

SCIENTIFIC COLLABORATION NETWORK IN “METRICAL STUDIES”: A CO-AUTHORSHIP STUDY USING THE SCIELO INFORMATION SCIENCE PERIODICALS

Ely Francina Tannuri de Oliveira
Universidade Estadual Paulista (UNESP)
Brasil

Maria Claudia Cabrini Grácio
Universidade Estadual Paulista (UNESP)
Brasil

ABSTRACT

This study is aimed at analyzing the scientific collaboration network in the Information Science area, regarding the theme “metrical studies”, based on institutional co-authorships in periodicals published on line by Scientific Electronic Library Online (SciELO), electronic library that comprises a selected collection of Brazilian scientific periodicals, in the following publications: *Ciência da Informação* and *Perspectivas em Ciência da Informação*. The adopted research procedure was the survey of published numbers, including a total of 53 papers related to the topic under study. Initially, 388 papers were worked on, and 53 of them (13,7%, between the two periodicals) are related to this topic. Software *Pajek* was used in order to construct the scientific collaboration network based on co-authorship, and Statistical Package for the Social Sciences (SPSS) was used for cluster analysis, by using Ward’s method and the distance measures were squared Euclidean with the standardized variables. The data were presented as an aggregation of communities, sometimes isolated, other times forming a configuration of a scientific collaboration network already established, but not dense.

Keywords: Metrical Studies; Scientific Collaboration Network; Co-Authorship Analysis; Bibliometry.

INTRODUCTION

In the last decades, studies on metrics in information, mainly Bibliometry and Scientometry, became consistent in the international scope and also in Iberoamérica scopes.

In Brazil, systematic studies on the evolution and trends of sciences were initiated in the 70s and they have been consolidated by researchers from different

areas, mainly from the Information Science area. This field is highlighted due to its important process of institutionalization observed in the last few years, with rich scientific production generated by the high offer of academic instruction in graduation and post-graduation courses, 42 graduation courses and 13 post-graduation courses, as well as by associations that focus on the research and on post-graduation *Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação* (ANCIB), and by a high number of periodicals, including the ones from Iberoamérica in ISI periodicals.

Based on the theory of social networks and on the analysis of Sociology relations, such as the sociogram and sociometry studies – in the mid 30s – as well as on the basis of the mathematical theory of graphs, the study on scientific collaboration networks has been obtaining importance as it gives visibility to the production of science, to the analysis of its domain, and to the more productive scientists, among other objectives.

The research on social networks, using the co-authorship structure, began in Spain and USA with studies intended to form the relations of scientific cooperation in network format, from the individuals, from groups and from institutions located in the same country or internationally.

Como aprenden los científicos? Como colaboran? Es posible establecer la estructura de influencias? [...] Para ello hemos desarrollado un procedimiento que permite, a partir de las publicaciones de un científico, identificar en olas sucesivas dos autores relacionados. El análisis de esta red de co-autorías ha permitido establecer indicadores significativos de la estructura de influencias, a menos como es percibida, por los propios científicos estudiados (MOLINA; MUNÓZ; DOMENECH, 2002)¹.

Therefore, based on the aforementioned presuppositions, to analyze the scientific collaboration networks aiming to verify the labor groups that are formed around a given topic is a vital question in the different knowledge areas, mainly in the Information Science area.

2 OBJECTIVES

¹ Disponível em: <http://revista-redes.rediris.es>

The present study is aimed at analyzing the scientific collaboration network in the information Science area, regarding the topic Metrical Studies, based on co-authorships, in the on line periodicals published by SciELO, in the following publications: *Ciência da Informação* and *Perspectivas em Ciência da Informação*. Therefore, it is intended to study how the scientists collaborate with each other, through the study of institutional co-authorships.

Thus, the specific aim is to provide a view of the scientific collaboration among the researchers who publish papers in these magazines, aiming to identify, to analyze and to describe the situation of the existing scientific collaboration networks, regarding their participation by means of the study on the original institutions and of possible co-authorships in scientific productions.

3 THEORETICAL MARK

The Scientific Electronic Library Online (SciELO) is an electronic library that comprises a selected collection of Brazilian scientific periodicals; it was the result of a research project carried out at The State of São Paulo Research Foundation (FAPESP) in partnership with Latin-American and Caribbean Center on Health Sciences Information (BIREME). From 2002 on, it was also supported by The National Council for Scientific and Technological Development (CNPq).

It is aimed at developing a common methodology for the preparation, storing, dissemination and evaluation of scientific production in electronic format. Nowadays, using this project, there is the publication of 204 periodicals that comprise eight areas of knowledge: Agrarian Sciences, Biological Sciences, Health Sciences, Exact and Earth Sciences, Social and Applied Sciences, Engineering, Linguistics and Arts. In the Social and Applied Sciences area, it is included the Information Science area with publications of two periodicals: *Perspectivas em Ciência da Informação* and *Ciência da Informação*.

The first one, *Perspectivas em Ciência da Informação*, is a quarterly publication of the Information Science School of Federal University of Minas Gerais (UFMG) that is available since 2006.

The second, *Ciência da Informação*, one of the publications with greater visibility in the Information Science area in Brazil, addresses quarterly publications of

unpublished work in this area or work that presents results of studies and researches on the activities of the science and technology information sector. The periodical *Ciência da Informação* has an important performance in the Information Science field and in the information sector in C&T due to the diffusion of significant contributions of national and foreign experts, and also because it reaches the academic community of researchers and professionals from the Information Science area and from correlated areas.

In the area of metrical studies, mainly on the “scientific collaboration” theme, since 1958, Michael Smith, referred to by researchers such as Balancieri et al. (2005), Glanzel and Schubert (2006), was one of the first to observe the growth of co-authorship publication. According to Spinak (1996, p.30),

[...] a co-autoría, también llamada autoría múltiple, se dice de documentos en que dos o más autores que participaron de su creación. Los autores de esos documentos pueden llamar-se *coautores*, pero algunos analistas prefieren reservar esa palabra para documentos en los que colaboraron exactamente dos autores.

Co-authorship has been used, by several researchers, as a measurement of the scientific collaboration among groups of researchers or institutions located in the same country or as international collaboration.

Therefore, in the 1960's, the beginning of studies on the scientific collaboration area was highlighted. During this period, it was observed that the scientific collaboration starts in the relationships between mentor and guiding student, and especially at “invisible schools” scope.

The scientific collaboration among authors and institutions (OLMEDA GÓMEZ; PERIANEZ-RODRIGUEZ; OVALLE-PERANDONES, 2008) presupposes the sharing of central ideas in a project, the goals and the consequences that generated these ideas. It is needed reliability, establishment of work division, and interaction among the researchers, as well as sharing of information and coordination of these different joint investment relationships and final essay. Therefore, the co-authorship reflects the possible list of exchanges among researchers, such as, for instance, the informal talks and discussions at congresses and speeches related to research presentations.

The important benefit resulting from multiple authorships, whether double, triple or *n-les*, when compared to the work of isolated researchers, is the broadening

of approaches and tools that arise from the information exchange and from the increased productivity that occurs when different groups, researchers or institutions join efforts regarding a certain area, promoting the interaction among the researchers.

According to Balancieri et al. (2005, p.2), “the scientific collaboration offers a support source to improve the result and maximize the potential of the scientific production” that is expressed by means of a network in which different collaborators are connected. The concept of social network analysis was developed from a confluence of social theory with mathematical, statistical and computational methodologies.

According to Wasserman and Faust (1994, p.9), “the term ‘social network’ refers to the group of actors and the bonds among them”. Therefore, the network analysis is aimed at modeling the connections among the actors in order to depict, describe and represent the structure of a group, whether formed by countries, institutions or people.

The metrical analyses offer subsidies and instrumentation for the study of social networks since, based on quantitative treatments, it is possible to evaluate some aspects of these relations by means of graphics, densities, proximities, similarities, vectors, intensities, centralities and equalities. Therefore, the link between two points can mean not only the existence of scientific collaboration among authors and scientific institutions, but also the intensity of this collaboration through co-authorship.

Studies carried out by several authors, in different knowledge areas (SPINAK, 1996) have demonstrated that multiple authorships are increasing constantly because the contemporary science shows a tendency to publish multiple authors. These productions cause a greater impact and, sometimes, they have more quality than papers with single authorship, even because sponsorship institutions most frequently support research groups and/or collaborations in multiple authorships.

Other studies, with applied nature, were conducted in several countries, mainly in Spain and they were led by Molina, Munoz and Domenech (2001),

Barragán, Guerrero Bote and Moya Anégon (2006), Liberatore, Herrero-Solana and Guimarães (2007) and Filippo, Casado and Gómez (2007), among others.

The first ones, in a co-authorship Spain/Brazil, carried out a bibliométrica analysis of the Brazilian periodical *Ciência da Informação*, during the period 2000-2004. By means of social networks analysis, they established relations among the periodicals most referred to, and also their grouping by using the technique of clusters.

Filippo, Casado and Gómez (2007), considering the importance of the investigators' mobility as a relevant element of the scientific collaboration, studied the production of co-authorship publications in a given institution: the University Carlos III of Madrid, from 1998 to 2003. The results pointed to the importance of researchers' mobility for the formation of co-authorship networks.

Barragán, Guerrero Bote and Moya Anégon (2006) studied the scientific collaboration from Spain with the countries in Latin America and with the Caribbean Islands, in different thematic domains. The obtained results point to a rare presence of Latin America countries and of Caribbean Islands in collaboration with Spain, constituting a fragile relationship network.

Molina, Munhoz and Domenech (2001) studied the networks of scientific publications by analyzing the structure of co-authorships in three measurement groups: centrality, cohesion and equality and they concluded that the collaboration through papers and communication in congresses is the most common.

Other papers could be mentioned in order to address the literature on this theme, but in general, in recent research the concept of social network and the analysis of these relations are developed as a way of measuring the scientific collaboration among researchers, institutions and countries in order to visualize the research basis in a given area. The scientific collaboration is not a novelty of the twentieth century, but from this period on, its growth was accelerated.

4 METHODOLOGICAL PROCEDURES

The research procedure adopted for identification, analysis and evaluation of data was the survey of on line published issues, in *Perspectivas em Ciência da*

Informação (only three years of on line publication), comprising a total of 68 papers, among which only 11 (16,2%) address the topic “Metrical Studies”. In the *Ciência da Informação* periodical, 11 years of on line publications, 320 papers were analyzed, but only 42 (13,1%) of them addressed the topic. Therefore, 388 papers were initially worked on, and 53 (13,7%, between the two periodicals) papers address the Metrical Studies topic.

Among the 53 articles, 33 of them (62%) presented the theme “metrical studies” applied to different areas of knowledge, as methodological resources or as tools for analyses. The remaining articles present a theoretical-conceptual nature, advancing the knowledge on the area or presenting a conceptual-methodological nature.

Based on the reading of the selected papers, all of them were classified as single and multiple co-authorships, especially in the institutions of origin of these researchers.

For the institutions that produced at least two articles, the following variables were analyzed: total number of articles, number of authorships and number of articles in co-authorship, both intra-group and extra-group (international and national collaborations), that are presented in Table 1.

Software *Pajek* was used to construct the scientific collaboration network based on co-authorships. The density of the constructed network was calculated by the rate among the network connections and the number of possible connections.

The data were analyzed by clusters analysis in order to regroup the different institutions according to their similarity, in relation to co-authorship.

The analysis of groups, also known as analysis of conglomerations, classification or cluster, is aimed at dividing in groups the sample elements, or population, so that the elements belonging to the same group are similar among themselves regarding the measured variables (characteristics), and the elements in different groups are heterogeneous in relation to these same characteristics (MINGOTI, 2007, p.155).

In addition, according the author in question, due to the hierarchy property, resulting from the similarity among data, it is possible to construct the dendogram

[...] that represents the tree or the history of grouping. The Dendogram is a graph in tree format in which the vertical scale indicates the similarity (or dissimilarity) level. In the horizontal axis, the sample elements are marked in an appropriate order related to

the grouping history. The vertical lines, coming from the grouped sample elements, present a height that corresponds to the level in which the elements were considered as similar [...] (MINGOTI, 2007, p.165).

Statistical Package for the Social Sciences (SPSS) was used for cluster analysis, using Ward's method and the squared Euclidean distance measure, with standardized variables.

5 PRESENTATION AND DATA ANALYSIS

Of the 53 analyzed papers, 25 (47%) were presented in single authorship. Therefore, 53% were presented in co-authorships; and among these, 57% were presented in double authorship, indicating that partnership work, in doubles, trebles or more, is present in the area.

The Table 1 shows the number of authorships in institutions that produced at least two papers, single authorship or intra/extra institution co-authorships.

Table 1: Number of Papers and Authorships in Institutions with at Least Two Papers Produced.

UNIVERSITY	NUMBER OF PAPERS WITH AUTHORSHIP					
	Number of papers	Number of authorships	Single	Own co-authorship	Co-Authorship brazilian institution	Co-Authorship foreign institution
USP	7	11	2	4	1	0
UFRGS	5	6	3	1	2	0
UFSC	3	9	0	3	0	0
UNB	3	7	0	2	1	1
UFMG	3	3	2	0	1	0
UEFEIRA SANT	3	3	3	0	0	0
EMBRAPA	2	9	0	2	0	0
NETIC	2	9	0	2	0	0
UFPR	2	5	0	1	1	0
UFSCAR	2	4	0	1	1	0
UNIV. GRANADA	2	4	0	2	0	0
UDESC	2	3	1	1	0	0
UN. MARSEILLE	2	2	0	0	2	0
UFBA	2	2	2	0	0	0
UFES	2	2	1	0	1	0
BIREME	2	2	1	0	1	1
PUC	2	2	0	0	1	1

Source: Own elaboration.

By analyzing Table 1, it is observed that among the 17 institutions that produced at least two articles, only three of them (EMBRAPA, NETIC and BIREME) are not universities, however, they conduct researches. Some of the universities have research groups that focus on the evaluation of production and scientific communication,

[...] very active and being of the utmost importance for the development of studies centered on Bibliometry. Among them, it is highlighted the research groups registered at CNPq: Nucleus of Scientific Production (ECA/USP), Scientia Group (UFSC), Indicators of C&T and Innovation (UFSCar), Communication and Scientific Dissemination (IBICT/UFF), Scientific Communication (UnB), Scientific Communication (UFRGS), Scientific Communication in Human Movement Sciences (UDESC), among others (NORONHA; MARICATO, 2008, p.125).

It is also observed that among the universities, seven of them that is: USP, UFRGS, UFSC, UnB, UFMG, UFBA and PUC offer post-graduation programs in Information Science, and the first four ones are among the institutions that also have research groups, already mentioned by Noronha and Maricato (2008).

The aspects previously emphasized reveal that the production of the approached theme is centered at institutions that present this theme in post-graduation programs and/or research groups, which constitute the formal spaces for the knowledge production; and therefore, they support the scientific production in the area.

Moreover, by analyzing the data on Table 1, University of São Paulo (USP) is highlighted in relation to the number of published papers and also in relation to the number of co-authorships, especially in the production among the group of researchers itself. This institution, considered as one of the greatest in the country, has research groups already consolidated; the post-graduation course, one of the first to be created in the area and the academic tradition of this university are strong indicators for the positioning in the research front. It is in the ranking of the 5 most productive university institutions, namely: USP (the first), University of Campinas (UNICAMP), Federal Universities of Rio Grande do Sul, Rio de Janeiro and Minas Gerais, in this order according to *Jornal Folha de São Paulo, Folha Ciência* (PRODUÇÃO..., 2008).

The following institutions are also important in the production of papers: Federal University of Rio Grande do Sul (UFRGS), already mentioned in the ranking, Federal University of Santa Catarina (UFSC), University of Brasília (UnB), Federal University of Minas Gerais (UFMG) all with co-authorship production, mainly inside the institution itself, and State University of Feira de Santana (UEFS) only with single authorship.

The Brazilian Agricultural Research Corporation (EMBRAPA) and the Research, Knowledge and Innovation (NETIC), in spite of having only two productions each, comprise a high number of inter-group researchers, suggesting relations' networks of native production. It is considered that this configuration comes from the agriculture/technological instrumentation and from the multidisciplinary area that are characteristics of these institutions, and also that the proximity among researchers favors, reinforces and makes natural the formation of research groups inside the institution.

Next, there is the graphical presentation of data, by using *Pajek* software, referring to co-authorships among the institutions, with at least two published papers, except UFBA, UEFS and UFMG: the first ones because of presenting only single authorship and the last one because of presenting co-authorship that does not belong to this group.

In Figure 1, the colored points indicate each institution; its size is related to the number of published papers and green color indicates the institutions that are part of the scientific collaboration network.

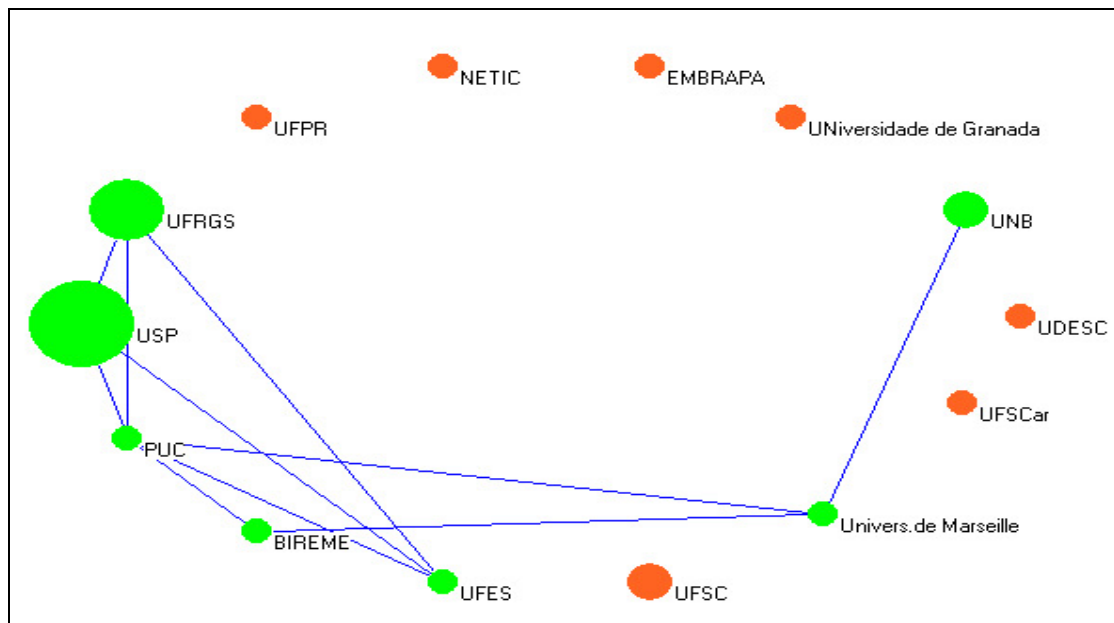


Figure 1: Scientific Collaboration Network among Institutions with at Least Two Published Papers.

Source: Own elaboration.

According to the following chart, it can be observed the configuration of a scientific production network among these institutions: USP, UFRGS, PUC, BIREME, UFES, University of Marseille and UnB. The institutions UFPR, NETIC, EMBRAPA, University of Granada, UFSCar and UFSC only have inter-group authorship. USP, UFRGS, UFPR, UnB, UFSCar, UFES and PUC have partnerships with other institutions, not mentioned here due to the lower productivity on this topic.

Besides, PUC deserves to be highlighted because, in spite of having a small number of total publications with only two articles, it presents a high level of centrality in the network, since it is the institution with the higher number of interinstitutional partnerships.

The density of the network, calculated by the ratio between the number of presented links (19) and the total number of possible links (21) resulted in $\cong 0.21$, representing a 21% possibility of the connections, indicator that suggests a fragile connection among the institutions.

Afterwards, the data were analyzed by *SPSS* statistical software in order to regroup the different institutions according to their similarity.

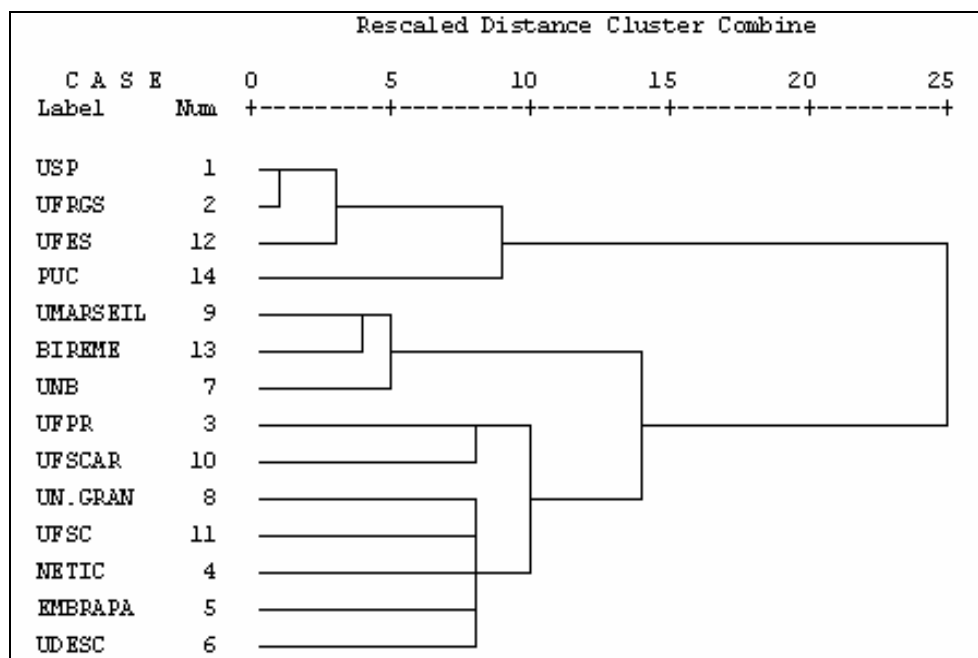


Figure 2: Dendrogram Using Ward Method.

Source: Own elaboration.

By means of the dendrogram analysis, in Figure 2, three groups constituting *clusters*, according to the similarity pattern, can be observed.

The first group is formed by USP, UFRGS, UFES and PUC; institutions that present a high level of inter-institutional collaboration, all of them presenting at least three assignments in collaboration and having the evaluation of scientific production of different areas and the evaluation of periodicals as the most relevant research topics, as analyzed on the basis of the themes and key words of the articles produced by them.

The second group is formed by the University of Marseille, BIREME and UnB, it is characterized by institutions that have little intra and inter-institutional collaboration and that deal mainly with topics related to the evaluation of scientific production at a macro level, in a greater scope and with international bibliographical basis.

The third group, formed by UFPR, UFSCar, University of Granada, UFSC, NETIC, EMBRAPA and UDESC, is characterized, in similarity, by the intragroup production with varied topics related to “metrics” in information, including the use of libraries, assessment of scientific production, information measurement, and social collaboration networks.

FINAL CONSIDERATIONS

The aim of this paper was to identify the scientific collaboration network in the Information Science area regarding the **Metrical Studies** topic, based on the study and analysis of institutional co-authorships, on the on line periodicals published by SciELO, under the following publications: *Ciência da Informação* and *Perspectivas em Ciência da Informação*. The specific purpose was to provide a view on the scientific collaboration among the researchers who publish papers in these periodicals, aiming to identify, to analyze and to describe the situation of existing scientific collaboration networks, regarding the authors' participation.

Data were presented as an aggregation of communities, some times isolated and other times forming the configuration of a scientific collaboration network, already established but not dense. It is hypothesized that this picture reflects the recent character of the advance in metrical studies, more in Brazil than in the international and Iberoamérica scopes.

From the methodological point of view, it can be said that the use of visualization techniques of social networks, using *Pajek* software, is highly recommended in order to represent the configuration of co-authorship networks graphically.

Regarding the use of multivariate statistical analysis by means of clusters, it was possible to group the institutions according to their similarities in the patterns of scientific collaboration, highlighting the most important topic researched by the groups.

The results point to and recommend other possible investigations, particularly the comparison with networks determined by other bases, including the periodicals that are indexed by ISI, as well as the scientific collaboration network formed by analysis and citation, the network determined by co-citations, and the analysis of different tendencies and conceptions in the area.

REFERENCES

BALANCIERI, R. et al. A análise de redes de colaboração científica sob as novas tecnologias da informação e comunicação: um estudo na Plataforma Lattes. **Ciência da Informação**, Brasília, v.34, n.1, p.64-77, 2005.

BARRAGAN, M. J.; GUERRERO-BOTE, V. P.; MOYA-ANÉGON, F. Colaboración científica de España con América Latina y el Caribe. In: ENCUENTRO ASOCIACIÓN DE EDUCADORES E INVESTIGADORES DE BIBLIOTECOLOGÍA, ARCHIVOLOGÍA, CIENCIAS DE LA INFORMACIÓN Y DOCUMENTACIÓN DE IBEROAMÉRICA Y EL CARIBE, 7., 2006. **Anais...** Marília: EDBICIC, 2006. (CD-ROM)

FILIPPO, D. D.; CASADO, E. S.; GÓMEZ, I. Movilidad de investigadores y producción en coautoría para el estudio de la colaboración científica. **Revista CTS**, v.3, n.8, p.23-40, April 2007.

GLANZEL, W.; SCHUBERT, A. Analyzing scientific networks troughs co-authorship. In: MOED, H. F. et al. (Eds.). **Handbook of quantitative science and technology research**. Netherlands: Kluwer Academic Publishers, 2004. 257-276p.

PRODUÇÃO científica cresce 133% em 10 anos no país. **Jornal Folha de São Paulo** (Caderno Ciência), A22, 4 jul. 2008.

LIBERATORE, G.; HERRERO-SOLANA, V.; GUIMARÃES, J. A. C. Análise bibliométrica do periódico brasileiro Ciência da Informação durante o período 2000-2004. **Brazilian Journal of Information Science**, Marília (SP), v.1, n.2, p.3-21, jul./dec. 2007. Available in: <<http://bjis.unesp.br/pt>>. Access: 24 Out. 2008.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte: Editora UFGM, 2007.

MOLINA, J. K.; MUÑOZ, J. M.; DOMENECH, M. Redes de publicaciones científicas: un análisis de la estructura de coautorías. **REDES** - Revista Hispana para el Análisis de Redes Sociales, Barcelona, v.1, 2002. Available in: <http://revista-redes.rediris.es/pdf-vol_1_3pdf>. Access: 24 Out. 2008.

NORONHA, D. P.; MARICATO, J. de M. Estudos métricos da informação: primeiras aproximações. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, n. esp., 1º sem., 2008.

OLMEDA GÓMEZ, C.; PERIANEZ- RODRIGUEZ, A.; OVALLE-PERANDONES, M. A. Estructura de las redes de colaboración científica entre las universidades españolas. In: GARCIA MARCO, F. J. (Ed.). **Avances y perspectivas en sistemas de información y documentación - 2008**. Zaragoza: Prensas Universitarias, 2008. p.129-140

SPINAK, E. **Diccionario enciclopédico de bibliometria, cienciometria e informetria**. Caracas: UNESCO, 1996. 245p.

WASSERMAN, S.; FAUST, K. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.

Ely Francina Tannuri de Oliveira

Universidade Estadual Paulista (UNESP)
Departamento de Ciência da Informação
Av. Hygino Muzzi Filho, 737
Caixa Postal 181
17.525-900 – Marília – SP – Brazil
etannuri@flash.tv.br

Maria Claudia Cabrini Grácio

Universidade Estadual Paulista (UNESP)
Departamento de Psicologia da Educação
Av. Hygino Muzzi Filho, 737
Caixa Postal 181
17.525-900 – Marília – SP – Brazil
cabrini@marilia.unesp.br