



Linking patterns in the European Union's Countries: geographical maps of the European academic web space

Jose Luis Ortega, Isidro F. Aguillo¹

Cybermetrics Lab, Centro de Información y Documentación Científica (CSIC), Madrid, Spain

Abstract

This national-level study intends to describe the existing relationships between the number of web pages, inlinks to and outlinks from Europe and national or internal links in the European Higher Education Area through a sample of 535 European universities web domains. Several geographical maps are introduced to summarize and visualize this statistical information. The main result shows that larger countries, in number of web pages, link less to the remaining European countries, while the smaller ones are characterized by their link profusion to the European zone. The great presence of national links in large and medium size countries confirms that the European academic web space is shaped by the aggregation of national sub-networks, while the similar low presence in small countries suggests that these are linked to another large country.

Keywords: Webometrics; European Higher Education Area; Link Analysis; Geographical maps

1. Introduction

Since the European Economic Community (EEC) foundation in 1957, the European Union (EU) has shown the clearest example of integration and cooperation in the Western Europe both in economic and political issues. This has favoured the building of a geographical space in which the member states are setting up a common identity in contrast to their historical national or local tradition. This duality between Europeanism and localism grants these countries an interesting feature in order to study them, because we may see integration levels in the European space.

These relationships among countries can be measured and visualized through multiple indicators as imports/exports from an economic point of view [1] or scientific collaboration from the S&T evaluation [2]. From

¹ Correspondence to: José Luis Ortega Priego, Cybermetrics Lab, Centro de Información y Documentación Científica (CSIC), Joaquín Costa, 22, 28002 Madrid, Spain. E-mail: jortega@cindoc.csic.es

José Luis Ortega and Isidro F. Aguillo

a webometric scope, these relationships may be represented through the hyperlink flows that point from one country to another. Hence, the links among the different European countries web spaces might be a good indicator to measure the European cohesion on the Web.

The World Wide Web was born and developed in the academic environment [3,4]. So the academic web is mature and it could be used to find relationships among European countries. Several studies have analysed and compared the academic web relationships at the country level, using the Web Impact Factor (WIF) in a recurrent way. For example, Almind and Ingwersen [5] compared the scientific visibility on the Web and on the print production from Denmark, Norway and Sweden, introducing WIF indicator. Smith also compared the Australasian [6] and the Latinamerican academic web space [7], founding a better WIF rate form countries with low web performance than the high ones. Similar results were obtained by Norouzi [8] in the Middle East academic web.

Other works have focused their attention on the relationships between different web elements (pages, links, domains, etc.) from several countries. Thelwall and Wilkinson [9] studied the web topology of three Anglo-Saxon academic networks: New Zealand, United Kingdom and Australia, following the Bow-tie model [10-12]. They found similar graph and power law properties although the size of the samples is quite different. Cothey [13] analysed the European Research Area on the web, and he also found that elements studied followed a power law distribution. From a visual approach, other studies have developed maps and graphs showing the link relationships among European countries. Polanco, Boudourides, et al. [14] used Multidimensional Scaling (MDS) maps to show the density and centrality of more than one thousand European university web sites. Musgrove, Binns, Page-Kennedy and Thelwall [15] used a probabilistic model to map the link relationship between European countries. Ortega, Aguillo, Cothey and Scharnhorst [16] analysed the European Higher Education Area web space and they found sub-networks or communities of national university networks, similar to the finding of Heimeriks and Van den Besselaar [17] which also mapped the European universities on the Web. However, the building of geographical maps in the Web research environment has not yet a significant development, although it has had in the Internet research [18,19].

2. Objectives

This study intends to show the existing relationships between the number of web pages, inlinks, outlinks and national or internal links in the European Higher Education Area (EHEA) at the country level through a sample of 535 European university web sites. We want to know how these countries are connected and if there is any pattern in the linking behaviour. This study also intends to analysis the link relationships among web spaces instead of web domains or sites where the internal structure is a considerable factor to know the out and inlink patterns in these spaces. We also want to introduce geographical maps to show these relationships and to study the suitability of these maps in the Web research environment.

3. Methodology

3.1. Data extraction

535 universities of the 14 European countries -EU except the Luxembourg universities because they are not included among the first 1,000- in 2004 were selected from the first 1,000 university domains ranked in Webometrics Ranking of World Universities web site (www.webometrics.info). This site ranks 3,000 universities according to two main criteria: size (number of pages and rich files) and visibility (number of incoming links). So we consider that both indicators let us to select the mainstream of the European academic web space. Then we

José Luis Ortega and Isidro F. Aguillo

extracted the number of pages and links between these 535 European universities from Yahoo! Search with the following query in August of 2005:

+site:{university domainA} +linkdomain:{university domainB}

Finally, this set of European universities were grouped by countries (Table 1), then the data obtained were analysed at the aggregate level of countries.

<i>Countries</i>	<i>Universities</i>	<i>%</i>
UK	117	21.87
Germany	110	20.56
France	76	14.21
Italy	50	9.35
Spain	50	9.35
Sweden	28	5.23
Finland	23	4.30
Netherlands	20	3.74
Austria	12	2.24
Greece	12	2.24
Denmark	11	2.06
Portugal	11	2.06
Belgium	8	1.50
Ireland	7	1.31
Total	535	100

Table 1. Number of universities by country

3.2. Geographical maps

Several geographical maps were built to show the spatial distribution of the web data obtained. Mapviewer 6.0 software was selected in order to design these maps. Two main components are necessary to build thematic maps: a base map and a thematic layer. The base map is a reference map that describes the geographical region where data will be projected. The thematic layer is the data projection in accordance with the cartographic coordinates of the base map [20]. As base map, a geographical representation of the 15 European Union member states in 2004 was used. The thematic layer was plotted from a spreadsheet file.

4. Results

We present several geographical maps showing the data extracted and some tables that quantify and explain those maps.

José Luis Ortega and Isidro F. Aguillo

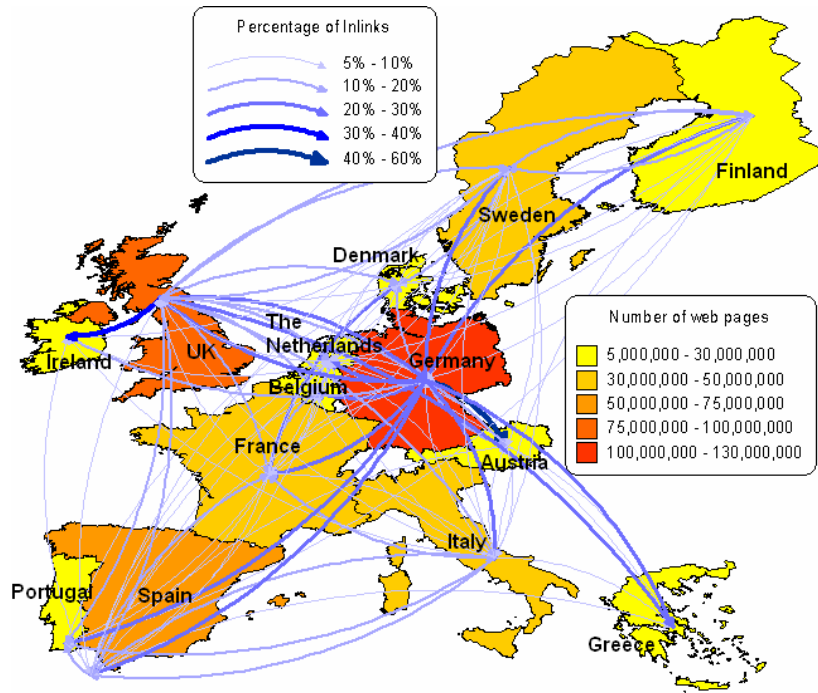


Figure 1. Geographical map of number of pages and inlinks percentage.

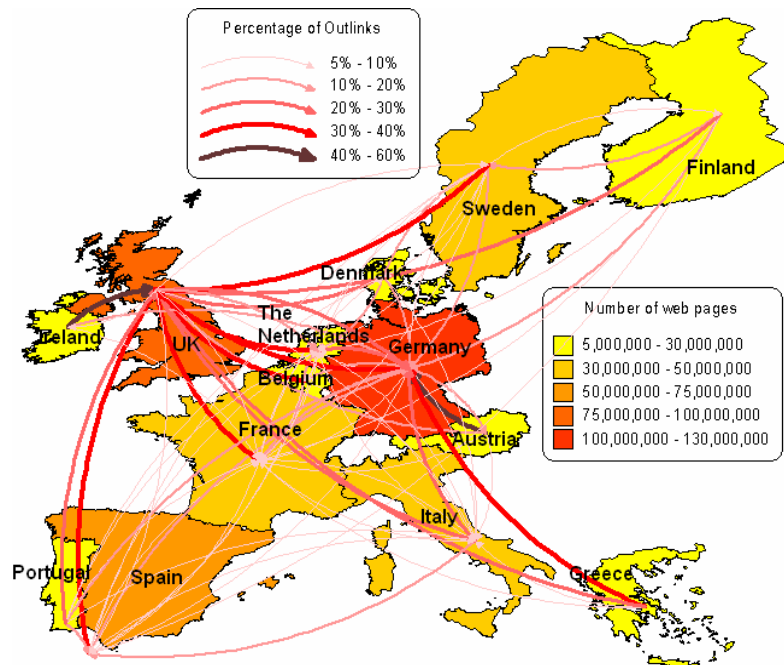


Figure 2. Geographical map of number of pages and outlinks percentage.

José Luis Ortega and Isidro F. Aguillo

Figures 1 and 2 show a map with two layers. The first one displays the number of web pages by each country. The second one presents the link relationships among the 14 EU countries. These link flows are standardized over the total inlinks of a country in Figure 1 and outlinks in Figure 2. The maps show highlighted neighbourhood relationships such as Austria with Germany and Ireland with the United Kingdom (UK), which have the strongest links in the maps. These relationships have a clear linguistic and cultural component. Figure 1 also shows the central position of Germany that has an important percentage of links to other European states, maybe caused by the large size of this web space. However, in Figure 2, UK is the country that attracts most number of outlinks from the rest of European ones. This strong attraction of links may be caused by the use of English language, the use of which attracts more links than other languages in an international environment [21,22].

<i>Countries</i>	<i>Inlinks</i>	<i>% country inlinks</i>	<i>Outlinks</i>	<i>% country outlinks</i>	<i>Pages</i>	<i>%</i>
Austria	47,556	35.24	78,178	47.22	11,925,000	2.52
Belgium	31,666	57.5	48,142	67.29	12,668,000	2.68
Denmark	27,957	49.97	26,114	48.27	11,359,800	2.4
Finland	32,914	25.16	46,628	32.26	22,677,000	4.79
France	65,507	18.94	71,272	20.27	33,080,700	6.99
Germany	169,096	19.86	217,140	24.14	127,821,190	27.01
Greece	14,277	37.05	17,844	42.39	6,448,390	1.36
Ireland	19,368	69.96	14,022	62.77	5,742,000	18.76
Italy	60,785	34.44	73,994	39.01	35,225,400	1.21
Portugal	16,593	47.85	25,774	58.77	9,042,000	5.28
Spain	55,990	24.55	78,147	31.23	51,796,200	10.95
Sweden	66,806	33.74	76,188	36.74	31,618,000	1.91
The Netherlands	64,834	47.22	70,890	49.45	24,997,700	7.44
UK	260,428	35.5	124,540	20.84	88,783,300	6.68
Total	883,386		597,560		473,184,680	100

Table 2. Pages and European links distribution by EU countries.

Table 2 also shows the distribution of web pages and European inlinks and outlinks over the total link by country (notice that the national links are not included). Both Germany and UK are the countries with more web contents, representing 27.01% and 18.76% of the total respectively. These countries hold almost half of the population, being the main countries in the European academic web space. Belgium is the country with the highest European outlinks percentage (67.29%), while Ireland is the country with the highest European inlinks percentage (69.96%). Conversely, France is the country with the lowest outlinks (20.27%) and inlinks (18.94%) percentage.

4.1. Outlinks

Appendix 1 shows the distribution and percentage of outlinks according to the national links and the percentage over the total European outlinks. Both Appendix 1 and Figure 3 show that the percentage between national and European outlinks vary according to the number of pages by each country. Thus, Belgium (E= 67.29%; N= 32.71%), Ireland (E= 62.77%; N= 37.23%) and Portugal (E= 58.77%; N= 41.23%) have a higher

José Luis Ortega and Isidro F. Aguillo

percentage of European outlinks than local ones. On the contrary, France (E= 20.27%; N= 79.73%), UK (E= 20.84%; N= 79.16%) and Germany (E= 24.14%; N= 75.86%) are the countries that display a lower percentage of European outlinks. Regarding the total number of outlinks, Germany (28.25%), UK (18.78%) and France (11.04%) are the countries with more outlinks, while Ireland (0.7%) and Greece (1.32%) are the countries with lower number of outlinks.

These results led us to suggest that the countries with fewer pages, i.e. small web spaces, tend to link more outside of their web spaces than inside, and the countries with a larger web space are more likely to link inside of their own web space than outside. Thus, Figure 3 addresses this suggestion showing the percentage of outlinks over the national links and related with the number of pages.

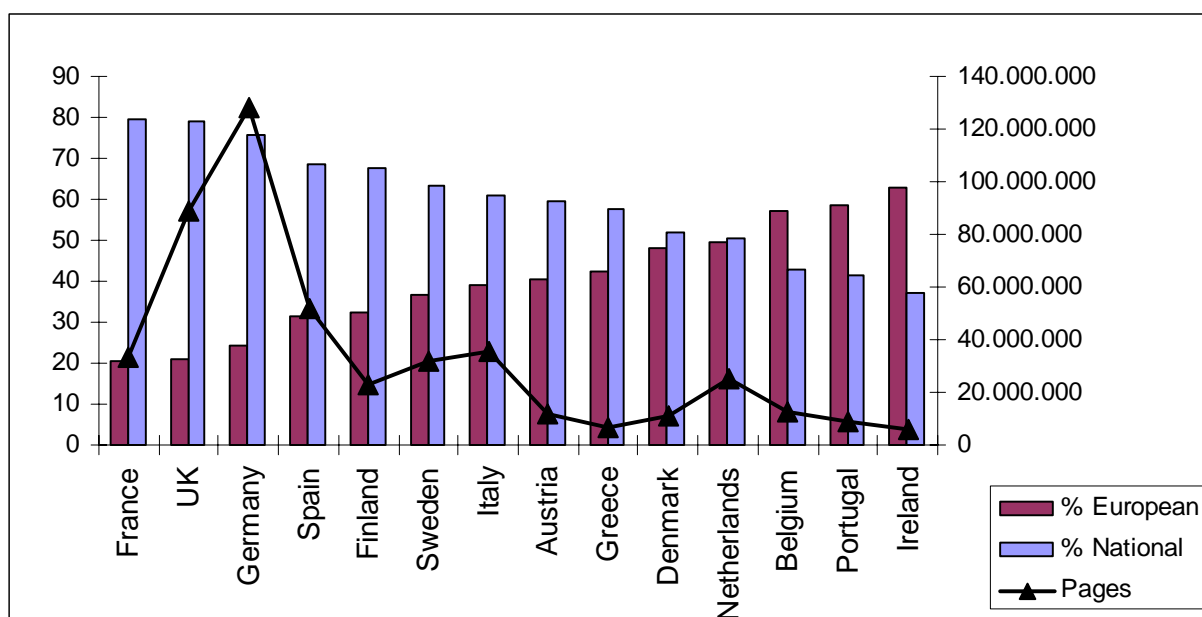


Figure 3. Percentage of European and national outlinks according to the number of pages.

Figure 3 clearly shows that countries with a larger number of pages tend to link more to national pages than European ones. Thus, as the web space of a country is enlarging, the proportion of national links increase and the proportion of European outlinks reduce. The main exception found is France, where the percentage of national links, the highest in the EU, is not proportional to the size of its web space.

4.2. Inlinks

Appendix 1 also shows the distribution and percentage of inlinks according to the national links. Clearly, UK (23.30%) and Germany (27.04%) receive more inlinks. France (81.06%), Germany (80.14%) and Spain (75.45%) have a greater proportion of national links than European ones, while Ireland (30.04%) and Belgium (42.50%) have a lower percentage of national links than European inlinks. These results suggest a double fact to the European outlinks. Thus, the small countries receive more links from Europe than from their own country and the countries with large web spaces have a low proportion of European inlinks.

José Luis Ortega and Isidro F. Aguillo

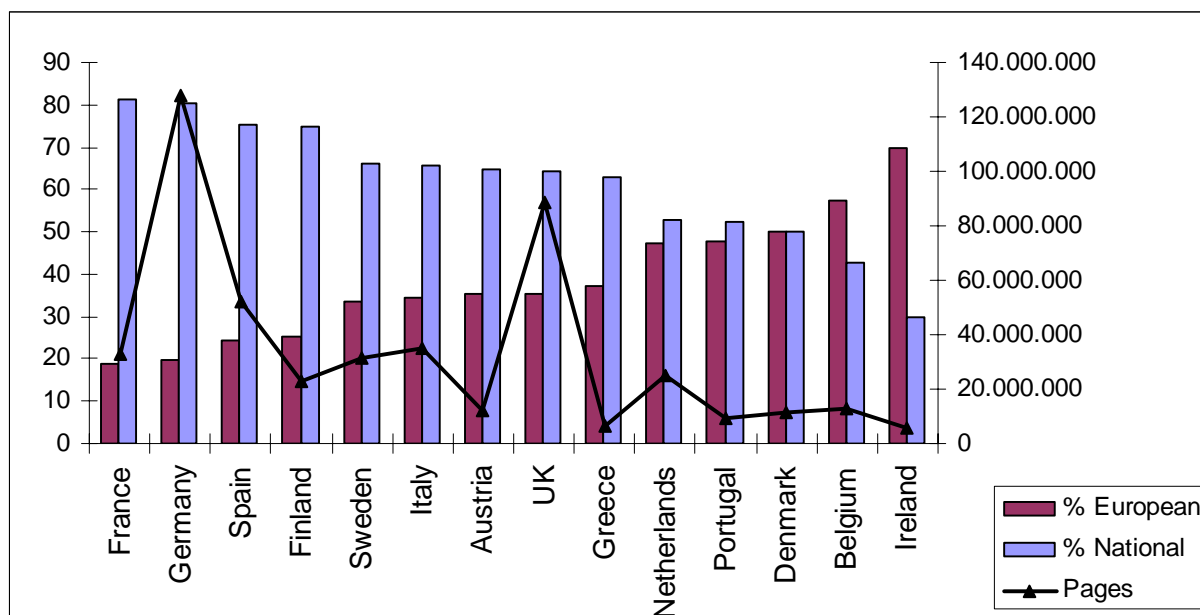


Figure 4. Percentage of European and national inlinks according to the number of pages.

Figure 4 illustrates the relationship between national and European inlinks and number of pages. Thus, the countries with a large size web space have more national links than European inlinks, while the small size countries have more proportion of European inlinks than national ones. However, two exceptions are observed in this distribution. France is the country with more proportion of national links, while its web space is medium-size. The other one, UK, is the second largest country web space; however it has a great percentage of European inlinks. As we have previously pointed out this atypical characteristic may be due to a linguistic aspect.

These results led us to think that the small countries with a limited web space not only establish more links to other external web spaces, but also they receive more links from other countries than their own one. This explains why Ireland (69.96%) and Belgium (57.5%) are the only countries which receive more European links than national ones. This may be caused by the linguistic aspect and because they have a small academic web space with fewer university web sites that link between them.

5. Discussion

An important factor emerges when we work with small populations. Their size and distribution may cause some distortions in the results. Thus countries with a lot of universities may have more pages and more links. Nevertheless, we have noticed that the distribution of web pages is not the same. For example, France is the third country in number of universities but it is fifth in number of pages. And Belgium is the second to last country in universities but the ninth in pages. So, we think that these results may be not affected by inherent characteristics in the population covered.

Especially, we want to discuss the outlier case of France. It presents few web pages according to the number of universities covered in this study and it is the country with less percentage of inlinks and outlinks. It seems that France is not connected with Europe. An extended argument that explains the France low performance is the prior use of a similar technology: Minitel [23, 24]. However, this fact may be also produced by the wide and scattered French dual academic system (Universities and *Écoles*) which does not have big universities such as UK and Germany [25] or by the low use of English language in their web pages [22]. Thelwall et al. [26] also noticed this

José Luis Ortega and Isidro F. Aguillo

low web page count in the French academic web space, using AltaVista. They attribute this 'shortfall' to the large number of web sites, although they cover a relatively small number of web pages.

Geographical maps have shown strong relationships between countries based on cultural and geographical aspects. Several works have detected similar relationships in academic web spaces as European Union [15] or Australasia [27]. So we can consider that the link relationships at the level of countries are determined by geographical and cultural factors, while the link relationships among web domains or sites are more specialized and thematic-oriented [17, 28].

Several linguistic aspects have appeared during this work. We have seen that UK and Ireland are the countries with more proportion of European inlinks according to their size. This can favour the web visibility of English-speaking countries and overestimate the web presence of these countries such as Ireland. Ortega and Aguillo [22] through several regression models found a strong correlation between international links and English language pages. Thelwall, Tang, and Price [21] also found that English pages are more likely to attract links than other languages. In the Asia-Pacific academic web space, Thelwall and Smith [29] also found similar trends, appreciating a pronounced dominance of the English language in the link relationships. Hence, this confirms the English language as an international *lingua franca* on the Web which helps the linking of contents internationally. Nevertheless, the significant weight of this language may be because we are analysing the academic web space where the English language is a well recognized scientific language, although there are languages as German, Sweden or French which have a certain weight in some local regions in Europe.

We have seen that the countries with a small web space tend to link to pages from other countries, because these countries have a low amount of web content and the relevant pages are located in external countries. But the web sites of these countries link to more national pages than foreign ones, as the number of contents increases. This trend follows an increasing rate, being so far as large countries where the proportion of outlinks is clearly lower than national links. Similarly, this also happens with inlinks. In countries with a small web space, the incoming links from other countries are proportionally more than national ones, because there are few pages linked themselves. For this reason, in large countries the incoming links are hardly significant because there are a lot of national nodes that link themselves and the growth rate of national links is higher than the inlinks. These findings were previously noticed by Thelwall and Smith [29] when they studied the web link patterns in the Asia-Pacific academic web space. The disproportionate size of the countries analysed let them to appreciate that large countries as Australia and Japan attracted a great amount of inlinks while they barely linked out of their boundaries. And the smaller countries addressed a huge proportion of their link to other countries. Thelwall and Smith argue that this phenomenon could be caused by a greater number of internal contents in large size countries. However, we think that this phenomenon is caused by the evolving and fractal property of the Web [30], in which the Web is formed by hierarchical sub-networks that evolve from a tiny network with few contents and connected to other networks to a large self-dependent network with their own contents. Similar patterns also detected by Heimeriks and Besselaar [17] and Ortega, et al. [16] allow us to suggest this statement about the country interlinking patterns in our results. However, this study is only based on 14 European countries and the detected link patterns have to be considered cautiously because the sample is not very large. Perhaps, future work with a larger country sample may confirm this assumption.

6. Conclusions

Germany and UK stand out from the remaining countries. The size of their web spaces constitutes more than the half of the European university web. Hence, we can conclude that both countries are the base in which the EU academic web space rests on. Around these countries are located Spain, Italy, France, Sweden, The Netherlands and Finland, middle size countries with a good position on the European academic web space which have an equal proportion of European links and national ones. Finally, the remaining small countries such as Belgium, Greece, Ireland, Portugal, Denmark and Austria which have a small web spaces are highly linked to other countries.

José Luis Ortega and Isidro F. Aguillo

Geographical maps have allowed us to summarize and present easily the results obtained, showing link relationships and frequencies between countries and displaying link distributions. They have also allowed us to detect linguistic and cultural relationships between countries such as Germany and Austria or UK and Ireland, discovering the strong significance of the linguistic and cultural relationships in the Web environment at the level of countries. They have also shown that whereas UK is the most linked country, Germany is the country that makes more links to the rest of European countries. Hence, we can conclude that the geographical maps are a suitable tool in order to better understand web relationships and to detect different roles among countries on the Web.

From this point of view, the use of English language in each country can have influence on its relationship with other countries and its visibility in and outside of the network. So we may consider the number of English language pages as an indicator of the international visibility of an academic web site or web space.

Finally, we conclude that the EU national web spaces behave differently among themselves, and these differences depend on the web size in number of pages by country. Thus, large countries make more links to their own network than the rest of European academic web, and on the contrary the small countries link more to European web sites than to national ones. Hence, the size of a web space, in this case a country web space, determines in a sense the number of links it makes or receives. The great presence of national links allows us to suppose that there is a European academic web space shaped by the joining of several sub-networks according to local criteria such as language or geographical proximity [16].

7. References

- [1] A. Porojan, Trade Flows and Spatial Effects: The Gravity Model Revisited, *Open Economies Review*, 12(3) (2001) 265-280.
- [2] M. Bonitz and A. Scharnhorst, National Science Systems and the Matthew Effect for Countries. In: P. Ohly, G. Rahmstorf and A. Sigel, *Globalisierung und Wissensorganisation* (Ergon Verlag, Würzburg, 2000) <http://www.virtualknowledgestudio.nl/staff/andrea-scharnhorst/documents/national-science-systems-and-the-matthew-effect-for-countries.pdf>
- [3] T. Berners-Lee, R. Cailliau and J. F. Groff, The Worldwide Web, *Computer Networks and Isdn Systems*, 25(4-5) (1992) 454-459.
- [4] T. Berners-Lee, R. Cailliau, J. F. Groff and B. Pollermann, World-Wide Web: the information universe, *Internet Research*, 2(1) (1992) 52-8.
- [5] T. C. Almind and P. Ingwersen, Informetric Analyses on the World Wide Web: Methodological Approaches to 'webometrics', *Journal of Documentation*, 53(4) (1997) 404-426.
- [6] A. G. Smith, ANZAC webometrics: exploring Australasian Web structures, *Proceedings of Information Online and On Disc 99: Strategies for the next millennium*, Sydney, Australia, 1999.
- [7] A. G. Smith, The Impact of Web sites: a comparison between Australasia and Latin America. *Proceedings of INFO'99, Congreso Internacional de Informacion*, Havana, Cuba, 1999.
- [8] A. Noruzi, Web Presence and Impact Factors for Middle-Eastern Countries, *Online Magazine*, 30(2) (2006) 22-28.
- [9] M. Thelwall and D. Wilkinson, Graph structure in three national academic webs: power laws with anomalies, *Journal of the American Society for Information Science and Technology*, 54(8) (2003) 706-712.

José Luis Ortega and Isidro F. Aguillo

- [10] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, Graph Structure in the Web, *Computer Networks-the International Journal of Computer and Telecommunications Networking*, 33(1-6) (2000) 309-320.
- [11] R. Baeza-Yates and C. Castillo, *Características de la web chilena 2004*, Technical Report, Center for Web Research, University of Chile, 2005.
- [12] R. Baeza-Yates, C. Castillo and V. López, *Characteristics of the Web of Spain*, *Cybermetrics*, 9(1) (2005).
- [13] V. Cothey, Some preliminary results from a link-crawl of the European Union Research Area Web. In: P. Ingwersen and B. Larsen, (eds.), *Proceeding of the 10th International Conference of the International Society for Scientometrics and Informetrics*, Stockholm, Sweden, 2005.
- [14] X. Polanco, M. Boudourides, D. Besagni and I. Roche, *Clustering and Mapping European University Web Sites Sample for Displaying Associations and Visualizing Networks*, NTTS&ETK 2001 Conference, Hersonissos (Crete), 2001.
- [15] P. B. Musgrove, R. Binns, T. Page-Kennedy and M. Thelwall, A method for identifying cluster in sets of interlinking Web spaces, *Scientometrics*, 58(3) (2003) 657-672.
- [16] J. L. Ortega, I. F. Aguillo, V. Cothey and A. Scharnhorst, *Maps of the academic web in the European Higher Education Area - an exploration of visual web indicators*, *Scientometrics*, 74(2) (2008).
- [17] G. Heimeriks and P. Van den Besselaar, *Analyzing hyperlinks networks: The meaning of hyperlink based indicators of knowledge production*, *Cybermetrics*, 10(1,1) (2006).
- [18] T. Munzner, E. Hoffman, K. Claffy and B. Fenner, *Visualizing the global topology of the Mbone*, IEEE Symposium on Information Visualization, San Francisco, California, 1996.
- [19] M. A. Zook, *The Web of Production: The Economic Geography of Commercial Internet Content Production in the United States*, *Environment and Planning A*, 32 (2000) 411-426.
- [20] C. Chen, *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*. (Springer-Verlag, London, 2003).
- [21] M. Thelwall, R. Tang and L. Price, *Linguistic Patterns of Academic Web Use in Western Europe*, *Scientometrics*, 56(3) (2003) 417-432.
- [22] J. L. Ortega and I. F. Aguillo, *Linear Analysis of cybermetric data: Quantifying the European university web space*, 11th International Conference of the International Society for Scientometrics and Informetrics, Madrid, 2007
- [23] H. Dauncey, *A cultural battle: French Minitel, the Internet and the Superhighway*, *Convergence: The International Journal of Research into New Media Technologies*, 3(3) (1997) 72-89.
- [24] H. Dumez and A. Jeunemaître, *Regulating and facilitating: European institutions, governments and the global Internet*, *The Global Internet Economy: A United States and France Comparison*, Paris, 2001 http://jonescenter.wharton.upenn.edu/events/Global_Internet_Economy/Dumez_Jeunemaitre.pdf
- [25] I. F. Aguillo, J. L. Ortega and B. Granadino, *Webometric evidence of an academic digital divide in France*, The 9th International Conference on Science and Technology Indicators, Leuven, Belgium, 2006.
- [26] M. Thelwall, R. Binns, G. Harries, T. Page-Kennedy, L. Price and D. Wilkinson, *European Union associated university websites*, *Scientometrics*, 53(1) (2002) 95-111.

José Luis Ortega and Isidro F. Aguillo

- [27] A. Smith and M. Thelwall, Web Impact Factors for Australasian Universities, *Scientometrics*, 54(3) (2002) 363-380.
- [28] L. Vaughan and J. You, Comparing business competition positions based on Web co-link data: The global market vs. the Chinese market, *Scientometrics* 68(3) (2006) 611-628.
- [29] M. Thelwall and A. Smith, Interlinking between Asia-Pacific University Web sites, *Scientometrics*, 55(3) (2002) 363-376.
- [30] G. Caldarelli, R. Marchetti and L. Pietronero, The fractal properties of Internet, *Europhysics Letters*, 52(4) (2000) 386

Appendix 1. European Out, European In and National links distribution by EU countries.



Countries	Outlinks				Inlinks				National Links
	<i>European</i>	<i>% country outlinks</i>	<i>Total</i>	<i>% total outlinks</i>	<i>European</i>	<i>% country outlinks</i>	<i>Total</i>	<i>% total outlinks</i>	
Austria	78,178	47.22	165,572	5.20	47,556	35.24	134,950	4.29	87,394
Belgium	48,142	67.29	71,544	2.25	31,666	57.50	55,068	1.75	23,402
Denmark	26,114	48.27	54,105	1.70	27,957	49.97	55,948	1.78	27,991
Germany	217,140	24.14	899,354	28.25	169,096	19.86	851,310	27.04	682,214
Spain	78,147	31.23	250,231	7.86	55,990	24.55	228,074	7.24	172,084
Finland	46,628	32.26	144,526	4.54	32,914	25.16	130,812	4.16	97,898
France	71,272	20.27	351,551	11.04	65,507	18.94	345,786	10.98	280,279
Greece	17,844	42.39	42,098	1.32	14,277	37.05	38,531	1.22	24,254
Italy	73,994	39.01	189,698	5.96	60,785	34.44	176,489	5.61	115,704
The Netherlands	70,890	49.45	143,370	4.50	64,834	47.22	137,314	4.36	72,480
Portugal	25,774	58.77	43,857	1.38	16,593	47.85	34,676	1.10	18,083
Sweden	76,188	36.74	207,388	6.51	66,806	33.74	198,006	6.29	131,200
UK	124,540	20.84	597,742	18.78	260,428	35.50	733,630	23.30	473,202
Ireland	14,022	62.77	22,338	0.70	19,368	69.96	27,684	0.88	8,316
Total	968,873	30.43	3,183,374	100	933,777	29.66	3,148,278	100	2,214,501