs.

7. Comments

This is a preliminary work on accesses to the ETD collection available from the Maxwell System. There is abundant additional information that can be known through applications offered by the system. If new ways of examining data are found necessary, the corresponding applications to process them can be implemented.

Since there were no other collections with the same conditions, i.e., ETDs in Portuguese hosted by systems that are data providers and have metadata available from international union catalogs, comparisons were not made.

Assuming that these comparisons will not be possible in the near future, the next steps are to analyze accesses by:

- Areas of knowledge – this can easily be done by grouping graduate programs.
- Country and area of knowledge.
- Level (D/M).

A second type of comparison is also targeted by the authors – accesses from the “pt+es group” to ETDs that are in other languages. In order to keep the consistency of analysis, the conditions of being OAI-PMH data providers and having metadata records available from the NDLTD union catalogs will be maintained. For this reason, digital libraries of institutions that contribute to the union catalogs will be searched for access data.

The final remark is to emphasize the importance of the “pt+es group”, specially the Portuguese speaking countries as “readers” of PUC-Rio’s ETDs.

8. References

- [01] Infoplease
<http://www.infoplease.com/ipa/A0775272.html>
Captured on April 1, 2010
- [02] Internet World Stats
<http://www.internetworldstats.com/stats7.htm>
Captured on April 1, 2010
- [03] VicNet – The State Library of Victoria
<http://www.librarybooksforeasttimor.org.au/pop.htm>
Captured on April 03, 2010
- [04] The Human Development Index (HDI)
<http://hdr.undp.org/en/statistics/indices/hdi/>
Captured on April 06, 2010
- [05] IBGE – Instituto Brasileiro de Geografia e Estatística
<http://www.ibge.gov.br/>
Captured on April 06, 2010