EXAMINING ACCESSES TO OPEN ACCESS SIMULATOR OBJECTS IN ENGLISH AND IN PORTUGUESE

Ana M B Pavani, *Member IEEE*, and William S Barbosa
Pontifícia Universidade Católica do Rio de Janeiro
Rio de Janeiro, RJ, Brazil

ABSTRACT

This work focuses on analyzing access patterns to a collection of simulators within an Open Educational Resources project. The simulators are originally developed in Portuguese and then translated into English. The investigation compares the numbers of accesses and the countries accessing each pair of simulators over a six-year period (2018-2023). The data utilized for this analysis are sourced from an openly available dataset in the references. Two distinct analytical approaches are employed. The first relies on calculating averages, while the second involves analyzing access patterns under specific conditions using yearly data. The findings reveal significant differences in access patterns, attributed not only to language variations but also to the thematic contents addressed by the simulators.

KEYWORDS

Open Educational resources, Simulator Objects, accesses, Simulators in Portuguese, Simulators in English

1. INTRODUCTION

Online materials play a pivotal role in e-learning and digital education, particularly within the framework of Open Educational Resources (OER). Various designations are associated with such materials, including Learning Object (LO) as defined by IEEE (2002), Shareable Content Object (SCO) according to the ADL Guide (2004), Reusable Chunks of Instructional Media proposed by Wiley (2000), and Reusable Learning Objects as discussed by Alsubaie and Alshawi (2009).

This work focuses on examining accesses to a select set of OERs that are available at Open Educational Resources @PUC-Rio (https://www.maxwell.vrac.puc-rio.br/projetosEspeciais/OER/Home.php). The OERs under consideration are simulators, though the collection hosts others of different natures. They have special characteristics that led to their choice for observation. The simulators and their special characteristics are presented in section 2. The objective of this analysis is to discern access patterns, including the numbers of countries accessing each OER and the total numbers of accesses, for both the Portuguese (pt) and English (en) versions. The primary research question guiding this study is:

Research Question (RQ): Are the patterns of accesses (numbers of countries and of accesses) different for simulators in Portuguese (pt) and their English (en) versions?

This work is an assessment of this project that started in May 2015 with simulators in Portuguese and the first English version made available in April 2016. The objective of the English versions is to support the sharing of online materials whose implementations are resource and time consuming, Since February 2023, Spanish (es) versions of the simulators have been developed and published with the same objective. The authors of the OERs all agree with a Creative Commons License (https://creativecommons.org/) in order to contribute to the worldwide effort based on Create-Share-Reuse.

It is important to remark that The Most Spoken Languages Worldwide in 2023 (Statista,2023), lists English (1st), Spanish (3rd), French (4th) and Portuguese (8th) among the 10 most spoken languages in the world; the other 6 are Russian and Eastern languages. Wang and Towey (2017) list Language and Localization as one of the challenges for the use of OERs in higher education. They also cite the existence of the Universia Digital Library that offers Learning Materials in Portuguese and Spanish. So, language is a wider concern.

This work is divided in 4 sections besides this Introduction. Section 2 describes the context at the university and the available data. Section 3 presents the data to be analyzed, the analysis and the results. Section 4 comments the results.

2. CONTEXT AT THE UNIVERSITY, AVAILABLE DATA AND THE ANALYSIS PERFORMED

2.1 The OER collection

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) operates an IR – Institutional Repository (Lynch, 2003) that hosts thousands of digital contents created at the university. It is the Maxwell System (https://www.maxwell.vrac.puc-rio.br) that publishes Electronic Theses and Dissertations (ETDs), Senior Projects (TFCs), Monographs, Research Data (RD), journals, research & technical reports, restricted courseware and OERs. In order to make OERs easier to find, an aggregator, Open Educational Resources @PUC-Rio (https://www.maxwell.vrac.puc-rio.br/projetosEspeciais/OER/Home.php), with all interfaces in English was deployed. There are aggregators for ETDs, TFCS and RD too.

Since the focus of this work is OERs, some additional information about them is necessary. The characteristics of the OER collection are:

01. The collection has 492 titles with 1,566 objects – many titles have multiple objects;
02. 452 titles are in Electrical Engineering (EE) and Science & Technology (S&T) topics while 40 are in Portuguese as a Second Language (PSL);
03. OERs in EE and S&T are of various natures – hypermedia objects, videos, podcasts and simulators;
04. The topics addressed by the EE and S&T OERs are part of the courses in the curricula of EE was well as Control Engineering and Computer Engineering in institutions all over the world;
05. Simulators are the focus of this work because they are the most interactive in the collection and stimulate students being proactive;
06. Simulators have been developed by faculty, undergraduate students, graduate students and technical staff; the participation of students was twofold – they came with the suggestions of topics they wanted to learn better and then became the authors of the simulators;
07. Simulators are first implemented in Portuguese and, afterwards, translated into English and Spanish;
08. Currently, there are 43 simulators in Portuguese, 22 in English and 22 in Spanish. Others are under development;
09. Simulators are powered by Scilab (https://www.scilab.org/) which is installed on a server and accessed by the IR;
10. Each simulator has a theoretical introduction with help and instructions of use – the interfaces are in HTML and they communicate with Scilab via PHP programs;
11. Users do not see the Scilab interface and do not have to learn how to use it, they just have to fill the boxes with functions and/or parameters.

In the previous section, it was stated that the usage of the simulators in English when compared to the originals in Portuguese could serve as an indication if the effort to translate is worth. It is important to mention that only the interfaces in HTML are translated. For this reason, the effort is small and the results presented later on in this work indicate it is worth. A second benefit of this analysis is to guide the authors in deciding which new simulators to develop; this has already been happening as mentioned later on in this work.

The first simulator was published in May.2015 and the first English version in April.2016. This time frame yields an important consequence for this work – data are available for e very limited number of years. Since the authors are comparing accesses to pairs of simulators – the original in Portuguese and the corresponding English version, the limitation grows bigger because it is necessary to examine when both are published.

It is important to mention that all OERs on the collection are included on the database of MERLOT – Multimedia Educational Resources for Learning and Online Teaching (https://www.merlot.org).

A final comment in this section respects the students participation, which was in a voluntary basis when the students suggest, in opposition to the work developed by Andone et al. (2020) where the creation of OERs were “renewable” assignments in contrast to “disposable” tasks. Their work addresses different types of OERs - multimedia objects, while this work focuses on simulators.

2.2 Available Data

The IR runs on a relational Database Management System (DBMS). A unique sequential number identifies each and all contents. The webserver is Apache that records accesses to the id numbers of the contents with the time stamps and the IP addresses where they come from. Every hour, the Apache server log is scanned to identify the numbers and the IP addresses, that are converted to countries using GeoIP Free. This information is stored on a database table that is organized by year-month-content id-country-number of accesses; accesses are added every time the log is scanned. This table contains information for all digital contents on the IR. Data on accesses are publicly available on the IR.

To organize data, a dataset was created compiling numbers of accesses for each original in Portuguese and the corresponding English version. The condition to be included in the dataset is that both had to had been available from Jan 01 to Dec.31 of the year under consideration. In order to have data on more simulators, the starting year was decided to be 2018. The dataset is available on the IR – Pavani and Temporão (2024). Data on the dataset start on Jan.01.2018 for the first 10 sets (Portuguese and English) and end on Dec.31.2023. Each set is in a sheet. The remaining 10 sheets have later starting dates due to the dates of publication of their English versions. Only the first 10 sets are examined in this work. The 10 sets under consideration are:

- Set 01 – Nyquist Stability Criterion;
- Set 02 – Root Locus Plot;
- Set 03 – Bode Plot for First and Second Order Systems;
- Set 04 – Fourier Transform of a Continuous Time Pulse Function;
- Set 05 – Inverted Pendulum;
- Set 06 – RL Series Circuit;
- Set 07 – RC Series Circuit;
- Set 08 – First Order Systems;
- Set 09 – Combinations of Continuous Time Functions;
- Set 10 – Combinations of Discrete Time Functions.

The 10 sets are clearly related to courses on Signals & Systems (S&S), Electric & Electronic Circuits (E&EC) and Control Systems (CS). The courses are core of the three engineering curricula mentioned before.

Table 1 classifies the sets according to the courses whose students are expected to use them.

Table 1. Classification of simulators according to the courses they support

Set	S&S	E&EC	CS
1			✓
2			✓
3		✓	✓
4	✓		
5			✓
6	✓	✓	✓
7	✓	✓	✓
8	✓	✓	✓
9	✓		✓
10	✓	✓	✓

The analysis of the Table 1 delineates several discernible patterns: (1)The numbers of simulators across the three courses are different; (2) Noteworthy, is the cross-applicability of specific topics, initially tailored for the E&EC course, such as RC Series Circuit and RL Series Circuits, which serve a pedagogical purpose in the CS and S&S domains, facilitating the comprehension of 1st Order Linear Systems dynamics; and (3) These topics can have recurrent utilization; this is evident when parameters and/or functions undergo modifications.

3. DATA ANALYSIS

Data were extracted from the dataset and organized in 10 tables, one for each set. Since all sets were observed from 2018 to 2023, all tables have the same numbers of rows. Each table is divided in two vertical blocks –

one for the numbers of countries that accessed the two versions and the other for the numbers of accesses. Objects in Portuguese are treated separately from their English versions – one column for each in each block. The tables reorganize data in the first 10 data sheets of the dataset. The last row of each table contains the averages of the numbers in the other rows in the same columns of the table.

Tables 2 and 3. Numbers of Countries and Numbers of Accesses Sets 1 (Nyquist Stability Criterion) and 2 (Root Locus Plot)

	Countries		Accesses	
	pt	en	pt	en
2018	8	14	359	148
2019	14	28	468	255
2020	9	26	761	205
2021	11	18	687	229
2022	8	18	392	130
2023	10	14	494	196
μ	**10.00**	**19.67**	**526.83**	**193.83**

	Countries		Accesses	
	pt	en	pt	en
2018	8	62	1,026	3,491
2019	7	65	1,353	6,611
2020	8	77	1,272	7,884
2021	19	61	2,141	2,365
2022	10	50	605	1,313
2023	8	56	907	1,904
μ	**10.00**	**61.83**	**1,217.33**	**3,928.00**

Tables 4 and 5. Numbers of Countries and Numbers of Accesses ts 3 (Bode Plot for 1st and 2nd Order Systems) and 4 (Fourier Transform of a Continuous Time Pulse Function)

	Countries		Accesses	
	pt	en	pt	en
2018	15	19	2,127	99
2019	21	32	3,415	247
2020	25	23	2,556	175
2021	28	42	2,457	332
2022	25	66	1,754	1,167
2023	28	81	2,653	3,464
μ	**23.67**	**43.83**	**2,493.67**	**914.00**

	Countries		Accesses	
	pt	en	pt	en
2018	14	6	315	44
2019	6	8	320	68
2020	5	23	233	84
2021	8	25	139	99
2022	9	27	282	135
2023	10	12	161	117
μ	**8.67**	**16.83**	**241.67**	**91.17**

Tables 6 and 7. Numbers of Countries and Numbers of Accesses Sets 5 (Inverted Pendulum) and 6 (RL Series Circuit)

	Countries		Accesses	
	pt	en	pt	en
2018	14	18	3,374	149
2019	16	12	3,515	91
2020	16	17	4,401	323
2021	16	7	4,374	173
2022	12	21	2,863	203
2023	15	28	3,029	600
μ	**14.83**	**17.17**	**3,592.67**	**256.50**

	Countries		Accesses	
	pt	en	pt	en
2018	8	9	609	23
2019	13	5	990	40
2020	7	14	1,174	63
2021	16	16	1,200	50
2022	14	10	1,100	67
2023	16	26	1,053	160
μ	**12.33**	**13.33**	**1,021.00**	**67.17**

Tables 8 and 9. Numbers of Countries and Numbers of Accesses Sets 7 (RC Series Circuit) and 8 (First Order Systems)

	Countries		Accesses	
	pt	en	pt	en
2018	13	6	485	32
2019	9	4	1,219	42
2020	10	6	980	45
2021	15	5	845	19
2022	13	4	961	21
2023	11	25	740	182
μ	**11.83**	**8.33**	**871.67**	**56.83**

	Countries		Accesses	
	pt	en	pt	en
2018	6	3	185	25
2019	7	3	286	28
2020	9	4	457	47
2021	14	6	691	37
2022	11	7	516	45
2023	12	20	453	89
μ	**9.83**	**7.17**	**431.33**	**45.17**

Tables 10 and 11. Numbers of Countries and Numbers of Accesses Sets 9 (Combinations of Continuous Time Functions) and 10 (Combinations of Discrete Time Functions)

	Countries		Accesses	
	pt	en	pt	en
2018	5	3	1,384	105
2019	9	4	3,246	97
2020	6	4	981	47
2021	13	5	1,753	40
2022	8	7	1,404	79
2023	5	5	1,335	55
μ	**7.67**	**4.67**	**1,683.83**	**70.50**

	Countries		Accesses	
	pt	en	pt	en
2018	6	2	734	25
2019	6	3	1,415	98
2020	5	4	93	40
2021	7	6	127	54
2022	6	7	416	79
2023	9	5	213	71
μ	**6.50**	**4.50**	**499.67**	**61.17**

A quick look at the 10 previous tables indicates that the numbers are quite different from one set to the others, between the two languages in the same set and even in different years of the time series of one object. As it can be observed, tables 2 to 11 show larger numbers of accesses to objects in Portuguese compared to those in English (an exception is table 3); similar numbers of countries in both languages, except for tables 2 and 3, where the numbers of countries that accessed simulators in English surpass the ones in Portuguese. The subsections that follow analyze some characteristics of the data.

3.1 Analysis based on the averages

This subsection examines data based on the averages of numbers of accesses and numbers of countries where they came from for each set. These data are presented in Table 12 consolidating the last rows of Tables 2 - 11.

Table 12. Averages of Numbers of Countries and Numbers of Accesses

		μ Countries		μ Accesses	
		pt	**en**	**pt**	**en**
Set 1	**Nyquist Stability Criterion**	10.00	19.67	526.83	193.83
Set 2	**Root Locus Plot**	10.00	61.83	1,217.83	3,928.00
Set 3	**Bode Plot for 1st and 2nd Order Systems**	23.67	43.83	2,493.67	914.00
Set 4	**Fourier Transform of a Continuous Time Pulse Function**	8.67	16.83	241.67	91.17
Set 5	**Inverted Pendulum**	14.83	17.17	3,592.67	256.50
Set 6	**RL Series Circuit**	12.33	13.33	1,021.00	67.17
Set 7	**RC Series Circuit**	11.83	8.33	871.67	56.83
Set 8	**First Order Systems**	9.83	7.17	431.33	45.17
Set 9	**Combinations of Continuous Time Functions**	7.67	4.67	1,683.83	76.50
Set 10	**Combinations of Discrete Time Functions**	6.50	4.50	499.67	61.17

The inspection of Table 12 yields interesting information:

01. Across six instances (sets 1 – 6), the averages of the numbers of countries accessing the English versions exceed these of the originals in Portuguese. Notably, this disparity reaches up to six times higher in one case and nearly double in three others.
02. The sets with highest average access numbers, both in Portuguese and in English, predominantly belong to the six aforementioned cases, with the exception of set 4, which typically pertains to S&S courses.
03. Four out of the six sets are utilized in Control Systems courses at PUC-Rio and globally.
04. Sets 6 and 7 are used in topics of the E&EC courses at PUC-Rio and all over the world.
05. Sets 4, 9 and 10 are used in S&S courses at PUC-Rio and all over the world
06. The English version of the Root Locus Plot, as depicted in Table 2, garnered accesses from 50 or more countries over the six-year period under scrutiny.
07. Except for the CS simulators, the others have significantly higher numbers of accesses to the Portuguese originals when compared to the English version.
08. As can be observed, sets 1-3, which exhibited greater diversity in access across countries, are centered on topics related to the knowledge base of CS.

09. Consistency in Accesses per Country: Except for Sets 2 and 3, where the numbers of countries accessing content in English exceed those accessing content in Portuguese, there generally is consistency in the number of accesses per country between the two languages.
10. Temporal Trends: While there are fluctuations in the numbers of accesses over the years, there does not seem to be a clear upward or downward trend across sets. However, some sets show an increase in accesses over time (e.g. Sets 1 and 2), while others show more variability (e.g. Sets 5 and 6).
11. Impact of Content Type: Different types of content (e.g. Nyquist Stability Criterion, Fourier Transform) seem to attract different levels of interest, as evidenced by variations in access numbers across sets.

The observed patterns in access numbers appear to be closely tied to the topics addressed by the simulators. Notably, CS emerges as the subject area with the highest access numbers, manifesting in some cases through the Portuguese versions and in others through their English counterparts.

3.2 Analysis based on specific accesses to some sets

This subsection examines sets in which one or both objects garnered a minimum average of 1,200 accesses, denoting at least 100 average accesses per month. This choice was to be able to analyze more data since the time frame of this work is small due to the dates of the creation of the simulators.

Four sets are under examination – Root Locus Plot, Bode Plot for 1st and 2nd Order Systems, Inverted Pendulum, and Combinations of Continuous Time Functions. To conduct this analysis, the tables in section 2 and the dataset are examined.

For a more comprehensive understanding of the relationships between the numbers of countries and of accesses across both versions of each simulator, a bi-dimensional graphic was generated for each set. Within each graph, the horizontal axis denotes the number of countries, while the vertical axis represents the number of accesses. Bullets indicating the Portuguese original are depicted as round, while those representing the English version are triangular. Control Systems emerges as the topic with the highest number of accesses for certain instances, observed in both the Portuguese and English versions.

3.2.1 Set 2 – Root Locus Plot and Set 3 – Bode Plot for 1st and 2nd Order Systems

The two figures that follow are bi-dimensional graphics for sets 2 and 3.

Figures 1 and 2 are the bi-dimensional graphics for sets 2 and 3.

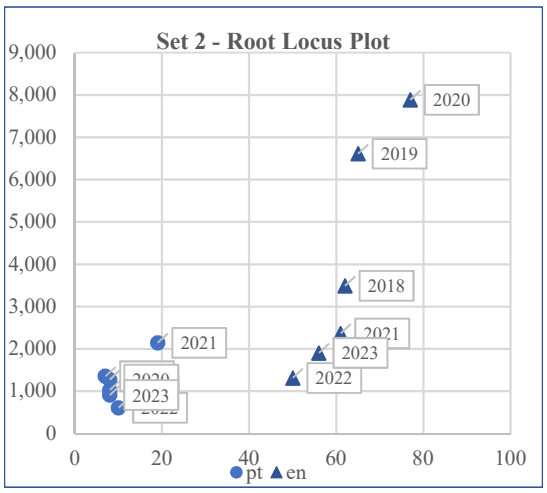

Figure 2. Numbers of countries and of accesses to the Root Locus Plot objects in pt and en 2018-2023

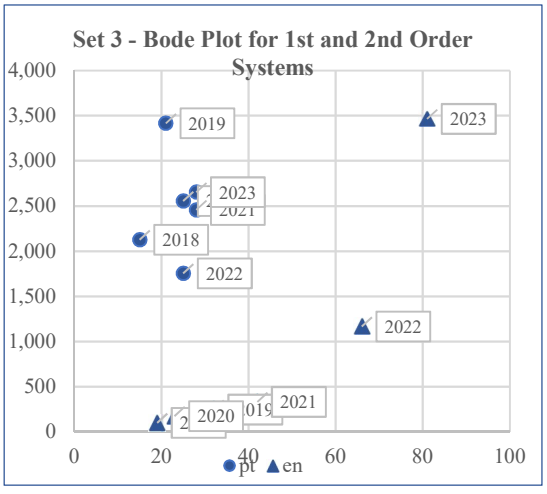

Figure 2. Numbers of countries and of accesses to the Bode Plot for 1st and 2nd Order Systems objects in pt and en 2018-2023

The first observation about Set 2 – Root Locus Plot is that average number of countries that accessed the English version is six times higher than the corresponding average to the Portuguese original. The average

number of accesses is three times as high. In order to examine the numbers with more detail, Table 2 and the dataset are considered. Figure 1 shows the numbers of countries and corresponding numbers of accesses for each of the years 2018-2023 as presented in Table 2. Figure 1 yields a better perception of the relations among numbers of countries and of accesses. A comment can be made: the bullets of countries and of accesses are very close together in the pt original and far apart from the bullets in the en version; bullets of the en version are far apart from each other too. This characteristic indicates that both the numbers of countries and of accesses are higher for the en version. Concerning Set 3 – Bode Plot for 1st and 2nd Order Systems, it is easily seen that the numbers of accesses to the Portuguese original are higher than the ones to the English version. On the other hand, the numbers of countries are either similar or higher for the English version.

Both sets are used in CS courses and the Bode Plot is used in E&EC too.

3.2.2 Set 5 – Inverted Pendulum and Set 9 – Combinations of Continuous Time Functions

Figures 3 and 4 are the bi-dimensional graphics for sets 5 and 9.

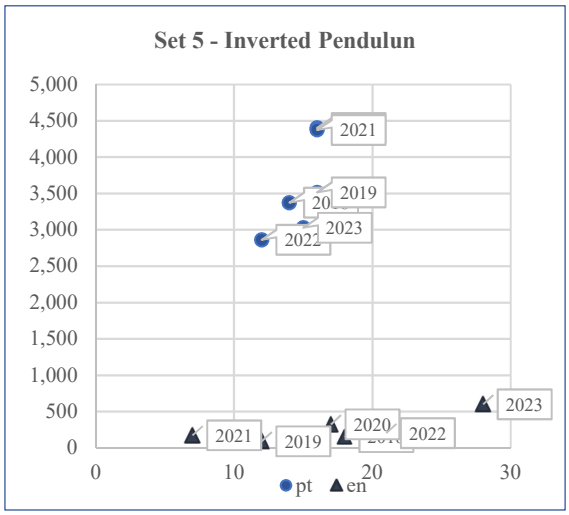

Figure 3. Numbers of countries and of accesses to the Inverted Pendulum objects in pt and en 2018-2023

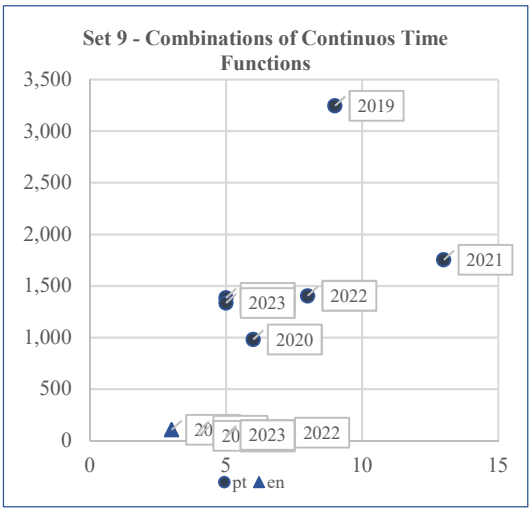

Figure 4. Numbers of countries and of accesses to the Combinations of Continuous Time Functions objects in pt and en 2018-2023

The first observation about Set 5 – Inverted Pendulum is that the numbers of accesses to the English version are much lower than the ones to the Portuguese original – the average number of accesses for Portuguese is 14 times the average for English. The high numbers of accesses in the six years under consideration are due to Brazil. This can easily be seen by averaging the number of accesses from Brazil as presented in the dataset; this number is 3,367.33. On the other hand, the average number of countries that accessed the objects is slightly higher for the English version. A possible explanation for these numbers is that the Inverted Pendulum, in general, is addressed in Control Systems courses in Brazil. Concerning Set 9, it is clear that both the numbers of countries and the numbers of accesses are higher for the Portuguese original. Once more, the higher numbers are due to Brazil. Combinations of functions, though necessary in many courses, is mainly focused in Systems & Signals courses.

3.2.3 Comments on the Four Sets

When the four sets are examined together, it is clear that the numbers of accesses to the CS simulators are higher regardless of where they originated from. If the numbers of countries are compared, the differences are not significant except in Set 9.

Accesses from Brazil to the English versions happen in significant numbers. This can easily be verified examining the dataset.

4. FINAL ANALYSIS, COMMENTS AND CONCLUSION

The numbers presented and discussed in this work were gathered from the Maxwell System using Open Access applications. The dataset was organized and made available in Open Access to make analysis and understanding easier.

The data present a complex picture of user behavior and preferences. While Portuguese content seems to be more popular overall, there are variations across sets and languages that warrant further investigation. Additionally, factors such as the use in undergraduate courses and impact factor of the repository over the world may influence these patterns and should be considered in a more comprehensive analysis.

Another factor that may have influenced data is the COVID pandemic. From 2020 to 2021 or 2022, many universities switched to online learning and online materials may have become more useful.

The Research Question seems to be answered, revealing differences in access patterns, both in terms of the number of accesses and the countries of origin, between simulators in Portuguese and their English versions. Additionally, variations in access numbers are noted across different topics addressed by the simulators, with CS emerging as a favored subject.

The CPLP – Community of Portuguese Language Countries (https://www.cplp.org/) includes Angola, Brazil, Cape Verde, Guinea-Bissau, Equatorial Guinea, Mozambique, Portugal, São Tomé and Príncipe, and Timor-Leste as Portuguese-speaking countries, all of which have Portuguese as one of the official languages. It is important to remark it is the only official language of Brazil that has over 200 million inhabitants.

Given the abundance of data available for examination, the authors intend to investigate the presence of Portuguese-speaking countries among the accesses. This will be discussed in future works. Another topic that will deserve attention comes from the fact that last March the Maxwell System deployed interfaces that can be switched between Portuguese and English. This may influence accesses from non-Portuguese speaking countries. Further work will analyze this. And last but not least, it is necessary to add more data to the analysis. This will be accomplished as time goes by, since it is necessary that the Simulators be exposed for more time.

REFERENCES

ADL, 2004, *ADL Guidelines for Creating Reusable Content with SCORM 2004"*, July 2008, available https://adlnet.gov/assets/uploads/ADLGuide2004_Final_073108.pdf. Last accessed May 2022.

Alsubaie, M. and Alshawi, M., 2009, Reusable Objects: Learning Object Creation Cycle, *Proceedibgs of the Second International Conference on Developments in eSystems Engineering*, 2009, pp. 321-325, doi: 10.1109/DeSE.2009.63. Last accessed May 2022.

Andone, D., Mihaescu, V., Vert, S. Ternauciuc A. and Vasiu, R. (2020), Students as OERs (Open Educational Resources) co-creators, 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT), Tartu, Estonia, 2020, pp. 34-38, doi: 10.1109/ICALT49669.2020.00017.

Institute of Electrical and Elesctronics Engineers, (2002), *IEEE Standard for Learning Object Metadata*, in *IEEE Std 1484.12.1-2002* , vol., no., pp.1-40, 6 Sept. 2002, doi: 10.1109/IEEESTD.2002.94128. Last accessed in May 2022.

Lynch, C. (2003), Institutional Repositories: essential infrastructure for scholarship in the digital age, ARL Bimonthly Report, nº 226, USA, 2003.

Pavani, A.M.N and Temporão, G.P., (2024), Comparative Accesses to Simulator OERs in Portuguese and their English Versions Available from Open Educational Resources@PUC-Rio – Extended Set and Time Frame, doi.org/10.17771/PUCRio.ResearchData.65826.

Statista (2023), The Most Spoken Languages Worldwide in 2023, https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/. Last accessed in January 2024.

Wang, T. and Towey, D., Open educational resource (OER) adoption in higher education: Challenges and strategies, 2017 IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Hong Kong, China, 2017, pp. 317-319, doi: 10.1109/TALE.2017.8252355. keywords: {Open Educational Resources;Conferences;Training;Government;Production;Technological innovation;Open Educational Resources (OERs);barriers to change;higher education},

Wiley, II, D.A. , (2000)*Learning Object Design and Sequencing Theory*, PhD dissertation presented at Brigham Young University, United States, June 2000, available https://opencontent.org/docs/dissertation.pdf. Last accessed May 2022.