Mapping the Web Relations of Science Centres and Museums from Northern Europe

Fábio Castro Gouveia[1] and Eleonora Kurtenbach[2]

[1] *fgouveia@fiocruz.br*
Museu da Vida – COC – Fundação Oswaldo Cruz, Av. Brasil, 4365, 21040-360 Rio de Janeiro (Brazil)

[2] *kurten@biof.ufrj.br*
Instituto de Biofísica Carlos Chagas Filho - Universidade Federal do Rio de Janeiro, Av. Brigadeiro Trompowisky s/ nº, CCS, Bloco G, CEP 21941-590, Rio de Janeiro (Brazil)

Abstract

Northern Europe is a region characterized by the highest index of science centres and museums visits per year. Their regional association is the *Nordiska Science Center Förbundet* – NSCF – and it is considered a key institution for science popularization in the area. In this study we explore the relations of its associated science centres and museums by performing a Web co-link analysis applied to 37 Websites of NSCF affiliates in order to map their relationships over the Internet. Clustering analysis, multidimensional scaling (MDS) and an analysis of the language versions of each Website were performed. Danish, Norwegian and Swedish, for its linguistic similarities, constitute a community with great mutual intelligibility. In addition, the analyzed data shows strong geographical and/or linguistic influences in the formation of major groups, with external recognition and thematic issues being responsible for secondary cluster formation.
Keywords: Webometrics; Museums and Science centres, Web Co-Link Analysis; Internet.

Introduction

The current technological changes and the need for access to scientific information have strengthened the role of museums and science centres as interfaces between knowledge and society. According to a survey conducted in 2001 by the National Science Foundation (NSF), 30% of Americans received information on science and technology from visits to museums and science centres (Mintz, 2005).

Over the last ten years, museums and science centres have developed new models for popularizing science on the Internet, using it as an electronic folder to disclose the institution or as a virtual space for visitation. The Internet has been gradually understood as an environment where space must not only be created, but established. If the information available is not correctly indexed and structured, it may be almost as inaccessible as those that are not on the network. When an object is presented without contextualization, it loses its strength. The same thing occurs on the Internet: when the information is presented without any reference it tends to be lost.

Local, regional and international networks have been active throughout the process of expansion in the creation of museums and science centres, supporting their development, grouping these institutions and promoting forums, seminars and conferences for discussions and exchange of experiences with regards to their achievements and the challenges they face. One of these associations is *Nordiske Science Center Förbundet* (NSCF), with focus on museums and science centres of Northern Europe.

According to NSCF´s website, the association was created in a meeting that took place in Finland in October 1987. It is a cooperative organization of science centres in the Nordic countries and their official languages are Danish, Norwegian and Swedish. The association aims to assist the establishment of contacts between the science centres in Scandinavia. The association promotes annual meetings for discussion and exchange of ideas among its members. Several of them work actively through the exchange of exhibitions and interactive modules.

Proceedings of ISSI 2009, pages 665–674.
Edited by B. Larsen and J. Leta.

Northern Europe is characterized by some of the highest index of science centres and museums visits per year among European countries according to the European Union surveys (European Commission, 2005). The European Commission produces reports on public opinion surveys since 1973 in all European countries, members or not of the European Union. One of these reports, called *Eurobarometer*, is focused on various issues related to the public understanding of science. In 2005, this results showed that the mean percentage of people that had visited a museum in a year in Europe (16%) is more than two and a half times lower than the results showed for the U.S. in 1999 (42%), and well above the one found in Latin America (2.3%) for the same year. Sweden (36%) and Norway (27%) are among the countries that have a higher percentage of annual visits to science and technology centres and museums, while Finland (16%) and Denmark (16%) are within the average percentage found for the European Union. On the other hand, Estonia (11%) has a low percentage of visitors.

The Internet and webometric studies

It is widely recognized that more than 85% of the Internet traffic is routed directly or indirectly by Websites devoted to provide links from query searches, having each one of them a different methodology for indexing the information (Hu *et al.*, 2001). Brin & Page (1998), the founders of Google, shifted the previous paradigm of information indexation on the Internet by creating a new method, called PageRank, which considers the network of links devoted to a particular page to rank the search results.

In parallel to that change of paradigm, emerges a new area of study within the science information field, which has the link as its central unit of study. Almind & Ingwersen (1997) defined the term Webometrics as the application of informetric methodologies on the Internet. One of the ways of collecting these webometric variables would be by using commercial search engines. According to Gulli & Signorini (2005), there were around 11.5 billion of public indexable pages in 2005, with Google indexing 68.2%, Yahoo / AltaVista, 59.1%, MSN, 49.2% and Ask / Teoma, 43.5% of this total.

Since the beginning of the Internet, many search engines have been trying to address the issue of indexing and structuring information available on the Web. Only a portion of them offers the possibility to perform advanced queries, using Boolean operators. One of the first approaches with Boolean operators for mapping the relationships between Websites on the Internet was used by Larson (1996) to perform an analysis that he called *co-link analysis*. According to Zuccala (2006), co-link studies are emerging now and few have been conducted to this date, resulting in still considerable innovative interpretations. Prime, Bassecoulard & Zitt (2002) pointed the co-link technique as a promising methodology to be applied to academic activities, but stressed the need for caution in interpreting the data. Herrero-Solana & Morales-del-Castillo (2004), while performing a co-link analysis with universities from Germany, the U.S.A. and Russia, found a geopolitical map in which the division of the "iron curtain" was still present on the Internet and considered co-link analysis one of the simplest methods to obtain information related to Websites. Co-link analysis can therefore assess the similarities and social networks of Websites on the Internet.

In this study we explore the relations of the *Nordiska Science Center Förbundet´s* associated science centres and museums Websites by performing a Web co-link analysis in order to map and investigate their relationships over the Internet. Clustering analysis, multidimensional scaling (MDS) and an analysis of the language versions of each Website were also performed.

Methodology

The *Nordiska Science Center Förbundet* (NSCF) is composed by 45 members - all considered as science centres or museums. For data collection, 41 members were selected for having their own Website domain (i.e. their Websites were not a subdirectory under a domain).

The webometric data collection was performed using AltaVista Search engine (http://www.altavista.com), due to its Boolean search capabilities (Björneborn, Jörnerbon & Ingwersen, 2001). Direct queries with two "link" clauses are not allowed in Google and Yahoo!. Altavista allows this kind of query directly from the search window and since it shares the database with Yahoo! it has the second biggest database according to the research done by Gulli & Signorini (2005). Searches were performed in order to obtain the estimative number of URLs with co-links to each pair of Websites, using the query string "link:URL*i* link:URL*j*" (i.e.: link:www.heureka.fi link:www.zoo.dk). To avoid search engine instabilities searches were carried out on a single day, on January, 14th 2009.
A symmetric matrix with the results was then generated. The values on the diagonal were the result of the sum of each column, according to an approach previously used by Gouveia & Kurtenbach (2009), and represent the sum of the links shared between the Website of the institution with each of the other Websites.

Table 1: Museums and science centres of NSCF selected to co-link analysis

Name	Country	URL	Abbr.*
Danfoss Universe	Denmark	www.danfossuniverse.com	DFUN-DK
Elmuseet	Denmark	www.elmus.dk	ELMS-DK
Experimentarium	Denmark	www.experimentarium.dk	EXPM-DK
Ferskvandscentret	Denmark	www.ferskvandscentret.dk	FCTR-DK
Kattegatcentret	Denmark	www.kattegatcentret.dk	KGCT-DK
Økolariet	Denmark	www.okolariet.dk	OKLR-DK
Steno Museet	Denmark	www.stenomuseet.dk	STNM-DK
Zoo København	Denmark	www.zoo.dk	ZOOC-DK
Energia Keskus	Estonia	www.energiakeskus.ee	ENGC-EE
Teaduskeskus AHHAA	Estonia	www.ahhaa.ee	SCAH-EE
Arktikum	Finland	www.arktikum.fi	ARKT-FI
Heureka – Suomalainen Tiedekeskus	Finland	www.heureka.fi	HFSC-FI
Lusto	Finland	www.lusto.fi	LSTO-FI
Tekniikan Museo	Finland	www.tekniikanmuseo.fi	TCHM-FI
Tiedekeskus Tietomaa	Finland	www.tietomaa.fi	TTSC-FI
Bergen Vitensenter	Norway	www.vilvite.no	BGSC-NO
Innlandets Vitensenter	Norway	www.vitensenteret.no	VTST-NO
Jærmuseet	Norway	www.jaermuseet.no	JAEM-NO
Norsk Teknisk Museum	Norway	www.tekniskmuseum.no	NMST-NO
Sciencenter Østfold	Norway	www.sciencecenter.no	SCOF-NO
Vitensenteret i Trondheim	Norway	www.viten.ntnu.no	SCTH-NO
Balthazar	Sweden	www.balthazar.skovde.se	BZMS-SE
Dalénium Upptäckarcentrum	Sweden	www.dalenium.com	DLSC-SE
Dropp In	Sweden	www.droppin.se	DRIN-SE
EXperimentLabbet	Sweden	www.xl.hik.se	XLAB-SE
Framtidsmuseet	Sweden	www.framtidsmuseet.se	FUTM-SE
Innovatum Teknikpark i Trollhättan	Sweden	www.innovatum.se	IVTP-SE
Kreativum	Sweden	www.kreativum.se	KRTV-SE
Molekylverkstan Stenungsund Science Centre	Sweden	www.molekylverkstan.com	MSSC-SE
Naturhistoriska riksmuseet	Sweden	www.nrm.se	SMNH-SE
Navet – Sjuhäradsbygdens science center	Sweden	www.navet.com	NVSC-SE
Technichus	Sweden	www.technichus.se	TECH-SE
Teknik Verksta'n	Sweden	www.teknikverkstan.com	TNKV-SE
Teknikens Hus	Sweden	www.teknikenshus.se	TNKH-SE
Tom Tits Experiment	Sweden	www.tomtit.se	TTEX-SE
Universeum	Sweden	www.universeum.se	UNIV-SE
UPPTECH Science Center	Sweden	www.upptech.se	UTSC-SE

In a first analysis of data collection results, four of the Websites showed lower adherence to the sample studied and were removed from the sample, leaving 37 for the final co-link analysis. The criteria used was a successive removal of Websites that had the sum of co-links

received lower than two times the number of Websites in the sample (*n*), as the NSCF Website linked all these institutions´ Websites, thus representing a minimum of *n* co-links for all of them. The set of 37 Websites is listed in Table 1 by name, country, Website URL and an abbreviation for further references.
The abbreviations were composed by a shortened name of the institution and a suffix indicating their country according to the standard ccTLD. To assist in data interpretation, each of the 37 Websites was visited in order to obtain the languages in which the content was provided. The primary language was determined by what was shown on the first page of the Website. When the choice of language appeared before any Website content, all languages offered were determined as primaries.
The co-link matrix with the results for the 37 museums and science centres´ Websites was imported to the Statistica 7.0 package and clustering analysis was performed. The method of amalgamation used was Ward's method (Ward, 1963) and the distance measures were 1-Pearson r.
A distance matrix was also generated for further multidimensional scaling analysis (MDS) and a two-dimensional scattered plot was created. According to Park & Thelwall (2006), the strength of MDS is its ability to transform a large table of data from the matrix of distances on a chart in which the Websites that have greater interconnection tend to be closer, allowing a general view of the distribution of links among them. The Websites that had a greater number of co-links with each other tend to appear grouped in the chart.

Results and discussion

Figure 1 presents the plot for the clustering analysis performed for the 37 Websites of NSCF. It is noteworthy that NSCF is a regional association that holds different countries with different languages, but with geographical proximity. One of the advantages of using clustering analysis is the ability to reduce the volume of information to be represented in a chart with groups of data sharing similar characteristics.
To assist in data interpretation, all languages available in each of the 37 Websites were identified (Table 2). We found thirteen different languages. For NSCF, there was a high frequency of Websites whose contents were being presented in English (24 Websites). Swedish was the second language most present (21 Websites), a not surprisingly result as 16 of the Websites were from Sweden. Similar patterns were also found by Thelwall, Tang & Price (2003) while examining the languages of academic areas of 16 European countries. They found that Swedish is highly available as another language in Scandinavian Websites and English was the language most widely used as a second language in European Web. In our study German was also highly present with ten Websites with versions in this language. There were also 11 Websites in a single language version. The average was 2.49 languages per Website.

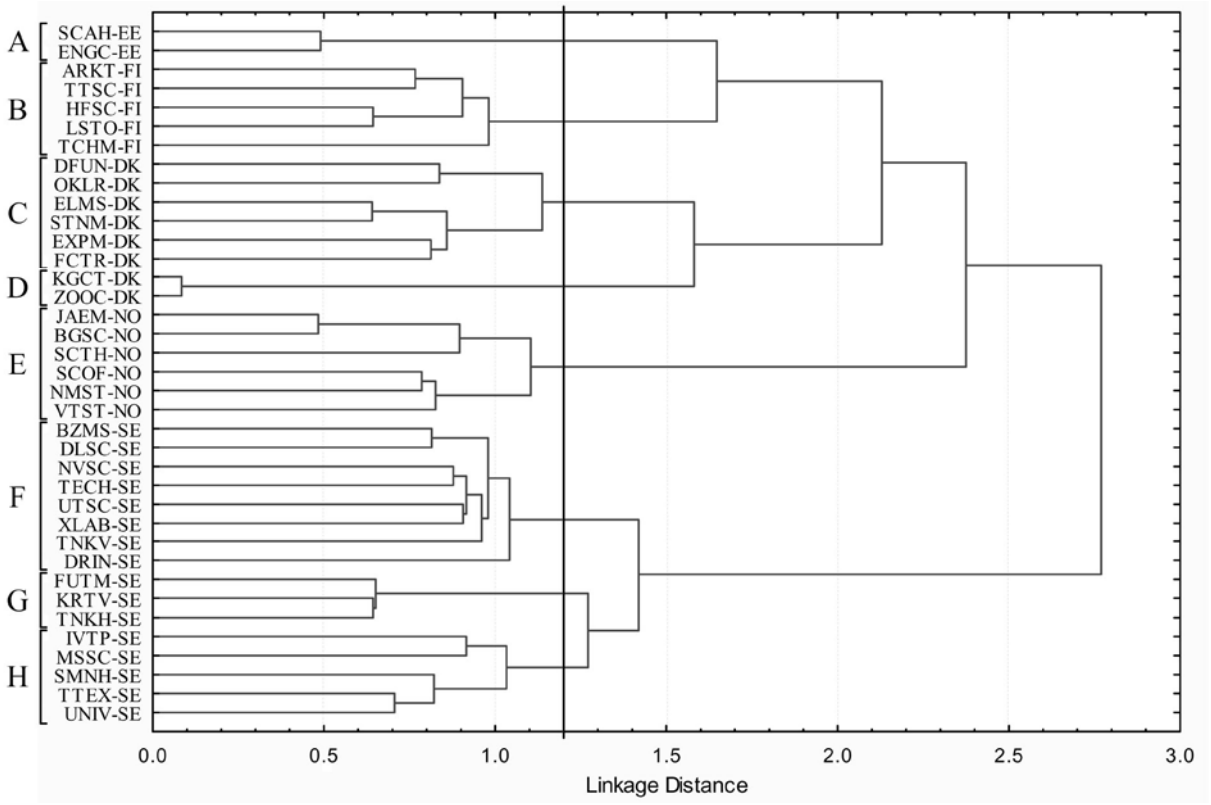

Figure 1: Clustering analysis plot from Web Co-link Analysis of 37 science centres and museums´ Websites from NSCF members.

The cluster division was made at a distance of 1.2, forming eight distinct groups, with a clear geographical/linguistic pattern in group formation. Cluster A was formed by two Estonian institutions. Both were among the less co-linked Websites. The five Websites with the greatest sum of the column were, in descending order, Universeum (UNIV-SE) with 1502, Kreativum (KRVT-SE) with 1076, Framtidsmuseet (FUTM-SE) with 1068, Tom Tits Experiment (TTEX - SE) with 984 and Zoo København (ZOOC-DK) with 912 total co-links. The five Websites with lesser sum were, in ascending order, Teaduskeskus AHHAA (SCAH-EE) with 86, Sciencenter Østfold (SCOF-NO) with 95, Dropp In (DRIN-SE) with 96, Tekniikan Museo (TCHM-FI) with 99 and Energia Keskus (ENGC-EE) with 110 total co-links (data not shown).

There are also some noteworthy linguistic aspects related to this cluster. The two Estonian Websites did not have versions in any other language from countries belonging to NSCF. On the other hand, an Estonian version was only available in these Websites and on the Finnish Website Heureka (HFSC-FI) which has a significant number of co-links with the other two. Five Finnish institutions form cluster B with two of their institutions being the ones with the highest number of language versions. The Websites with the greatest variety of languages were the Arktikum (ARKT-FI), with ten, and Heureka (HFSC-FI), with seven. However, the availability of more language versions apparently did not influence cluster formation, since Finland Websites formed a strong cluster.

Clusters C and D are both formed by Danish institutions as the result of geographical/language influences. The two institutions in cluster C (KGCT-DK and ZOOC-DK) are separated from the ones in cluster D probably due to their strong relation, as they share more than 500 co-links, and because of thematic influences since the first is an Aquarium and the second a Zoo.

Table 2: Languages available in the 37 Websites of museums and science centres studied from NSCF*

	Danish	Estonian	Finnish	Norwegian	Swedish	English	French	German	Italian	Japanese	Polish	Russian	Spanish	Total
DFUN-DK	1	0	0	0	0	1	0	1	0	0	0	0	0	3
ELMS-DK	1	0	0	0	0	1	0	0	0	0	0	0	0	2
EXPM-DK	1	0	0	0	1	1	0	0	0	0	0	0	0	3
FCTR-DK	1	0	0	0	0	1	0	0	0	0	0	0	0	2
KGCT-DK	1	0	0	0	0	1	0	1	0	0	0	0	0	3
OKLR-DK	1	0	0	0	0	1	0	1	0	0	0	0	0	3
STNM-DK	1	0	0	0	0	1	0	1	0	0	0	0	0	3
ZOOC-DK	1	0	0	0	0	1	0	0	0	0	0	0	0	2
ENGC-EE	0	1	0	0	0	0	0	0	0	0	0	0	0	1
SCAH-EE	0	1	0	0	0	1	0	0	0	0	0	1	0	3
ARKT-FI	0	0	1	1	1	1	1	1	1	1	0	1	1	10
HFSC-FI	0	1	1	0	1	1	1	1	0	0	0	1	0	7
LSTO-FI	0	0	1	0	1	1	0	1	0	0	0	1	0	5
TCHM-FI	0	0	1	0	0	0	0	0	0	0	0	0	0	1
TTSC-FI	0	0	1	0	1	1	0	0	0	0	0	0	0	3
BGSC-NO	0	0	0	1	0	1	0	0	0	0	0	0	0	2
VTST-NO	0	0	0	1	0	1	0	0	0	0	0	0	0	2
JAEM-NO	0	0	0	1	0	1	0	0	0	0	0	0	0	2
NMST-NO	0	0	0	1	0	0	0	0	0	0	0	0	0	1
SCOF-NO	0	0	0	1	0	0	0	0	0	0	0	0	0	1
SCTH-NO	0	0	0	1	0	0	0	0	0	0	0	0	0	1
BZMS-SE	0	0	0	0	1	0	0	0	0	0	0	0	0	1
DLSC-SE	0	0	0	0	1	0	0	0	0	0	0	0	0	1
DRIN-SE	0	0	0	0	1	0	0	0	0	0	0	0	0	1
XLAB-SE	0	0	0	0	1	1	0	0	0	0	0	0	0	2
FUTM-SE	0	0	0	0	1	1	0	1	0	0	0	0	0	3
IVTP-SE	0	0	0	0	1	1	0	0	0	0	0	0	0	2
KRTV-SE	0	0	0	0	1	1	0	1	0	0	1	0	0	4
MSSC-SE	0	0	0	0	1	1	0	0	0	0	0	0	0	2
SMNH-SE	1	0	0	0	1	0	0	0	0	0	0	0	0	2
NVSC-SE	0	0	0	0	1	1	0	0	0	0	0	0	0	2
TECH-SE	0	0	0	0	1	0	0	0	0	0	0	0	0	1
TNKV-SE	0	0	0	0	1	0	0	0	0	0	0	0	0	1
TNKH-SE	0	0	0	0	1	1	0	0	0	0	0	0	0	2
TTEX-SE	0	0	0	0	1	1	0	0	0	0	0	0	0	2
UNIV-SE	0	0	0	0	1	1	1	1	0	0	0	0	1	5
UTSC-SE	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Total	9	3	5	7	21	25	3	10	1	1	1	4	2	***2.49***

*The primary languages in which the content is offered are marked in gray. In the lower right, in italics and bold, is the average of languages per Website (2.49).

Cluster E was composed by six Norwegian institutions, also showing the same influences observed before. There are two official forms of written Norwegian, Bokmål ("book language") and Nynorsk ("new Norwegian"), with Norwegians being educated in both. It is noteworthy that five of the Websites were written in Bokmål and only one, Jærmuseet (JAEM-NO), was mainly in Nynorsk with some pages internal pages in Bokmål.

Finally, clusters F, G and H are formed by Swedish institutions. Cluster G is constituted by three institutions that are highly co-linked together forming a strong cluster. They receive several co-links from a single Website related to a national education development project and this is probably the cause of the segmentation between clusters G and H. Cluster H is headed by Universeum (UNIV-SE), the most co-linked Website from this sample, and secondly by Tom Tits Experiment (TTEX-SE), a highly co-linked Website. Several lists of links join the

Websites from cluster H together. Cluster F is formed by Websites that share co-links together but are not closely related to the more highly cited ones. In this sense, it appears that the external recognition and legitimacy of the Swedish museum and science centres´ Websites are playing a central role in cluster formation among, grouping together those that are highly cited and with segmentations as a result of high co-link frequency. As a result, we can observe a map of these Websites with a strong group formation.

Using the Web co-link matrix, a multidimensional scaling (MDS) analysis was performed. Figure 2 shows the two-dimensional MDS analysis plot of the 37 Websites. According to Park & Thelwall (2006), the strength of MDS is its ability to transform a large table of data from the matrix of distances on a chart in which the Websites that have greater interconnection tending to be close in the graphic projection, allowing a general view of the distribution and showing the main links between them. The technique was used to obtain a simple two-dimensional representation of relations between the studied Websites of NSCF. This technique can uncover "hidden structure" in databases resulting in a projection of our data in a two-dimensional plane as a special representation. The stress value was within an acceptable range (approx: 0.161) and the major clusters could be observed clearly.

The cluster letters represented here are the same as identified in the previous cluster analysis. The clusters from the Swedish Websites could not be clearly separated in this visualization. A representation of the Multidimensional Scaling in three dimensions could not improve the visualization of these cluster groups. On the other hand, the separation between the clusters from different countries and the two Danish groups can clearly be observed.

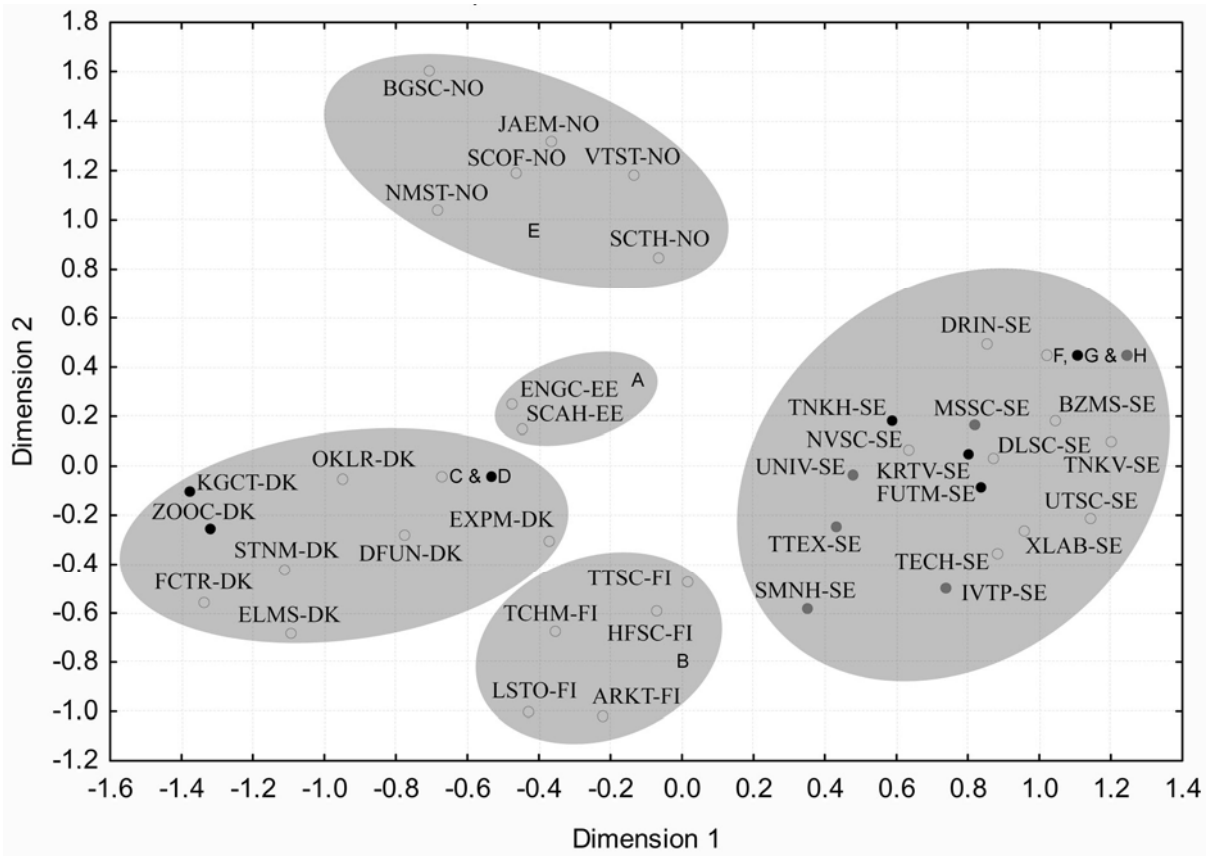

Figure 2: MDS 2D plot using distance matrix from Web co-link analysis of 37 science centres and museums´ Websites from NSCF.

Thelwall (2003) had concerns regarding link motivation as they may have a trivial nature and therefore could pose problems to the implementation of bibliometric techniques in the Web or to the use of links or their counting to infer relationships between individuals or

organizations. On the other hand, Musgrove *et al.* (2003) inferred that there must be certain patterns in the connections between Websites, including those based on political ties, geographical proximity, cultural or linguistic and technological level of recognition. These patterns could reveal how countries differ in terms of their involvement with a larger community, the measure of the forces influencing the underlying facts that govern the involvement of communities on the Web, or the monitoring of the dynamics of social change through studies over time.

The result of clustering analysis for NSCF shows a great geographical and language influence on cluster formation. Similar findings were also brought up by Heimeriks & van den Basselaar (2006) in a study on links between different Websites of European academic institutions. Thelwall, Tang & Price (2003), in an analysis of scholars´ Websites languages from 16 European countries, found a major incidence of links being exchanged between pages of the same language, with English as the dominant one, followed by German, Spanish and Swedish. Also, Thelwall (2002), in a study among universities in the United Kingdom, found that the mutual exchange of links maintained an inverse relation to the distance between the institutions.

The results of Vaughan & You (2006) for a co-link analysis on telecommunication companies suggested that co-links represented the thematic proximity between these companies. In that case, companies with many links in common would have similarities as to the type of product and market. Despite the fact that the Websites in this study are not competitors among themselves, some of them share the same target audience, as in the case of the segmentation of cluster C & D.

Vaughan & Thelwall (2005), while assessing the relationship of the links received between Canadian, English and French universities, noted that cultural issues, along with linguistic influences, were a significant factor for link attraction. In addition, the French-language Canadian universities received a significantly lower number of links when compared to the English-language Canadian universities.

In a previous work, Gouveia & Kurtenbach (2009) found an isolation of Brazilian science centres and museums´ Websites (the only Portuguese speaking country of Latin America) from the other Latin American institutions. Websites in Spanish were found grouped by thematic reasons. In the present study we found that geographical/language issues played a prominent role in cluster formation. All of the clusters were formed by Websites from the same country.

It is important to point out that according to Moberg (2007) the three mainland Scandinavian languages (Danish, Norwegian and Swedish) are so alike that they constitute an interesting linguistic community with respect to mutual intelligibility with its citizens being capable of using their native language when communicating to their neighbours. On the other hand, mutual understanding among these language pairs is asymmetric. In his study it is shown that Swedish is more easily understandable for a Dane than Danish for a Swede and that Norwegians understand their neighbour´s languages better, while Danes and Swedes have more difficulties in understanding Norwegian. As for the written form of these languages, someone with full reading skills in one of them and with contact with the written form of the others is often able to understand the content without riddle. In this sense, the lack of versions on other sibling languages should not be as critical as the findings of Park & Thelwall (2006) with Asian institutions. In their study there was a low level of link exchange between Japan, China and South Korea, countries which do not have a common language, as a result of an inability to understand the native language of the neighbouring country.

As none of the countries studied have English as its main language and considering the proximity of Danish, Swedish and Norwegian languages, Websites versions in English would help more in the aggregation with each other and also with the other Finno-Ugric language

speaking countries (Finland and Estonia) Websites as a result of common understanding and international visibility due to the fact that English is the dominant language on the Internet and for scientific communication.

Conclusions

From the Web Co-link Analysis of the 37 Websites selected from NSCF it is clear that the groups were all formed by institutions from the same country. Considering the proximity of the Scandinavian languages it could be expected a higher thematic influence in cluster formation. This was only observed among Websites from the same country. Also it was observed that external recognition was important as highly co-linked institutions were playing a key role in sub-clusters formation. With content in the same language, science centres and museums were clustered together in terms of their area of activity (cluster D) and differences in link popularity (clusters F, G & H).

One of the common expectations on the Internet is the transposition of physical distance, allowing the exchange of knowledge. It was observed in this study that, though geographically close, museums and science centres from different countries were primarily grouped in a geographical/language pattern.

As previously stated, the Internet is currently an environment in which information should not only be offered, but its value must be established. Without working to strengthen the networks of links between Websites these pages may be kept at a disadvantage in regards to so many other ones of pseudo-science that receive dozens of links because of its popularity. It is up to the Websites of museums and science centres, among other Websites in the field, to strive in the democratization and popularization of scientific knowledge.

Acknowledgments

The authors wish to thank Pamela Lang for the valuable critical reading and revising of the manuscript and the support from CNPq and FAPERJ.

References

Almind, T.C. & Ingwersen, P. (1997) Informetric analyses on the world wide web: methodological approaches to 'Webometrics'. *Journal of Documentation*, 53(4), 404-426.

Björnerbon, L., Jörnerbon, L. & Ingwersen, P. (2001) Perspectives of webmetrics. *Scientometrics*, 50(1), 65-82.

Brin, S. & Page, L. (1998) The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, 30, 107-117.

European Union (2005) *Special Eurobarometer 224 / Wave 63.1 – TNS Opinion & Social - Europeans, Science and Technology*. Retrieved July 10, 2007 from: http://www.mct.gov.br/upd_blob/0013/13512.pdf.

Gouveia, F.C. & Kurtenbach, E. (2009) Mapping the web relations of science centres and museums from Latin America. *Scientometrics*, preprint.

Gulli, A. & Signorini, A. (2005) The Indexable Web is More than 11.5 billion pages. *International World Wide Web Conference,* 902 – 903. Retrieved July 10, 2007 from: http://www.cs.uiowa.edu/~asignori/web-size/size-indexable-web.pdf

Heimeriks, G. & van den Basselaar, P. (2006) Analyzing hyperlinks networks: The meaning of hyperlink based indicators of knowledge production. *Cybermetrics*, 10(1), article 1.

Herrero-Solana, V. & Morales-del-Castillo, J. (2004) "Geopolitical" maps of the Internet: application of new information representation methods. *Ciência da Informação*, 33(3), 69-75.

Hu, W.C.; Chen, Y.; Schmalz, M.S.; Ritter, G.X. (2001) An overview of the World Wide Web search technologies, In: *Proceedings of 5th World Multi-conference on System, Cybernetics and Informatics, SCI2001*, Orlando, Florida, July 22-25.

Larson, R. (1996) Bibliometrics of the World Wide Web: An Exploratory Analysis of the Intellectual Structure of Cyberspace. *Proceedings of ASIS96*, (pp. 71-78). Retrieved July 10, 2007 from: http://sherlock.berkeley.edu/asis96/asis96.html

Musgrove, P. B., Binns, R., Page-Kennedy, T. & Thelwall, M. (2003) A method for identifying clusters in sets of interlinking Web spaces. *Scientometrics*, 58 (3), 657-672.

Mintz, A. (2005) Science, society and science centres. *História, Ciências, Saúde – Manguinhos*, 12 (sup.), 267-280.

Moberg, J., Gooskens, C., Nerbonne, J. & Vaillette, N. (2007) Conditional Entropy Measures Intelligibility among Related Languages. In: Peter Dirix, P., Schuurman, I. Vandeghinste, V. & van Eynde, F. (eds.). *Computational Linguistics in the Netherlands 2006: Selected papers from the 17th CLIN Meeting*. (pp. 51-66) Utrecht: LOT.

Park, H.W. & Thelwall, M. (2006) Web–science communication in the age of globalization. *New Media & Society*, 8(4), 629-650.

Prime, C., Bassecoulard, E. & Zitt, M. (2002) Co-citations and co-sitations: A cautionary view on an analogy. *Scientometrics*, 54(2), p. 291-308.

Thelwall, M. (2002) Evidence for the existence of geographic trends in university Web site interlinking. *Journal of Documentation*, 58 (2), 563-574.

Thelwall, M. (2003) What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, 8(3), article 151. Retrieved July 10, 2007 from: http://informationr.net/ir/8-3/paper151.html

Thelwall, M., Binns, R., Harries, G., Page-Kennedy, T., Price, L. & Wilkinson, D. (2002) European Union associated university websites. *Scientometrics*, 53(1), 95–111.

Thelwall, M., Tang, R. & Price, L. (2003) Linguistic patterns of academic Web use in Western Europe. *Scientometrics*, 56 (3), 417–432.

Vaughan, L. & Thelwall, M. (2005) A modeling approach to uncover hyperlink patterns: the case of Canadian universities. *Information Processing and Management*, 41, 347–359.

Vaughan, L. & You, J. (2006) Comparing business competition positions based on Web co-link data: The global market vs. the Chinese market. *Scientometrics*, 68(3), 611-628.

Ward, J.H. (1963) Hierarchical Grouping to optimize an objective function. *Journal of American Statistical Association*, 58 (301), 236-244.

Zuccala, A. (2006) Author Cocitation Analysis is to Intellectual Structure As Web Co-link Analysis is to… ? *Journal of the American Society for Information Science and Technology*, 57(11), 1487-1502.