DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Internal Research Reports

ISSN

Number 21 | November 2011

# Examining Accesses by Country, Language and Area of Knowledge

Ana Maria Beltran Pavani

PUC RIO

**Internal Research Reports**

Number 21 | November 2011

# Examining Accesses by Country, Language and Area of Knowledge

Ana Maria Beltran Pavani

**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

This article was originally published in the Proceedings of the 14th International Symposium on Electronic Theses and Dissertations, National Research Foundation, South Africa, a3 September 2011, ISBN 978-0-51049-3. Reference: file:///E:/papers/pavani.pdf

It is also available from http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011_pavani.pdf .

# Examining Accesses by Country, Language and Area of Knowledge

**Ana M B Pavani**
Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil
apavani@lambda.ele.puc-rio.br

**ABSTRACT**

This paper addresses the analysis of accesses to an ETD collection whose items are mostly written in Portuguese. It is concerned about examing if Spanish and, specially, Portuguese speaking countries are important readers of the works. This analysis for the whole collection is an extension of a previous work that innovates by breaking the problem in 3 subsets of data – ETDs in Humanities & Theology, ETDs in Social Sciences and ETDs in Science & Technology. The differences among the subsets are quantized and considered, so that accesses can be viewed in this context. This article also updates the index created in the previous work to identify potential readers in countries where either or both languages are official languages. An additional result is the identification of the differences among areas in terms of numbers of ETDs, of partitions of the works and of profiles of accesses.

**Keywords**

Accesses to ETDs, accesses by countries, accesses by language groups, areas of knowledge.

**INTRODUCTION**

Pontifícia Universidade Católica do Rio de Janeiro has had an ETD program since 2000; the first ETD was published in May. In August 2002, ETDs became mandatory to all graduate programs of the university. Over 99% of the ETDs are in Potuguese because theses and dissertations written in other languages have only been allowed in the last few years.

The Maxwell System (http://www.maxwell.lambda.ele.puc-rio.br/) that hosts and makes ETDs available has been an OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting data provider since December 2002. The university is a member of both BDTD – Biblioteca Digital de Teses e Dissertações (http://bdtd.ibict.br/), the Brazilian national ETD consortium, and NDLTD – Networked Digital Library of Theses and Dissertations (http://www.ndltd.org/).

A previous study of accesses to ETDs with a special focus on Portuguese and Spanish speaking countries (Pavani and Mazzeto, 2010) viewed the collection as a single unit with no separation in areas of knowledge. Available data were from June 2004 to April 2010. June 2004 was the first month when access logs were collected and processed to yield information on accesses to the collection. Access information is available to other digital contents that are available from the Maxwell System as well. This system is an institutional repository of PUC-Rio. The analysis used 71 data sets - one for each month with information on countries, numbers of accesses by country, numbers of ETDs and average numbers of partitions. For each month, raw data were combined to yield other indicators used in the work.

This paper is an extention of results of a subset of the analysis performed in 2010. Not all indicators are calcultated, but for the selected ones calculations address 3 subsets of the ETD collection. These sets are the areas of knowledge of each center of the university. The centers are:

- CTCH – Centro de Teologia e Ciências Humanas (Center for Humanities & Theology).
- CCS – Centro de Ciências Sociais (Center for Social Sciences).
- CTC – Centro Técnico-Científico (Center for Science & Technology).

All centers have graduate programs and the oldest ones are in Science &Technology and started in the 1960s.

In this work, once again, the focus is on accesses and languages but going beyond the collection as a whole. ETDs of each center are grouped in a subset named after the corresponding center.

At the same time, 14 data sets – from May 2010 to June 2011 – are added. The reason for this is that these 14 months represent almost 20% additional data sets. These are the newest and bring more up-to-date characteristics of accesses; it is expected that they are more significant than the oldest ones. A second reason to include new data was the increase in the

number of ETDs available from the BDTD – Biblioteca Digital de Teses e Dissertações union catalog. In December 2004, there were almost 5.4K metadata records; in December 2010, the number was a little over 150K; and, in Jun 2011, it was over 162K. A visit to the VTLS Visualizer (http://thumper.vtls.com:6090/) shows that there are over 177K records of ETDs in Portuguese while the number of the BDTD colletion is approximately 137K. This means that other countries publish ETDs in Portuguese. As the offer of ETDs in Portuguese increases, it is expected that numbers of accesses to PUC-Rio's collection vary since there is a bigger offer of ETDs in this language; so it is important to consider the newest data.

Another change from the previous work is that the way the HDI – Human Development Index is computed has been modified by the UNDP – United Nations Development Program; this change is important in the expectations of accesses.

Both topics will be addressed in later sections. An overview of the main characteristics of the collection is presented in the next section. It is important for the understanding of the treatment applied to access data..

## UNDERSTANDING THE ETD COLLECTION

The ETD collection of PUC-Rio has some characteristics that must be addressed before accesses are considered. At the end of June 2011, it held 5,694 ETDs of "different generations".

The first generation is of ETDs published between April 2000 and August 2002 – each one was published in one single file except for being too big (at the time bigger than 1.2 MB) or for having some restricted content. In these cases, they were partioned to allow saving the file on a diskette or temporarily protecting part of the work while the remaining part would immediately be made public. Except for eventual partial temporary restrictions, they were all of public access since ETDs were voluntary and only authors who wanted to make their works public would participate in the ETD program..The graduate programs that had joined the program were in different areas of Engineering and in Business Administration.

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

The third generation is of ETDs that came from retrospective digitization of theses and dissertations in Electrical Engineering. They were made available with the same policy of the first generation, but they differ because some 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 be contacted to seek authorization for public access.

As stated before, PUC-Rio's ETD collection holds works from graduate programs in 3 areas. The 3 subsets are quite different in size and in the numbers of digital objects. In the last 8 years, students have been required to submit their ETDs in one separate file for each chapter. The separation by area of knowledge has shown that they differ in the average number of files per work. Table 1 shows some characteristic numbers of the collection when considered as a whole and for each subset. Data up to June 2011 were included because of the extended time frame. Numbers related as of April 2010 are mentioned because they were the end point of the previous work.

| | Collection | CTCH | CCS | CTC |
|---|---|---|---|---|
| Number of ETDs – June 2004 | 1,171 | 233 | 215 | 723 |
| Number of ETDs – April 2010 | 5,140 | 1,297 | 1,090 | 2,753 |
| Number of ETDs – June 2011 | 5,694 | 1,442 | 1,291 | 2,961 |
| Average number of ETDs – June 2004 to April 2010 | 3,181.5 | 791.7 | 633.0 | 1,576.6 |
| Average number of ETDs – June 2004 to June 2011 | 3,553.1 | 888.9 | 726.1 | 1,938.3 |
| Average number of partitions – June 2004 | 5.9 | 7.9 | 5.5 | 5.4 |
| Average number of partitions – April 2010 | 7.2 | 7.9 | 7.2 | 6.9 |
| Average number of partitions – June 2011 | 7.3 | 7.9 | 7.3 | 7.0 |
| Average of averages number of partitions – June 2004 to April 2010 | 6.90 | 7.81 | 6.75 | 6.49 |
| Average of averages number of partitions – June 2004 to June 2011 | 6.93 | 7.82 | 6.83 | 6.57 |

**Table 1. PUC-Rio's ETD Collection Profile**

Numbers of ETDs in the last column of table 2 show that the Engineering graduate programs were the first to join the ETD

program and also indicate the Electrical Engineering did retrospective digitization of all its theses and dissertations – the numbers for the CTC subset are higher than the sum of the corresponding numbers in the other 2 subsets.

The average numbers of partitions are important because, since numbers of accesses are counted in tens of thousands, it is possible to associate 6,930 accesses (to the collection) to 1,000 ETDs being completely accessed; this is an average number and does not mean that all accessed ETDs were completely accessed.

The differences among the numbers in different areas of knowledge indicate that, if comparisons are to be made among centers, a normalization of data is necessary. Computations and/or comparisons of results in a subset do not require such normalization. The identification of different profiles of ETDs in terms of numbers of partitions is an interesting by product of this work. It is important to remark that:

- Most partitioned ETDs belong to the second generation.
- First generation ETDs are in Engineering (CTC) and Business Administration (CCS).
- Third generation ETDs are in Electrical Engineering (CTC).

For the above reasons, it is expected that the average number of partitions in the CTC subset be lower. The averageof averages number of partitions for the CTC set is 84% that of CTCH. The current relation (June 2011) is 88.61%; no more first and third generation ETDs are being added to the collection. This may change in the near future because the Graduate Program in Mechanical Engineering has been digitizing theses and dissertations that will be made available in the next few months.

## PORTUGUESE AND SPANISH SPEAKING COUNTRIES

Information available under the title Most Widely Spoken Languages in the World (infoplease , 2011) indicates that Spanish and Portuguese are, respectively, the 1st and the 3rd most widely spoken Western languages; English is the 2nd.

When Internet users are considered, as available on Internet World Users by Language (Internet World Stats , 2011a), the numbers of Spanish and Portuguese speakers rank, respectively, 3rd and 5th.

The growth of the numbers of Internet users, between December 31, 2000 and March 31, 2011, is published under the title Internet World Usage and Population (Internet World Stats , 2011b). The numbers show that Africa is the region with the highest growth rate – 2,527.4%, and Latin America and the Caribbean has the 3rd highest – 1,037.4%. This is important because most es- and pt-speaking countries are in these regions. If this trend is maintained, it means that more people in es- and pt-speaking countries will have access to the Internet meaning that potential of PUC-Rio's ETDs readers may increase.

An interesting aspect of Portuguese and Spanish is that they are quite similar languages; educated speakers of either one 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 both groups; they are considered potential readers.

### Countries and Groups

Countries that have at least one of the 2 languages as an official language are classified in 2 groups. The groups are:

- Es-speaking countries – 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.
- Pt-speaking countries – Angola, Cape Verde, East Timor, Equatorial Guinea, Guinea-Bissau, Macau, Mozambique, Portugal and Sao Tome and Principe.

The countries in the 2 groups are quite diverse in size and in geographic situation. Portugal and Spain are in Europe; most es-speaking countries are in Latin America; most pt-speaking countries are in Africa; some have large populations (Mexico, for example, has over 109M inhabitants) and others have small populations (Sao Tome and Principe, for example, has only a little over 162K inhabitants).

East Timor was not included in the previous work due to reasons that are clearly stated in it. In the last 14 months, conditions seem to have changed and East Timor was included because accesses from this country have been observed not only to ETDs but to other contents too.

### Human Development Index

Another important difference among the countries is in human development levels.

The United Nations Development Program created the HDI – Human Development Index (UNDP, 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)".

This index is available for most countries of the world. Data used by Pavani and Mazzeto were the latest available at the time; it had been published in October 2009 and covered up to 2007. At the same time, it is important to remark that the way HDI is computed has been changed by the UNDP; the latest published indices (UNDP, 2010) were computed with the new definition. Up to 2010, education contributed to HDI with adult literacy rate (2/3) and the combined primary, secondary, and tertiary gross enrollment ratio (1/3). The 2010 HDI definition considers as its education component a set of 2 indicators – mean years of schooling and expected years of schooling, as presented in pages 13-15 and Technical Note 1 in pages 216-217 (UNDP, 2010). This new definition lowered all HDIs when compared with the previous report. It is used in this work because its education component is more precise and this is important when ETDs are concerned.

Once again, countries in the 2 groups are quite diverse in terms of HDIs. In the 2010 report, the range of indices is from 0.289 (Guinea-Bissau) to 0.863 (Spain).

## ETDS AND POTENTIAL ACCESSES

Pavani and Mazzeto were concerned about examining accesses to the ETD collection from countries that are pt- or es-speaking. ETDs are a very special type of literature and some factors influence the potential accesses to them. The authors considered 3 factors. The first factor was the size of the country in terms of population. The level of education of the population was a second important factor since ETDs are highly specialized resources; the new definition of the HDI will be more precise to represent this factor. A third factor also considered was how easy/difficult it is for the population to access the Internet; ETDs are made available from digital libraries / institutional repositories connected to the Internet.

In order to assign numbers to expectations of accesses it was necessary to find an index taking into consideration the 3 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 2 factors had to be measured at the same time for all countries under consideration so that comparisons were valid. Due to these reasons, 3 decisions were made by the authors:

- To use the population of the country as one of the numbers.
- To use the HDI 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). Access to the Internet is dependent on the living standard too. This decision was maintained but updated – the new numbers for the HDIs and populations are used in this work; they are from the last report (UNDP, 2010).
- To create an index that combined both the population and the HDI. The index was:

$$I = (HDI) \times (Population) \quad (1).$$

Table 2 shows the numbers for both groups. The numbers in the last row of each column are not the product of the ones in the first and the second rows – countries differ in both population and HDI.

| | Es-speaking group | Pt-speaking group |
|---|---|---|
| Total population | 420,281,000 | 57,858,800 |
| Average HDI | 0.707 | 0.527 |
| **Index I** | **309,420,871** | **25,114,111** |

**Table 2. Index I for the Es- and Pt-speaking Groups**

Comparing index I for both groups it is easy to see that es-speaking countries would yield 12 times the expected accesses when compared to the pt-speaking group. The total es-speaking population is almost 8 times that of pt-speaking group and its average HDI is 34% higher. This was not what the results computed for the accesses to the collection as whole showed. This happened both in the previous work and in this one. Those from the pt-speaking group were much higher. The next sections present the highlights of the results.

## ACCESSES TO THE COLLECTION

In order to understand the results of accesses to the collection and to compare them with the ones to ETDs in each area of knowledge, the next subsection addresses the way data were combined and analyzed.

### Grouping Countries of the World

From June 2004 to June 2011, users from 204 countries accessed ETDs. As expected, most of them were from Brazil. The second largest group was from the United States. Brazil was not considered in the analysis because it is the "home country" of the collection. The United States were not considered due to its large numbers of both Spanish and Portuguese speakers, including a large number of graduate students from all over the world. There were accesses from all the countries belonging to the es- and pt-speaking groups.

The decision was to divide the countries in the following groups:

- Brazil and United States – accesses from the 2 countries were not considered.
- International group – all countries that accessed ETDs except Brazil and the United States.
- Es+pt-speaking group – all countries that have one of the languages as an official language, as listed in a previous section. This group is a subset of the international group.

The same grouping of countries is used to analyze data for each area of knowledge so that comparisons can be made.

### Accesses to the Collection

Some of the numbers presented in the previous work were computed for the new set of 85 months. The most important numbers related to the accesses to the collection as a whole are :

- Total number of countries of the international group that accessed ETDs – 202.
- Maximum number of countries of the international group that accessed ETDs in a single month – 143.
- Maximum number of countries in the es+pt-speaking group that accessed ETDs in a single month – 28 (the total is 30 because Equatorial Guinea has both Portuguese and Spanish as official languages).
- Number of months with 100 or more countries accessing ETDs – 42.
- Percentage of accesses that came from the international group – 8.48%
- Percentage of accesses of the international group that came from the es+pt-sepaking group – 69.03%.
- Percentage of accesses of the es+pt-speaking group that came from pt-speaking countries – 82.07%.
- Percentage of accesses of the international group that came from pt-speaking countries – 56.65%
- Percentage of accesses in the int group from Portugal – 49.74%
- Percentage of accesses of the es+pt-group that came from Portugal – 72.05%
- Percentage of accesses of the pt-speaking group that came from Portugal – 87.89%

The results showed that language seems to be an important factor in the accesses; Pt-speaking countries accounted for a little over 82% of the accesses from the es+pt-speaking group though its Index I was (1/12) of that of the es-speaking group. Accesses from pt-speaking countries accounted for more than 56% of the international accesses. When the percentages of accesses from Portugal are considered, it is clear that they are dominant – almost 50% of all international accesses and a little over 70% of the pt-speaking group. The population of Portugal is almost 11M in a pt-speaking population of almost 58M, accounting for a little less tha 20%. Portugal had almost 88% of the accesses from this group. It seems reasanoble to conclude that within the language group, HDI is a key factor.

## ACCESSES BY AREA OF KNOWLEDGE

This section is divided in 2 subsections. The first addresses accesses in each area of knowledge separately; they are analyzed in the same way the collection was. The second compares the way international accesses split among the areas.

### Accesses to ETDs in Each the Area of Knowledge

The accesses by area of knowledge are analyzed in 2 different manners The first repeats the same analysis performed for the whole collection but for each area of knowledge and the second compares results among the areas.

| | Collection | CTCH | CCS | CTC |
|---|---|---|---|---|
| Total # of countries in the international group | 202 | 181 | 181 | 187 |
| Max # of countries of the int group in a single month | 143 | 112 | 108 | 132 |
| Max # of countries in the es+pt-speaking group in a single month | 28 | 27 | 27 | 27 |
| # of months with 100 or more countries | 42 | 18 | 15 | 32 |
| % of accesses from the int group | 8.48 | 7.99 | 7.89 | 9.12 |
| % of accesses in the int group from the es+pt-sepaking group | 69.03 | 73.27 | 68.56 | 66.32 |
| % of accesses in the es+pt-speaking group from pt-speaking countries | 82.07 | 87.11 | 84.44 | 77.35 |
| % of accesses in the int group from pt-speaking countries | 56.65 | 63.83 | 57.89 | 51.30 |
| % of accesses in the int group from Portugal | 49.74 | 57.39 | 49.54 | 44.69 |
| % of accesses in the es+pt-speaking group from Portugal | 72.05 | 78.27 | 72.26 | 67.39 |
| % of accesses in the pt-speaking group from Portugal | 87.89 | 88.92 | 85.57 | 87.12 |

**Table 3. Profile of the Accesses to the Collection as a Whole and to the Subsets of Each Center**

Some remarks are interesting concerning data in table 3..

- Rows 1-4 contain absolute numbers and CTC shows the highest values among centers – this can be explained by the fact that the average number of ETDs of this center is higher than the sum of the other 2 averages (table 1). It is possible to suppose that if the other 2 centers had similar numbers of ETDs the first 4 rows could show similar results. The average number of partitions is the lowest for this center, but the rows are concerned with numbers of countries and are not related to the numbers of accesses, so it is expected that this fact does not impact the results.
- Rows 5-11 show percentages of accesses in the subsets; there are no computations relating data for different areas. CTC seems to be more international – the percentage of accesses from the international group is the highest, the percentage of accesses from the es+pt-group in the international group is the lowest and the same happens with other percentages related to accesses from pt-speakers. This is interesting because there is a higher percentage of accesses from the international group and in this group es and pt are not as significant as in the other areas.

The overall impression is that ETDs in Science & Technology behave differently from ones in Humanities and Social Sciences. In the case they are "more international".

### Accesses and How They Split Among Areas

This is a first attempt to examine data from this point of view – how accesses to ETDs are split among the 3 areas of knowledge. Most probably this examination is incomplete and more work should be devoted to it in the future.

Table 1 shows that the number of ETDs and corresponding partitions are quite different in the 3 areas. In order to try to compare how each area "attracts" readers from all over the world, a normalization is used. An index of equivalence is proposed. It is:

$$EI = \frac{1}{(\text{Average Number of ETDs}) \times (\text{Average of Averages Number of Partitions})} \quad (2)$$

There will be four different values for this index – one for the whole collection and one for each area of knowledge. They are computed for the 85 month time frame and the values are shown in table 4. If total numbers of accesses are multiplied by the corresponding index, the average number of accesses per ETD will be computed. This result is an average since there is no information concerning the individual files that were accessed. Another imperfection comes from restricted ETDs that have not be identified to be taken away from the computations; there is even a more difficult problem – ETDs that have been under restriction and afterwards became public. This last type of restriction is an option to students who want to protect their works while a patent is application is under exaimnation or an aritcle / book is waiting for publication for example.

| | Collection | CTCH | CCS | CTC |
|---|---|---|---|---|
| **Index EI** | **0.000041** | **0.000144** | **0.000202** | **0.000079** |

**Table 4. Values of Index EI**

Table 5 shows the average number of accesses per ETD. The numbers were computed using accumulated accesses in the 85 months. Since there are over 5.5K ETDs and data collected for 85 months, the results are to be viewed as a general behavior of the collection and of the subsets.

| | Collection | CTCH | CCS | CTC |
|---|---|---|---|---|
| Average number of accesses | 740.40 | 901.73 | 755.93 | 644.32 |
| Average number of accesses from the int group | 62.77 | 72.01 | 59.63 | 58.79 |
| Average number of accesses from the es+pt-speaking group | 43.33 | 52.76 | 40.88 | 38.99 |
| Average number of accesses from the pt-speaking countries | 35.56 | 45.96 | 34.52 | 30.16 |
| Average number of accesses from Portugal | 31.22 | 41.32 | 29.54 | 26.27 |

**Table 5. Average Accesses per ETD from June 2004 to June 2011**

If numbers in table 5 are divided by 85, the average monthly behaviors can be determined.

The percentage variations among rows in the same column of table 5 are the same as corresponding percentages shown in table 3. This is an exepcted result due to the definiton of EI given by (2).

A comparison of the results of tables 5 and 3 yields a curious comment. In table 3, the analysis of percentages within an area of knowledge, Science & Technology is more international because its percentage of accesses from the international group was the highest. In table 5, where normalized data is shown, the average number of accesses (per ETD in Science & Technology) from the international group is the lowest among all; the same happens with accesses from the es+pt and pt-speaking groups, and Portugal as well. The reason is that ETDs in this group have the lowest average of accesses per ETD among the 3 subsets.

## CONCLUSIONS

This is an initial analysis of accesses to ETDs in the 3 different areas of knowledge that PUC-Rio offers graduate courses. No specific graduate programs were examined, each area is dealt with as a single unit to be compared to the whole collection. A quick overview of raw data indicated that in each are graduate programs behave differently in terms of accesses and this may be interesting to examine.

Data generated during this work yielded a by product – the computation of the average numbers of partitions in each area. CTCH has the highest number and CTC the lowest.

Although there are differences among areas, the general behavior of the collection and the behaviors of each area are similar concerning the importance of language in the accesses. Results indicate that language is an importante factor because:

- Percentages of accesses of the international group that came from the es+pt-group (table 3) are over 66% in all cases.
- Percentages of accesses of the international group that came from pt-speaking countries (table 3) are over 51% in all cases.
- Percentages of accesses of the es+pt-speaking group that came from pt-speaking countries (table 3) are over 77% in all cases.

Within the language group, HDI is important because:

- Percentages of accesses of the international group that came from Portugal (table 3) are over 44% in all cases.
- Percentages of accesses of the es+pt-speaking group that came from Portugal (table 3) are over 67% in all cases.
- Percentages of accesses of the pt-speaking group that came from Portugal (table 3) are over 85% in all cases.

In the previous work, the authors tried to find collections with the same characteristics of PUC-Rio's in order to compare accesses, but there were none that offered public access statistics with analogous functionality.

If systems hosting ETDs offered a similar set of access statistics, comparisons would be possible. This could be a topic to be discussed by NDLTD – a suggestion of a set of statistics (production and accesses) for ETD systems. This would allow to compare behaviors by language, country, area of knowledge, etc.

## REFERENCES

1. infoplease – The Most Widely Spoken Languages in the World 2011. Available http://www.infoplease.com/ipa/A0775272.html.
2. Internet World Stats – Internet World Users by Language 2011a. Available http://www.interne2rldstats.com/stats7.htm.
3. Internet World Stats – Internet World Usage and Population 2011b. Available http://www.internetworldstats.com/stats.htm. .
4. Pavani, A. M. B. and Mazzeto, A. C. E. 2010. Examining Accesses by Country and Language, presented at ETD 2010 – International Symposium on Electronic Theses and Dissertations, 16-18 June, Austin, TX, USA. Available https://conferences.tdl.org/index.php/utlibraries/etd2010/paper/viewFile/34/53.
5. UNDP – United Nations Development Program 2010. 2010 Human Development Report. The Real Wealth of Nations: Pathways to Human Development, Available http://hdr.undp.org/en/reports/global/hdr2010/.
6. UNDP – United Nations Development Program, HDI – Human Development Index, 1990. Available http://hdr.undp.org/en/statistics/indices/hdi/.