(Accepted for publication in World Patent Information)

Ibero-American patent information with Latipat: public search systems and their differences

Authors:
Björn Jürgens[1] ; Victor Herrero-Solana[2]

[1] CITPIA PATLIB Centre
Agency of Innovation and Development of Andalusia
Seville, Spain
E-Mail: bjurgens@agenciaidea.es

[2] SCImago-UGR (SEJ036)
Universidad de Granada
Granada, Spain
E-Mail: victorhs@ugr.es

Abstract

The Latipat project unifies patent information from 20 Ibero-American countries making this content available and free of cost via several major public patent search systems: namely Espacenet, Patentscope and Invenes. These different accesses are presented and their data coverage is analysed. A comparison showed that, although the patent information comes from the same origin, the country coverage and actuality of the patent information varies substantially in the three compared search systems. Regarding the technology coverage, an exemplary search for nanotechnology patents revealed different results depending on the keyword and classification support of each search system.

Keywords

Latipat; Iberoamerican; Latinamerican; Patent Information; Coverage; Patent Database

1 Introduction

The Ibero-American countries form an area of great interest to industry. On one side stands Latin America, a market with over 580 million people and growing economies that actively promote policies to stimulate innovation [9]. On the other side lies the countries of the Iberian Peninsula which, despite their economical crisis, remain important markets due to their cultural and linguistic connection to Latin America and consequent position as a gateway between the European and Latin American markets and vice versa [2-3].

For innovative companies wishing to invest in this region, having access to patent information of these countries is essential – not only in order to avoid patent infringements but also to monitor, discover and exploit new technological opportunities. Taking the latter into account, the Spanish Patent and Trademark Office (SPTO), the European Patent Office (EPO), the World Intellectual Property Organization (WIPO) and patent offices from 19 Latin American countries[1] created the Latipat project in 2003. The aim of this project was to offer the public a centralized and free access to Ibero-American patent information in Spanish and Portuguese languages [15].

By taking part in this project, under the lead of the SPTO, the participating patent offices committed themselves to convert their patent information to a unified data format[2] and regularly send it to a central database, the *Latipat Central Repository* [1]. As visualized in Figure 1, this data is then integrated by WIPO in their patent search system Patentscope and by EPO in their master documentation database *Docdb* (accessible publicly via the Espacenet-Worldwide database) and to a dedicated server administered by the SPTO (Espacenet-Latipat database). Furthermore the STPO imports the Latipat data into their own patent search system Invenes (Invenes-Latipat database).

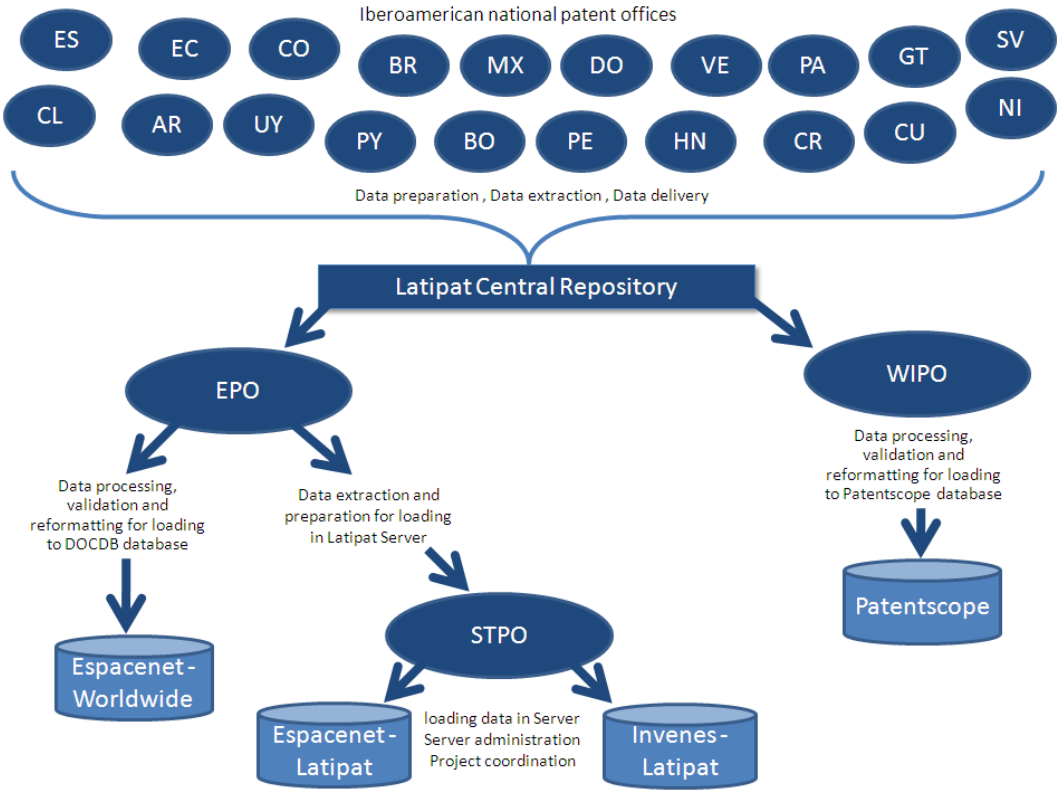

Fig. 1 : Latipat Data flow

[1] Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, Venezuela

[2] WIPO standards ST32 , ST33 und ST36

Therefore the Latipat patent information is made available by the patent authorities WIPO, EPO and SPTO via the following databases:

- Patentscope (https://patentscope.wipo.int)
- Espacenet-Worldwide (http://worldwide.espacenet.com)
- Espacenet-Latipat (http://lp.espacenet.com)
- Invenes (http://invenes.oepm.es)

All four databases are accessible online and free of charge which means that the patent searcher has the choice of four systems which nominally have the same data and purpose. In order to find out if the systems show any significant differences which might be of interest for a user, the patent data, country coverage and some specific patent search functionalities were compared. In this context the authors compared the ability of the search systems to launch sector wide searches of an emerging technological field like Nanotechnology, since this paper is related to a project about this sector (see chapter 5). General differences in functionality and features of the Espacenet and Patentscope Search system have been discussed in an earlier paper of the authors [5] and therefore are not mentioned here.

2 Materials and Methods

The four different Latipat web access points were compared and their language support analysed. For the data and country coverage comparison, statistical patent data was downloaded from the patent authority's statistical public data sources [16-19] and in a second step analysed and compared using the following criteria:

- number of countries covered
- number of bibliographic records
- number of full patents (in PDF)
- timeliness of the patent data (latest update year).

In order to compare the coverage of the Latipat sources of specific technological sectors, a comparison of a patent search for Nanotechnology-related patents was conducted and the following search strategies were used:

- Searching for patents using the Patent Classification *B82Y* which is the dedicated classification describing Nanotechnology patents used in the international patent classification (IPC) and cooperative patent classification (CPC) [8].
- Searching for patents with keywords using the truncated term *Nano** in title and abstract.
- Searching for patents with keywords using a specific Nano-related search query in title and abstract. This query, developed by Maghrebi (2011), excludes some terms which have the Nano string but are not relevant, and includes some terms which are related to nanotechnology but do not have the Nano string in the name (e.g. "supramolecular" or "quantum dot").

3 Results

3.1 Latipat Search Interfaces

3.1.1 Patentscope

The Latipat data is integrated into the database of the Patentscope search system and thus can benefit from all the features that this system has to offer (e.g. the built-in statistical analysis). To search in Latipat countries the user can restrict the offices to be searched via checkboxes (Figure 2), where the whole Latipat country group or single countries can be selected. This showed to be an advantage compared to the other Latipat accesses making it very easy to restrict patent searches to specific countries or exclude countries from the search.

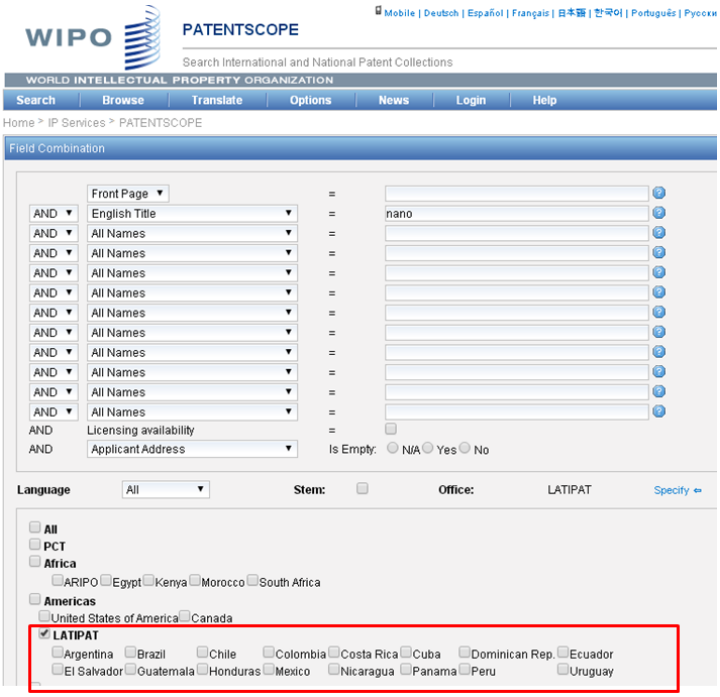

Fig 2: Latipat via Patentscope Screenshot

3.1.2 Espacenet

As illustrated earlier (Figure 1) the Latipat patent data is integrated by the EPO into their database *Docdb* which is the data source for Espacenet via its worldwide database. Hence, the Ibero-American patent information forms part of this database, although the keyword search is only possible on a reduced set of patents which have English titles or abstracts (see chapter 3.2).

Furthermore the Latipat data is accessible via the dedicated Espacenet-Latipat webpage that provides the patent data via the database LP-Espacenet. As one of the main advantages using this database we found that keyword search in title/abstract in original language (Spanish/ Portuguese) is possible and also full text search is available for Spanish patents. As a downside, contrary to the Espacenet-worldwide database, the database LP-Espacenet does not support searches using the Cooperative Patent Classification (CPC) as can be observed in Figure 3 (see also chapter 3.4).

All other search features and limitations of the Espacenet search system, as discussed in an earlier article [5], also apply to the Espacenet-Latipat webpage.

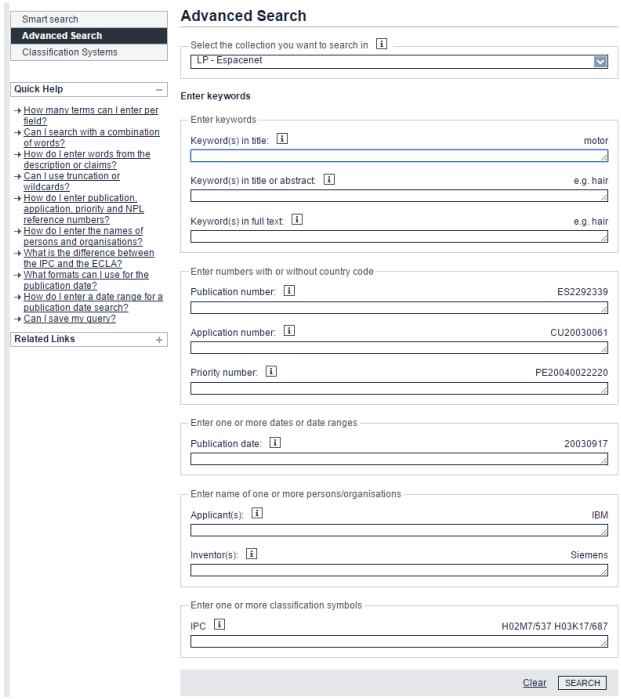

Fig 3: Espacenet-Latipat Advanced Search Interface

3.1.3 Invenes

Finally, Invenes [18] is the public patent search system of the Spanish Trademark and Patent Office (SPTO) and provides access to two separate databases: *Interpat* which contains patents filed in Spain and *Latipat*, which contains data from the Latipat countries (Figure 4).

Like Espacenet-Worldwide, it allows searches using the CPC classification, which can be useful for individual sectors or technologies that are specifically covered by this classification (as shown in the case study in chapter 3.4.).

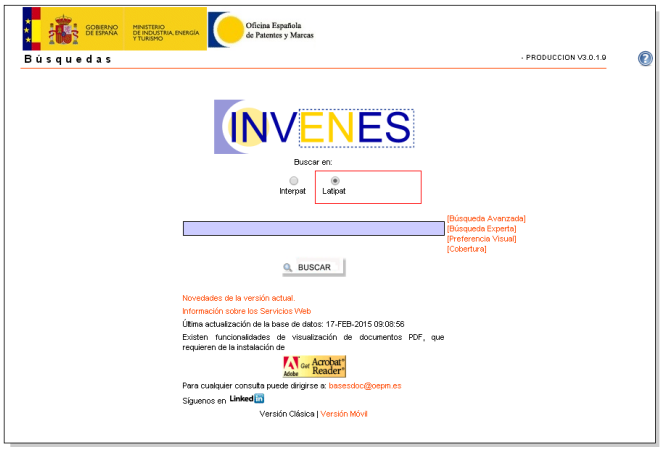

Fig 4: Screenshot of Invenes search system with selected Latipat database

3.2 Language support

One of the aims of the Latipat project was to make patent information from Iberoamerica more accessible to their users by offering the patent information in their language. As a consequence all compared search systems not only offer an English user interface but also an Interface in Spanish and in the case of Espacenet-Latipat and Patentscope also a Portuguese user interface (since Brazil forms part of the Latipat project). The exception is Invenes, which provides only a Spanish interface.

More important than the interface language is that patents in Latipat search systems have to be searched with Spanish keywords in title or abstract (and Portuguese in the case of Brazilian Patents). This enables users of the Latipat countries search in their native language, but requires a previous translation for all other users.

As shown in Table 1, searchable English titles and abstracts are only available in Patentscope on Latipat patents which have a PCT application and in Espacenet-Latipat and Espacenet-Worldwide on a reduced set of patents from Mexico, Argentina, Cuba and Spain [20] and with patents which have a patent in their family with English Title/Abstract (e.g. an US patent).

		ESPACENET LATIPAT	ESPACENET WORLDWIDE	PATENTSCOPE LATIPAT	INVENES LATIPAT
Interface Language	English	yes	yes	yes	no
	Spanish	yes	no	yes	yes
	Portuguese	yes	no	yes	no
Searchable Patent Data	English	only MX,AR,CU,ES	only MX,AR,CU,ES	only WO	no
	Spanish	yes	no	yes	yes
	Portuguese	only BR	no	only BR	no

Table 1 : Language support for Latipat interface and patent data

3.3 Country Coverage

Although all Latipat patent search systems described in this study use the same data origin, the amount of data available turned out to differ – and indeed significant differences were found in some countries.

With regards to total numbers of records, all compared databases had similar totals, with Espacenet-Latipat having nearly 2.5 million bibliographic patent records –the highest total coverage of the 20 Latipat countries, followed by Espacenet-Worldwide, Patentscope and Invenes (Table 2).

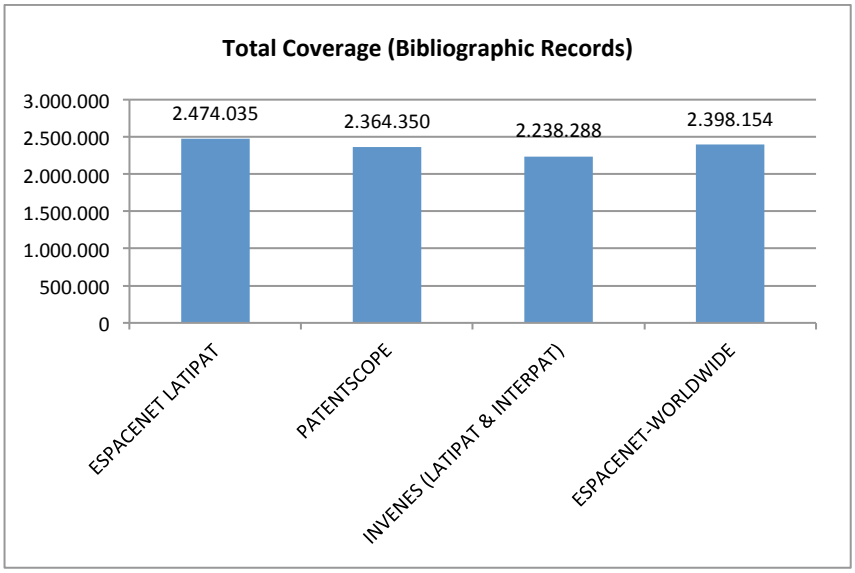

Table 2: Latipat countries coverage comparison – Total coverage (Bibliographic Records)

Regarding full document[3] coverage, Invenes was found to have the highest total (Table 3). This is mainly due to the fact that Invenes contains all Spanish patent documents in full text (via its Interpat database), and that the Spanish patent collection is the biggest of all Latipat countries in terms of record numbers.

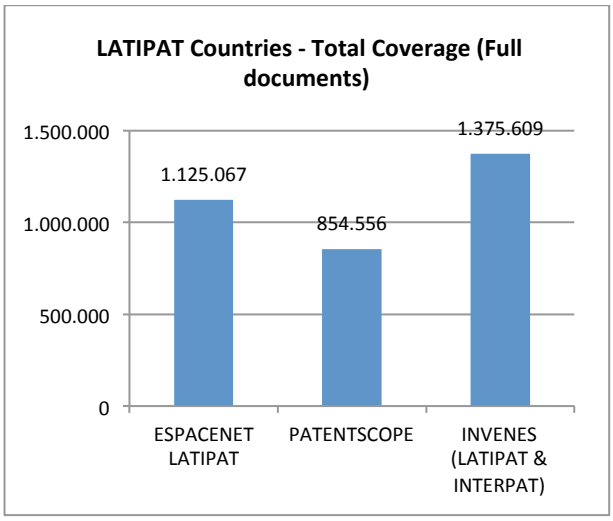

Table 3: Latipat countries coverage comparison – Total coverage (Full documents)

Regarding the specific country coverage, Invenes showed to have the best coverage of Spanish patents and significantly better bibliographic coverage for Argentina, Colombia, Mexico and Venezuela than the other databases, whereas Espacenet-Latipat had the best coverage in the Central American countries and in Paraguay, Uruguay and Ecuador (as showed in Table 4, where the significantly highest country coverage with more than 10% difference is marked in inverted colours).

Remarkably, although using the same search interface, Espacenet-Latipat and Espacenet-Worldwide also had several coverage differences. In most cases Espacenet-Latipat had significantly more patent records for the Latipat countries, and in the case of Bolivia, Paraguay and Venezuela both Espacenet-Worldwide and Patentscope did not cover the countries at all. This could be explained by the fact that the patent data of these countries is not regularly updated by their respective patent offices with their last updates executed more than five years ago (see also Table 5).

Invenes, on the other hand does not include a major country in its database: Brazil. The reason being that Invenes has no compatibility with non-Spanish language patent information (i.e. Portuguese, as is the case of Brazilian patent data) [10].

[3]Usually the PDF version (or an image file) of the full patent document as it is in its original, printed out version.

	PATENTSCOPE	ESPACENET LATIPAT	ESPACENET-WORLDWIDE (DOCDB)	INVENES (LATIPAT & INTERPAT)
Argentina	134.940	139.400	139.554	161.982
Bolivia	n.a.	316	n.a.	281
Brazil	534.812	506.059	599.232	n.a.
Chile	3.826	59.098	9.754	56.700
Colombia	12.028	23.620	22.282	32.969
Costa Rica	691	7.684	7.090	7.300
Cuba	2.815	4.111	4.391	4.488
Dominican Republic	2.361	3.247	3.243	3.270
Ecuador	2.858	19.731	10.112	11.006
El Salvador	1.582	2.147	1.591	1.626
Guatemala	5.949	11.778	6.666	7.239
Honduras	286	1.131	1.022	1.094
Mexico	216.229	264.960	238.999	292.020
Nicaragua	197	490	197	481
Panama	2.312	4.510	2.386	3.412
Paraguay	n.a.	1.550	n.a.	231
Peru	6.415	26.458	18.541	26.242
Spain	1.436.353	1.356.087	1.322.287	1.587.657
Uruguay	696	14.499	10.807	12.821
Venezuela	n.a.	27.159	n.a.	27.469
Total	**2.364.350**	**2.474.035**	**2.398.154**	**2.238.288**

Table 4: Country coverage comparison

Comparing the data actuality (by checking the last updated year) Invenes and Espacenet (Latipat & Worldwide) were shown to have rather recent data from the Latipat countries (latest update year 2014 or 2015 marked in grey in table 5), whereas in the case of Patentscope most country collections were not up to date (latest update pre-2014).

Some countries, namely Bolivia, Ecuador, Nicaragua, Panama, Paraguay and Venezuela, do not have recent patent data in all compared search systems. This is most probably due to these countries not sending data to the "Latipat Central Repository" for reasons not known to the authors.

	PATENTSCOPE	ESPACENET LATIPAT	ESPACENET-WORLDWIDE (DOCDB)	INVENES (LATIPAT & INTERPAT)
Argentina	2013	2014	2014	2014
Bolivia	n.a.	2008	n.a.	2008
Brazil	2013	2015	2015	n.a.
Chile	2008	2014	2014	2014
Colombia	2010	2014	2014	2014
Costa Rica	2013	2014	2014	2014
Cuba	2012	2014	2014	2014
Dom. Republic	2012	2015	2014	2015
Ecuador	2009	2012	2012	2012
El Salvador	2013	2014	2014	2014
Guatemala	2011	2014	2014	2014
Honduras	2010	2014	2014	2014
Mexico	2011	2014	2014	2014
Nicaragua	2009	2009	2009	2009
Panama	2010	2010	2010	2010
Paraguay	n.a.	1995	n.a.	1995
Peru	2011	2015	2015	2014
Spain	2014	2015	2015	2015
Uruguay	2013	2014	2014	2014
Venezuela	n.a.	1997	n.a.	1997

Table 5: Patent Data actuality comparison (timeliness data marked with dark background)

3.4 Technology Coverage – Nanotechnology Case Study

Regarding the search using keywords, searching with the truncated keyword "Nano*" in the patent abstracts naturally retrieved more documents than in the patent titles, but also included results which are not relevant, since not all terms beginning with Nano are also related to nanotechnology as defined in the literature [7,11].

The Nanotechnology search using the specific CALQ Query [7] showed that Invenes retrieved most patent records, whereas searching with this Query was not possible in Latipat-Espacenet since the Espacenet search platform restricts the number of maximum keywords used in one query. In addition, Espacenet search does not allow the user to separate queries and unite them at a later stage since it does not allow combining saved queries in the search history [5].

Table 6 shows a comparison of the obtained search results in the different search systems, where the following measures had to be taken into account:

- Since the Spanish patents in Invenes are stored on a separate database (Interpat) the search for Nanotech patents had to be done separately in both, Interpat and the Invenes-Latipat database.
- Espacenet-Worldwide was excluded from this specific comparison because of the limitations of the Espacenet Search Interface which does not allow to search in all Latipat countries at once.
- Since in Patentscope the user has to choose the language of the Title/Abstract to be searched, the search was done separately in Title/Abstracts in English and Spanish and then the results combined.

		B82Y	Nano* in Title	CALQ in Title	Nano* in Abstract	CALQ in Abstract
INVENES	INTERPAT (SPAIN)	1798	1007	1116	2134	2335
	LATIPAT (LATAM COUNTRIES)	1367	1036	1110	2248	2394
	INVENES TOTAL (SPAIN + LATAM)	**3165**	**2043**	**2226**	**4382**	**4729**
ESPACENET	**ESPACENET LATIPAT (SPAIN + LATAM)**	**260**	**2387**	**n.a.**	**4801**	**n.a.**
PATENTSCOPE	PATENTSCOPE (SPAIN + LATAM Title/Abstract EN)	199	506	506	1017	987
	PATENTSCOPE (SPAIN + LATAM Title/Abstract ES)	199	1508	1508	2326	2271
	PATENTSCOPE SPAIN + LATAM TOTAL	**199**	**1528**	**1528**	**2499**	**2430**

Table 6: Technology coverage comparison (Nanotechnology)

By using patent classifications, which are most used if entire technological fields are searched in patent databases [12, 13], the search for nanotech patents retrieved far less patents in Patentscope and Espacenet-Latipat than in Invenes. That this is because most of the nanotech-related patent documents seemed to be classified with the B82Y symbol in the CPC classification instead of the more common IPC classification and that most IPC-classified nanotech patents are also retrieved with a CPC search as a snapshot search, shown in table 7, revealed. The reason that nanotech patents are classified with the B82Y in the CPC might be related to the ongoing effort of the patent offices which manage the CPC (EPO and the US patent office) to classify and reclassify patents from emerging technologies like nanotechnology.

In the following table we compared the number of patent records in the top 5 Latipat countries (according to their patent output) classified with CPC and IPC in the class B82Y. As we can see Invenes and Espacenet-Worldwide were the only databases which allowed searching using CPC and retrieved far more patent records in this particular classification.

Publication country	B82Y Class	INVENES-LATIPAT	ESPACENET LATIPAT	ESPACENET WORLDWIDE	PATENTSCOPE
ES	IPC	213	166	166	158
	CPC	1767	n.a.	1634	n.a.
MX	IPC	26	54	54	0
	CPC	859	n.a.	939	n.a.
BR	IPC	n.a.	84	84	39
	CPC	n.a.	n.a.	1204	n.a.
AR	IPC	1	2	2	2
	CPC	247	n.a.	260	n.a.
CL	IPC	1	1	1	0
	CPC	34	n.a.	38	n.a.

Table 7: Patent records classified with CPC/IPC class B82Y in the top 5 Latipat countries [4]

[4] Searched with country code (publication number) in combination with a truncated B82Y classification

4 Conclusion

Although, thanks to the Latipat project, Ibero-American patent information is unified and made available free of cost via the public search systems Espacenet, Patentscope and Invenes, this paper showed that the data available on these platforms is not entirely the same. Significant differences were found in country coverage, language support and search possibilities.

Regarding the country coverage some countries are not covered at all in some Latipat search systems (such as Brazil in the case of Invenes and Bolivia, Paraguay and Venezuela in the case of Patentscope and Espacenet-Worldwide). In general, Latipat-Espacenet showed to have the best country coverage in absolute numbers and also the best actuality of the patent data. However, Invenes has similar results with the exception of Argentina, Colombia, Cuba and Mexico where Invenes showed to have better coverage. Patentscope had less patent records in most Latipat countries except for the country collections of Brazil and Spain.

Concerning the language support, the user should be aware that when doing a keyword search, these have to be introduced Spanish (or in Portuguese if Brazilian patents are searched) since English keywords will only retrieve a reduced set of patents.

With regards to the search possibilities, some downsides of the compared search systems were revealed. Only Espacenet-Worldwide and Invenes allows searching using the Cooperative Patent Classificacion (CPC). It surprised us that Latipat Espacenet did not support the CPC as it does in its Espacenet-Worldwide database. For a non-classification patent search, Espacenet-Latipat can be recommended since it has the best overall Latipat country coverage and also gives access to patent information to "exotic" patent collections like Bolivia, Venezuela and Paraguay which, although not updated lately, are not present in Patentscope or Espacenet-Worldwide.

5 Acknowledgements

We would like to thank the Spanish Ministry of Education for funding the framework project „Technology Watch of Spanish Nanotechnology via its patents" (Project number: CSO2012-38801) for which this analysis was used.

References

1. Berlicki, G. (2014). Descripción general del flujo de datos. Espacio LATIPAT. https://www3.wipo.int/confluence/pages/viewpage.action?pageId=27754647 (Accessed 13.01.2015)
2. Fornés, G. and Cardoza, G. (2009). Foreign exchange exposure in emerging markets. A study of Spanish companies in Latin America", International Journal of Emerging Markets , Vol. 4 No. 1, pp. 6-25.
3. Guillén, M. (2005). The rise of Spanish multinationals: European business in the global economy. Cambridge University Press.
4. Huang, Z., Chen, H., Chen, Z. K., Roco, M. C. (2004). International nanotechnology development in 2003: Country, institution, and technology field analysis based on USPTO patent database. Journal of nanoparticle Research, 6(4), 325-354.
5. Jürgens, B., Herrero-Solana, V. (2015). Espacenet, Patentscope and Depatisnet: A comparison approach. World Patent Information. In Press.
6. Kay L (2008), Developing nanotechnology in Latin America, Journal of Nanoparticle Research Feb 2009; 11(2): 259–278.
7. Maghrebi M. et al. (2011) A collective and abridged lexical query for delineation of nanotechnology publications, Scientometrics, Vol: 86 2011, 15~25
8. Michalitsch R. (2013) Nanotechnology and patents. European Patent Office Publications, Munich

9. Rivas G, Rovira S (2014), Nuevas instituciones para la innovación, CEPAL – Economic Commission for Latin America and the Caribbean, United Nations Publications
10. Roncero, J. Personal communication, 26.02.2015, Spanish Patent and Trademark Office
11. Porter AL, Youtie J, Shapira P, Schoeneck D (2008) Refining search terms for nanotechnology. J NanopartRes10(5):715–728.
12. Leydesdorff, L. (2008). Patent classifications as indicators of intellectual organization. Journal of the American Society for Information Science and Technology, 59(10), 1582-1597.
13. Schmoch, U. (2008). Concept of a technology classification for country comparisons. Final report to the world intellectual property organisation (wipo), WIPO.
14. Tunzelmann, G. V. (1998). Localized technological search and multi-technology companies. Economics of Innovation and New Technology, 6(2-3), 231-256.
15. http://lp.espacenet.com/?locale=en_LP (Accessed: 20.01.2015)
16. http://sitadex.oepm.es/LatipatCobertura/initCobertura (Accessed: 25.01.2015)
17. https://patentscope.wipo.int/search/en/help/data_coverage.jsf (Accessed: 25.01.2015)
18. http://invenes.oepm.es/InvenesWeb/faces/busquedaInternet.jsp (Accessed: 25.01.2015)
19. http://www.epo.org/searching/data/data/tables/weekly.html (Accessed: 25.01.2015)
20. http://lp.espacenet.com/help?locale=es_LP&method=handleHelpTopic&topic=abstract (Accessed: 12.01.2015)