

A colaboração científica intra e inter-institucional no Brasil^{1,i}

Samile Andréa de Souza Vanz (PPGCOM/UFRGS)

Ida Regina Chitto Stumpf (PPGCOM/UFRGS)

Resumo: Analisa a colaboração científica brasileira através dos artigos indexados pelo ISI em 2004. Os dados foram analisados com técnicas MDS, Agrupamentos e Fatorial combinadas com ARS. Os resultados mostraram que 29,9% dos artigos são escritos por uma única instituição, enquanto 70,1% por autoria múltipla, tendo como média 2,7 instituições/artigo. A análise das 15 instituições mais produtivas mostrou similaridades e uma separação em sub-grupos. A colaboração intra-institucional é notável em todo o grupo.

Palavras-chave: Colaboração científica. Co-autoria. Ciência brasileira.

Abstract: The Brazilian scientific collaboration was analysed through the papers indexed by ISI in 2004, using MDS, Cluster and Factor Analysis, combined with SNA techniques. The results showed that 29,9% of papers were written by one institution, meanwhile 70,1% by multiple co-authors. The mean is 2,7 institution/paper. Analysis of the 15 most productive institutions showed similarities and a cluster division. The intra-institutional collaboration is remarkable in the whole group.

Keywords: Scientific Collaboration. Co-authorship. Brazilian science.

Introdução

A colaboração entre duas pessoas é um processo social e de interação humana que pode existir de muitas maneiras. A colaboração científica, por sua vez, “[. . .] pode ser definida como o trabalho conjunto de pesquisadores para atingir um objetivo comum de produzir novos conhecimentos científicos.” (KATZ; MARTIN, 1997, p.7). Entre as formas de se obter alguma medida da colaboração entre os cientistas, a mensuração dos artigos publicados em co-autoria vem sendo utilizada como técnica desde a década de 50, quando foram publicados os primeiros relatos do fenômeno do aumento da co-autoria na ciência (SMITH, 1958; PRICE; BEAVER, 1966). Desde então, a co-autoria tem sido usada para investigar a colaboração entre indivíduos, instituições e países, e tem como vantagem o fato de os dados estarem disponíveis em bases de dados que permitem diversos tipos de análises em grandes amostras (KATZ; MARTIN, 1997; NEWMAN, 2001).

Objetivos e Metodologia

Com a intenção de aprofundar o entendimento da ciência brasileira e avançar no conhecimento acerca da formação das redes de colaboração científica, analisou-se a produção científica brasileira indexada no Institute for Scientific Information (ISI) em 2004. Buscaram-

¹ Resumo de pôster apresentado ao GT N°. 7 - Produção e Comunicação da Informação em CT&I.

se artigos científicos que continham ao menos um (1) endereço brasileiro no campo *Address* do referido banco de dados. Os dados foram organizados e analisados através do software Bibexcel². Para verificar a similaridade entre as instituições, foram aplicadas análise MDS (escalonamento multidimensional), Fatorial e de Agrupamentos, realizadas com o software SPSS versão 14. Para enriquecer o entendimento das relações entre elas, os resultados foram combinados com a metodologia de Análise de Redes Sociais, o que possibilitou a visualização das redes de colaboração através do software Pajek³.

Resultados

Foram recuperados 15.597 artigos, reduzidos a 15.350 após a identificação e remoção de 223 registros duplicados e 24 artigos sem especificação de endereço. Foram identificadas 5.464 entradas (instituições) diferentes no campo *Address*. Após processo de normalização dos dados restaram 3.854 diferentes instituições, entre nacionais e estrangeiras. A Tabela 1 apresenta um resumo dos resultados:

Tabela 1 – Resumo dos resultados

N. total de artigos	15.350
N. total de instituições co-autoras	3.854
Instituição por artigo	
Média	2,7
Mediana	2
Moda	2
N. Mínimo	1
N. Máximo	114
Artigos em autoria única (1 instituição)	29,9%
Artigos em autoria múltipla	
2 instituições	33,1%
3 instituições	20,3%
≥ 4 instituições	16,7%

Os artigos em co-autoria perfizeram 70,1%, enquanto 29,9% dos artigos são escritos por um única instituição. A média de instituições vinculadas aos artigos é 2,7 instituições/artigo.

No entanto, é importante salientar que nem sempre as instituições vinculadas são diferentes entre si, muitas vezes os autores pertencem a diferentes departamentos de uma mesma instituição, portanto, a instituição aparece mais de uma vez no campo endereço. A mediana e

² Disponível em <http://www.umu.se/inforsk/Bibexcel/>

³ Disponível em <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

moda são constantes (2) e próximas ao valor apresentado como média, indicando uma distribuição simétrica. O número mínimo de instituições encontrado é um, enquanto o máximo é 114.

Visando entender a colaboração entre as instituições, foram conduzidas algumas análises com as 15 instituições mais produtivas, conforme ranking apresentado na Tabela 2:

Tabela 2 – Ranking das instituições por número de ocorrências

Posição	Total de ocorrências	Instituição
1º	5172	USP
2º	1915	UNICAMP
3º	1852	UFRJ
4º	1420	UNESP
5º	1248	UFRGS
6º	986	UFMG
7º	777	UNIFESP
8º	598	Fiocruz
9º	526	UFSCAR
10º	501	UFSC
11º	492	EMBRAPA
12º	465	UFPR
13º	462	UFPE
14º	411	UERJ
15º	406	UNB

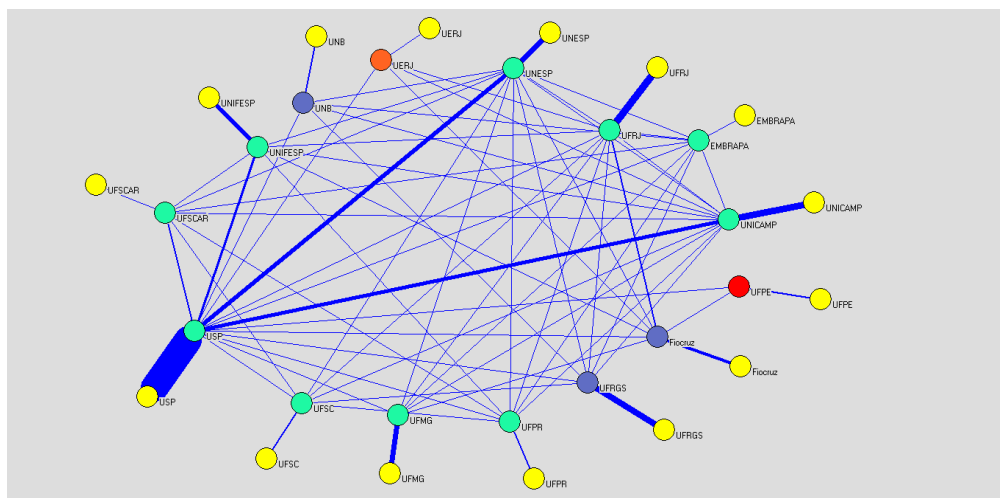
Observa-se que as universidades e instituições de pesquisa de reconhecida tradição lideraram o topo das instituições mais produtivas, sendo marcante a diferença na frequência da USP (5172) e a segunda colocada no ranking, UNICAMP (1915). As universidades federais ocupam posição de destaque. A Fiocruz é a primeira instituição de pesquisa a aparecer, ocupando a oitava posição, com 598 ocorrências, seguida pela Embrapa (11º colocada, com 492 ocorrências).

Quando aplicada análise fatorial às 15 instituições, encontrou-se 95,7% da variância total explicada por 4 fatores, indicando que as variáveis possuem similaridades e estão de fato estreitamente relacionadas umas com as outras. A Figura 1 apresenta o mapa resultante da análise MDS onde são destacados os quatro componentes (componente 1 composto pela USP; componente 2 composto pela UERJ e Fiocruz; componente 3 composto pela UFRJ e UFMG e componente 4 formado pela UFRGS, Embrapa, UFPE, UNESP, UNICAMP, UFSCAR,

Scatter plot showing the distribution of Brazilian universities (labeled) across two dimensions, Dimension 1 (X-axis) and Dimension 2 (Y-axis). The plot is divided into four quadrants by a vertical line at Dimension 1 = 0 and a horizontal line at Dimension 2 = 0. The universities are grouped into four distinct clusters, each enclosed by an ellipse.

- Top-Left Cluster (Quadrant II):** UERJ, Flocruz, UFMG.
- Top-Right Cluster (Quadrant I):** UFPE, UNIFESP, UNICAMP, UFRGS, UFSC, UFPR, UNESP, EMBRAPA, UNB, UFSCAR.
- Bottom-Left Cluster (Quadrant III):** USP.
- Bottom-Right Cluster (Quadrant IV):** UFRJ.

Ao fazer a análise do número total de artigos publicados em co-autoria entre as 15 instituições, é possível observar a predominância da colaboração intra-institucional, ou seja, colaboração dentro de uma mesma instituição. Observam-se percentuais de colaboração intra-institucional em torno de 50%, variando do mínimo de 17,3% apresentado pela UFSCAR ao máximo de 78,2% apresentado pela UFRGS. Entretanto, a análise dos dados não permite afirmar que quanto menos colaboração intra-institucional maior colaboração inter-institucional, como se observa no caso da UFSCAR, que não possui colaboração com três das instituições do grupo no ano estudado. A Figura 2 apresenta a rede de colaboração, gerada no Pajek:



4

A densidade das linhas é proporcional ao número de colaborações e as cores dos nós indicam a divisão das instituições em sub-grupos dentro da rede. Os nós amarelos indicam a colaboração intra-institucional. A fim de facilitar o entendimento da rede de colaboração foram retirados os laços com valores menores que 7 (mediana de colaborações apresentada pelo grupo das 15 instituições), e os laços foram dispostos usando o algoritmo Kamada Kuwai. A colaboração intra-institucional pode significar a colaboração entre diferentes faculdades, departamentos, programas de pós-graduação, que não são objetos de análise deste estudo.

Conclusão

A análise dos dados permite algumas conclusões preliminares. A co-autoria parece ser uma prática que está acontecendo entre a comunidade científica brasileira, visto que 70,1% dos artigos contêm mais de uma instituição autora. Entretanto, as 15 instituições estudadas revelaram uma forte presença de colaboração intra-institucional, sendo esta uma prática predominante entre instituições como a UFRGS, UFMG, UFRJ e USP. Tal fato pode indicar razões regionais e merece ser aprofundado. Aparentemente, as 15 instituições seguem determinados padrões de colaboração, formando quatro diferentes agrupamentos (um composto pela USP – a instituição mais colaborativa do grupo; outro composto pela UERJ e Fiocruz; outro formado pela UFRJ e UFMG e o último formado pela UFRGS, Embrapa, UFPE, UNESP, UNICAMP, UFSCAR, UNB, UNIFESP E UFSC). Os dados requerem análises detalhadas que serão futuramente publicadas pelas autoras.

Referências

- KATZ, J.S.; MARTIN, B.R. What is research collaboration? **Research Policy**, Amsterdam, n. 26, p. 1-18, 1997.
- SOLLA PRICE, D.J.; BEAVER, D.B. Collaboration in an invisible colleges. **American Psychologist**, v. 21, p. 1011-1018, 1966.
- NEWMAN, M.E.J. Scientific collaboration networks I: network construction and fundamental results. **Physical Review E**, New York, v. 64, n.1, 2001.
- SMITH, Mapheus. The trend toward multiple authorship in Psychology. **American Psychologist**, v.13, p. 596-599, 1958.

ⁱ Estes resultados fazem parte da pesquisa de tese, desenvolvida durante o estágio sanduíche financiado pela CAPES