

Uma plataforma para análise e visualização de redes de colaboração científica institucional

Roberth S. Gomes¹, Jhonatan F. Oliveira¹, Thiago M. R. Dias², Gray F. Moita²

¹Departamento de Computação - Universidade do Estado de Minas Gerais – (UEMG)
Av. Paraná 3001 – Jardim Belvedere – 35501-170 – Divinópolis – MG - Brasil

²Departamento de Computação - Centro Federal de Educação Tecnológica de Minas Gerais – (CEFET-MG)

Av. Amazonas 7675 – Nova Gameleira – 30510-000 - Belo Horizonte – MG - Brasil

rsg_roberth@yahoo.com.br, joliveira@divinopolisuemg.com.br,
thiago@div.cefetmg.br, gray@dppg.cefetmg.br

Resumo. *A colaboração científica entre instituições distintas possibilita o intercâmbio de conhecimento e efetiva troca de experiências, impulsionando a evolução da ciência. Analisar as redes de colaboração, pode proporcionar conhecimento sobre como vem ocorrendo as colaboração científica e o quanto estas tem contribuído para a excelência das pesquisas. Este trabalho tem como objetivo analisar as redes de colaborações institucionais a partir dos currículos Lattes. As redes são caracterizadas pelas colaborações entre pesquisadores de diferentes instituições, após caracterizar as redes, técnicas de visualização de grande volume de dados e métricas de análise de redes são aplicadas para verificar como a colaboração científica nacional tem ocorrido.*

Abstract. *The scientific collaboration between different institutions enables the exchange of knowledge and effective exchange of experiences, promoting the development of science. Analyze the collaborative networks can provide knowledge about the scientific collaboration as it has and how these have contributed to the excellence of research. This work aims to analyze the network of institutional collaborations from Lattes curricula. The networks are characterized by collaboration between researchers from different institutions, after characterizing the networks, large data volumes and network analysis metrics visualization techniques are applied to verify as the national scientific collaboration has occurred.*

1. Introdução

O conhecimento atualmente é fator primordial na definição e análise do desenvolvimento de um país e de sua população como um todo. A construção do conhecimento ocorre inevitavelmente através de colaborações, onde pesquisadores se reúnem em prol de objetivos em comum, formando as redes de colaboração científica.

A colaboração científica, objeto de estudos de vários trabalhos vem evoluído ao longo dos anos. Os autores Beaver e Rosen (1978), em seu trabalho, afirmam que o primeiro artigo científico escrito de forma colaborativa entre diferentes pesquisadores, surgiu em 1678.

Com o passar dos anos as redes de colaboração científica foram se consolidando e passaram a ter relação direta com o crescimento do conhecimento científico e de sua comunidade como um todo.

Nudelman e Landers (1972), sugerem que a credibilidade concedida pela comunidade científica a artigos de autoria conjunta é consideravelmente maior do que a creditada a artigos de autoria única, o que posteriormente é comprovado por William Goffman e Kenneth (1980), ao mostrarem que as pesquisas desenvolvidas por grandes grupos tendem a ter mais influência no meio científico.

Portanto, visualizar a colaboração ocorrida na construção do conhecimento passa a ser algo relevante para que se possa mensurar o grau de importância de cada região ao longo deste processo, podendo auxiliar instituições, agências de fomento ou órgãos governamentais no direcionamento mais assertivo de seus investimentos.

A plataforma *Lattes* tem como objetivo representar a experiência do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) na integração de bases de dados de currículos, grupos de pesquisa e de instituições em um único sistema de informações. Cadastrar um currículo na plataforma e mantê-lo atualizado, atualmente, é indicado e está disponível a todos aqueles inseridos no contexto acadêmico e/ou científico; isto se deve em função de sua adoção pela maior parte das instituições de fomento, universidades e institutos de pesquisa do país. A riqueza de informações providas pela plataforma em função de seu constante crescimento, abrangência e confiabilidade, torna-a indispensável e compulsória na análise meritória a pleitos de financiamentos de pesquisas científicas e tecnológicas.

Atualmente não há ferramentas que possibilitem a visualização de mapas de colaboração entre as instituições de pesquisa no cenário nacional. Essa ausência ferramental corrobora para inexistência de trabalhos compostos por embasamento visual que demonstre o grau de colaboração entre instituições de pesquisa no que tange ao conhecimento provido em determinada área do conhecimento.

A plataforma *Lattes* possui atualmente mais de 4.300.000 currículos cadastrados. Estes currículos estão disponíveis para consulta, porém, a consulta do currículo de forma isolada fragmenta o conhecimento disposto no conjunto de currículos, além do fato que a ferramenta de consulta disponibilizada pela própria plataforma *Lattes* não permite a identificação de colaborações científicas, por este motivo, o processamento e análise do grande volume de dados torna-se essencial para transformação destes em conjunto de dados que possam auxiliar na compreensão de como ocorre a colaboração científica institucional no Brasil e nas tomadas de decisões sobre a distribuição de recursos e investimentos.

Diversos trabalhos de redes de colaboração tem explorado a plataforma *Lattes* como principal fonte de dados [DIAS (2013 e 2014), MOREIRA (2014) e SIDONE(2014)].

Sidone (2014) analisa o período de 1990 á 2010 a fim de identificar a evolução das redes de colaboração científica do país e os padrões geográficos do fluxo do conhecimento. Sua análise é quantificada através de tabelas que demonstram as colaborações por cidades/regiões e posteriormente ilustrada com mapas e gráficos demonstrando a intensidade destas colaborações e sua evolução, inclusive por área do conhecimento.

Neste contexto, o objetivo deste trabalho consiste na exploração da plataforma *Lattes* como principal fonte de dados a fim de caracterizar redes de colaboração e aplicar técnicas de visualização de grande volume de dados através da geração de mapas de colaboração institucionais.

2. Metodologia

Para análise e visualização das informações, foi necessário o desenvolvimento de uma plataforma utilizando a linguagem Python. Os principais motivos pela escolha da linguagem, consistem na praticidade / otimização dos códigos e no alto poder de processamento de grandes volumes de dados, ambos obtidos através de bibliotecas nativas da linguagem ou desenvolvidas por terceiros.

A plataforma desenvolvida, permite processar e analisar os dados do repositório *Lattes*, podendo ser parametrizada para trabalhar com áreas específicas do conhecimento como, por exemplo, as nove grandes áreas: Ciências Exatas e da Terra; Ciências da Saúde; Linguística, Letras e Artes; Ciências Biológicas; Engenharias; Ciências Humanas; Ciências Agrárias; Ciências Sociais Aplicadas e Outras. Os dados do repositório *Lattes* foram obtidos através do extrator desenvolvido por DIAS e MOITA (2014). Figura 1.

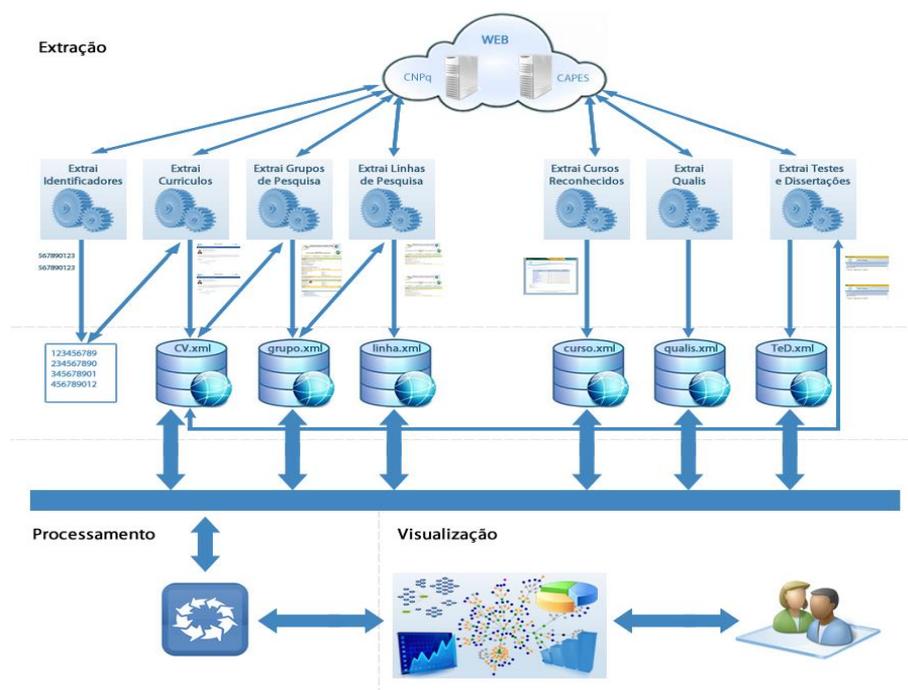


Figura 1 – Arquitetura da Plataforma de Extração e Integração

Na arquitetura desenvolvida pelos autores, eles demonstram que o processo de extração e integração dos dados é dividido em três partes principais denominadas de Extração, Processamento e Visualização. Porém, neste trabalho só foram utilizados os resultados da etapa de extração dos currículos Lattes, tendo em vista, que a identificação das colaborações utiliza-se dos dados das publicações cadastradas em cada um dos artigos, como título e autores.

Com base nos arquivos de pesquisadores e colaborações, ambos obtidos através do extrator desenvolvido por DIAS e MOITA (2015), a análise e manipulação das informações é realizada pela plataforma desenvolvida. As etapas descritas abaixo relatam o funcionamento da plataforma.

Arquivo de instituições: como o foco do trabalho é nas instituições de pesquisa, julga-se necessário uma base que contenha essas informações. Para este fim, um arquivo CSV foi gerado tendo como referência as instituições informadas nos currículos dos pesquisadores, acrescidas de um ID (identificador único), sigla, endereço completo e coordenadas geográficas. As coordenadas geográficas foram obtidas através do endereço completo da instituição.

Seleção dos dados: consiste na definição de um filtro para os dados que deseja trabalhar, por exemplo, uma grande área. Através desta definição, novos arquivos de pesquisadores e colaborações são gerados.

Manipulação dos dados:

- **Arquivo de pesquisadores:** ocorre uma comparação entre o arquivo de pesquisadores e o arquivo de instituições, onde o ID dos pesquisadores na plataforma *Lattes* é substituído pelo ID da instituição a qual estão vinculados. As coordenadas geográficas das respectivas instituições dos pesquisadores, também são acrescidas ao arquivo.
- **Arquivo de memória:** o arquivo é gerado com base na comparação anterior e contém: ID antigo (da plataforma *Lattes*) e o novo ID (da instituição de pesquisa).
- **Arquivo de colaborações:** tendo como referência o arquivo de memória, os ID's do arquivo de colaborações são substituídos pelos ID's das instituições de pesquisa correspondente dos pesquisadores. Posteriormente, há uma separação dos arquivos de colaborações, gerando dois arquivos, sendo: um arquivo contendo as colaborações entre pesquisadores de instituições distintas e outro contendo as colaborações dos pesquisadores das mesmas instituições.

Processamento dos dados: manipulação e tratamento das informações. Nesta etapa ocorre a geração das informações tabulares (*ranking's* e matriz de colaboração) e dos arquivos a serem importados no Gephi, sendo um arquivo de nós (pesquisadores) e um arquivo de arestas não direcionado (colaborações).

Mapa de colaborações / gráfico de representatividade: através do *Gephi*, ocorre a geração do grafo de colaborações entre as instituições de pesquisa e do gráfico correspondente a representatividade de cada instituição.

Os mapas foram gerados, utilizando como parâmetro a instituição atual informada no currículo dos pesquisadores, logo todas as colaborações realizadas (independente do período) foram acrescidas a instituição atual a qual estão vinculados, o que na prática também ocorre. Quando o pesquisador migra de uma instituição, a atual recebe todas as suas produções realizadas nas instituições anteriores.

As colaborações computadas são apenas de coautoria doméstica (entre pesquisadores brasileiros). Colaborações ocorridas com estrangeiros foram consideradas somente se ocorreram simultaneamente com pesquisadores brasileiros.

3. Resultados e discursão

Para representar a análise e processamento dos dados, utilizamos como amostra, os bolsistas que em seus currículos definiram Ciências Exatas e da Terra como grande área; e suas respectivas colaborações no ano de 2014. Os bolsista desta grande área, correspondem a 23% do total de bolsistas do país, sendo a maioria, mestre e doutores. A escolha da grande área se justifica por ser a com o maior número de bolsistas, conforme apresentado na tabela 1.

Tabela 1 – Bolsistas por grande área

Grande Área	Total de bolsistas
Ciências Exatas e da Terra	3296
Ciências Biológicas	2604
Engenharias	1856
Ciências Humanas	1735
Ciências Agrárias	1656
Ciências da Saúde	1595
Ciências Sociais Aplicadas	909
Linguística, Letras e Artes	547
Outros	99

Para melhor representação visual e apresentação tabular dos resultados, apenas as colaborações entre 30 instituições de pesquisa foram consideradas. O critério para definição das instituições é número de bolsistas a elas vinculados. O ranking das 50 instituições está representado na tabela 2.

Tabela 2 – Ranking bolsistas por instituições de pesquisa

Instituição	Quant. de bolsistas
Universidade de São Paulo	2017
Universidade Federal do Rio de Janeiro	902
Universidade Federal de Minas Gerais	695
Universidade Estadual Paulista Julio de Mesquita Filho	695
Universidade Estadual de Campinas	693
Universidade Federal do Rio Grande do Sul	659
Universidade Federal de Santa Catarina	418
Universidade de Brasília	317
Universidade Federal de Pernambuco	313
Universidade Federal do Paraná	290
Universidade Federal Fluminense	276
Universidade do Estado do Rio de Janeiro	269
Universidade Federal de Viçosa	244
Universidade Federal do Ceara	242
Universidade Federal de São Paulo	226
Universidade Federal de São Carlos	215
Fundação Oswaldo Cruz	214
Universidade Federal de Santa Maria	205
Universidade Federal da Bahia	199
Pontifícia Universidade Católica do Rio de Janeiro	188
Universidade Federal da Paraíba	163

Universidade Federal do Rio Grande do Norte	159
Empresa Brasileira de Pesquisa Agropecuária	152
Universidade Estadual de Maringá	148
Universidade Federal de Lavras	138
Universidade Federal de Goiás	132
Universidade Federal do Para	120
Universidade Federal de Uberlândia	115
Pontifícia Universidade Católica do Rio Grande do Sul	115
Universidade Federal de Pelotas	106

O mapa de colaboração, demonstrado na figura 2, representa todo processamento e análise realizada sobre os dados da grande área Ciências Exatas e da Terra. Nele, constam apenas as colaborações interinstituições (entre pesquisadores de instituições distintas). Informações relevantes podem ser extraídas através da sua análise visual, como por exemplo, a maior concentração das publicações nas regiões sudeste do país, intensificadas em sua grande maioria por instituições federais, como: Universidade de São Paulo, Universidade Federal do Rio de Janeiro e Universidade Federal de Minas Gerais; instituições de grande renome no cenário nacional. Em contra partida, temos a imaturidade da rede, tendo em vista a relação direta de dependência de instituições para proliferação do conhecimento.

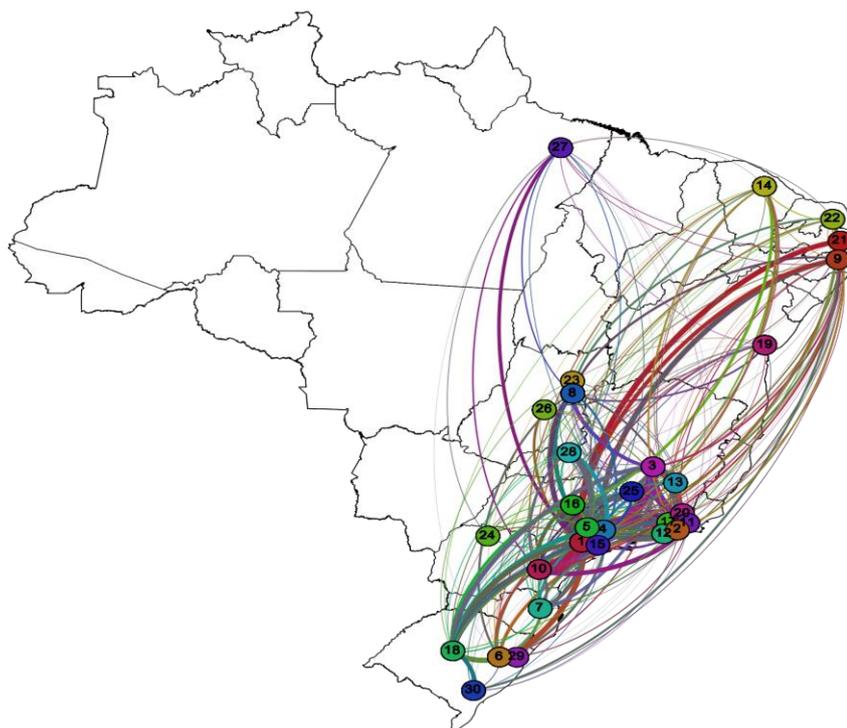


Figura 2 – Mapa de colaboração interinstituições

Tabela 3 – Instituições

Código	Instituição	Sigla
1	Universidade de São Paulo	USP
2	Universidade Federal do Rio de Janeiro	UFRJ
3	Universidade Federal de Minas Gerais	UFMG
4	Universidade Estadual Paulista Julio de Mesquita Filho	UNESP

5	Universidade Estadual de Campinas	UNICAMP
6	Universidade Federal do Rio Grande do Sul	UFRGS
7	Universidade Federal de Santa Catarina	UFSC
8	Universidade de Brasília	UNB
9	Universidade Federal de Pernambuco	UFPE
10	Universidade Federal do Paraná	UFPR
11	Universidade Federal Fluminense	UFF
12	Universidade do Estado do Rio de Janeiro	UERJ
13	Universidade Federal de Viçosa	UFV
14	Universidade Federal do Ceara	UFC
15	Universidade Federal de São Paulo	UNIFESP
16	Universidade Federal de São Carlos	UFSCAR
17	Fundação Oswaldo Cruz	FIOCRUZ
18	Universidade Federal de Santa Maria	UFSM
19	Universidade Federal da Bahia	UFBA
20	Pontifícia Universidade Católica do Rio de Janeiro	PUC-RIO
21	Universidade Federal da Paraíba	UFPB
22	Universidade Federal do Rio Grande do Norte	UFRN
23	Empresa Brasileira de Pesquisa Agropecuária	EMPRAPA
24	Universidade Estadual de Maringá	UEM
25	Universidade Federal de Lavras	UFLA
26	Universidade Federal de Goiás	UFG
27	Universidade Federal do Para	UFPA
28	Universidade Federal de Uberlândia	UFU
29	Pontifícia Universidade Católica do Rio Grande do Sul	PUCRS
30	Universidade Federal de Pelotas	UFPEL

Importante mencionar que a proximidade das instituições de pesquisa em alguns estados, em termos, acabou prejudicando a visibilidade e suas respectivas colaborações no mapa apresentado. Como exemplo, podemos mencionar as instituições de pesquisa do Rio de Janeiro. O estado possui cinco, destas, quatro estão em sua capital. Portanto, ao gerarmos o grafo utilizando as coordenada geográficas, em função da escala do mapa político do Brasil, as instituições ficam uma sobre as outras. Para não comprometer o objetivo do trabalho, omitindo as instituições em função da pouca visibilidade, o grafo foi modificado, onde os nós das mesmas foram reajustados para que pudessem ser visualizadas. Tal ajuste foi realizado respeitando às limitações geográfica dos estados, mas não conseguiu tratar o problema de visualização das colaborações entre estas instituições, onde, por exemplo, a UFRJ e UERJ, possuem 20 colaborações entre si, no mapa tal informação não pode ser visualizada. Situação semelhante também ocorreu no estado de São Paulo.

Em virtude da dificuldade mencionada e para complementar à informação visual apresentada no mapa, uma matriz de colaborações na tabela foi gerada e está sendo representada na figura 3. Através desta matriz é possível verificar com quais instituições de pesquisa houve as colaborações e as respectivas quantidades.

	USP	UFRRJ	UFMG	UNESP	UNICAMP	UFRGS	UFSC	UNB	UFPE	UFPR	UFF	UERJ	UTFV	UFC	UNIFESP	UFSCAR	FIOCRUZ	UFESM	UFBA	PUC-RIO	UFPB	UFERN	EMPRAPA	UEM	UFPA	UFG	UFPA	UFU	PUCRS	UFPEL
USP	488	37	58	126	93	32	26	30	25	92	56	2	16	13	44	84	0	28	7	8	33	14	3	8	10	16	20	38	0	3
UFRRJ	37	161	38	33	26	20	13	13	5	19	63	20	6	11	7	14	7	13	6	23	3	11	0	1	8	7	3	3	0	1
UFMG	58	38	235	51	31	10	19	24	12	33	56	0	19	8	12	29	7	29	3	10	4	3	0	5	12	2	2	9	0	0
UNESP	126	33	51	217	42	16	16	13	32	75	52	0	19	9	42	57	2	34	6	8	1	6	0	7	16	4	2	26	0	1
UNICAMP	93	26	31	42	159	9	9	26	11	31	14	0	4	10	3	11	5	4	6	4	2	1	0	8	14	5	5	3	0	0
UFRGS	32	20	10	16	9	276	19	13	3	17	8	2	2	3	3	5	0	35	7	2	2	1	0	0	3	2	2	3	13	6
UFSC	26	13	19	16	9	19	60	6	2	15	12	0	2	1	3	5	2	9	8	4	8	3	0	0	3	0	0	7	2	0
UNB	30	13	24	13	26	13	6	171	15	1	4	0	1	1	0	0	5	0	12	1	5	14	7	0	0	1	8	5	0	0
UFPE	25	5	12	32	11	3	2	15	115	11	7	0	2	1	3	5	0	6	0	5	5	4	0	0	3	0	1	3	0	0
UFPR	92	19	33	75	31	17	15	1	11	118	31	0	11	8	10	24	2	23	1	6	1	7	0	0	13	0	9	14	0	0
UFF	56	63	56	52	14	8	12	4	7	31	79	2	15	6	12	26	6	29	0	8	0	0	1	0	7	0	2	6	0	3
UERJ	2	20	0	0	0	2	0	0	0	0	2	51	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
UTFV	16	6	19	19	4	2	2	1	2	11	15	0	47	4	4	8	2	8	0	2	0	0	0	0	4	0	0	2	0	0
UFC	13	11	8	9	10	3	1	1	1	8	6	0	4	117	2	18	0	4	0	1	2	7	0	0	1	0	3	2	0	2
UNIFESP	44	7	12	42	3	3	3	0	3	10	12	0	4	2	11	10	0	8	0	2	2	0	0	2	2	0	2	0	0	0
UFSCAR	84	14	29	57	11	5	5	0	5	24	26	0	8	18	10	100	0	19	1	4	8	1	4	6	6	0	0	7	0	5
FIOCRUZ	0	7	7	2	5	0	2	5	0	2	6	0	2	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	2
UFESM	28	13	29	34	4	35	9	0	6	23	29	0	8	4	8	19	0	127	4	4	0	7	1	0	4	0	0	5	0	25
UFBA	7	6	3	6	6	7	8	12	0	1	0	1	0	0	0	1	0	4	51	0	0	0	0	0	0	0	4	2	0	2
PUC-RIO	8	23	10	8	4	2	4	1	5	6	8	0	2	1	2	4	0	4	0	24	0	1	0	0	1	1	0	4	0	0
UFPB	33	3	4	1	2	2	8	5	5	1	0	0	0	2	2	8	0	0	0	0	66	6	0	0	0	2	1	2	0	1
UFERN	14	11	3	6	1	1	3	14	4	7	0	0	0	7	0	1	0	7	0	1	6	44	0	0	2	0	1	2	0	14
EMPRAPA	3	0	0	0	0	0	0	0	7	0	0	1	0	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0
UEM	8	1	5	7	8	0	0	0	0	0	0	0	0	0	2	6	0	0	0	0	0	0	69	0	1	1	1	0	0	0
UFPA	10	8	12	16	14	3	3	0	3	13	7	0	4	1	2	6	0	4	0	1	0	2	0	0	18	0	0	5	0	0
UFG	16	7	2	4	5	2	0	1	0	0	0	0	0	0	2	0	0	0	4	1	2	0	0	1	0	25	0	0	0	0
UFPA	20	3	2	2	5	2	0	8	1	9	2	1	0	3	0	0	0	0	4	0	1	1	0	1	0	0	10	6	0	0
UFU	38	3	9	26	3	3	7	5	3	14	6	0	2	2	2	7	0	5	2	4	2	2	0	1	5	0	6	57	0	0
PUCRS	0	0	0	0	0	13	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0
UFPEL	3	1	0	1	0	6	0	0	0	0	3	0	0	2	0	5	2	25	2	0	1	14	0	0	0	0	0	0	0	59

Figura 3 – Matriz de colaborações

As colaborações das instituições selecionadas, correspondem a um total de 6275. Destas, 3304 (52,65%) são interinstituições e 2971 (47,35%) intra-instituição, ou seja, a diferença apresentada é relativamente pequena, o que demonstra que na prática o intercambio do conhecimento não vem sendo tão explorado.

A tabela 4 detalha como as intuições de pesquisa contribuíram através das suas colaborações. Nela é possível verificar que as instituições de pesquisa UFRGS, UERJ e UEM; possuem menos de 50% das suas publicações com instituições distintas, ou seja, a maior parte do conhecimento provido nas instituições, ocorre entre seus próprios pesquisadores. Em contra partida, temos a EMPRAPA (100%), UNIFESP (94%), UFPA (88%) e FIOCRUZ (82%) que tem grande parte do conhecimento provido realizado de forma conjunta com outras instituições. USP, UNESP e UFMG que estão no topo do ranking de colaborações interinstituições, possuem respectivamente os seguinte percentuais: 65%, 76% e 67%; ou seja, mesmo tendo uma intensidade considerável nas colaborações com outras instituições, ainda possuem um percentual relevante de publicações intra-instituições.

Tabela 4 – Detalhamento das colaborações

Cód.	Instituição	T. de colab.	Intra-Instituição	%	Inter-Instituições	%
1	USP	1410	488	35%	922	65%
2	UFRJ	572	161	28%	411	72%
3	UFMG	721	235	33%	486	67%
4	UNESP	913	217	24%	696	76%
5	UNICAMP	536	159	30%	377	70%
6	UFRGS	514	276	54%	238	46%
7	UFSC	254	60	24%	194	76%
8	UNB	376	171	45%	205	55%
9	UFPE	276	115	42%	161	58%
10	UFPR	572	118	21%	454	79%
11	UFF	505	79	16%	426	84%
12	UERJ	79	51	65%	28	35%
13	UFV	178	47	26%	131	74%
14	UFC	234	117	50%	117	50%
15	UNIFESP	186	11	6%	175	94%
16	UFSCAR	457	100	22%	357	78%
17	FIOCRUZ	49	9	18%	40	82%
18	UFSM	426	127	30%	299	70%
19	UFBA	125	51	41%	74	59%
20	PUC-RIO	123	24	20%	99	80%
21	UFPB	154	66	43%	88	57%
22	UFRN	149	44	30%	105	70%
23	EMPRAPA	16	0	0%	16	100%
24	UEM	109	69	63%	40	37%
25	UFLA	132	18	14%	114	86%
26	UFG	72	25	35%	47	65%
27	UFPA	81	10	12%	71	88%
28	UFU	214	57	27%	157	73%
29	PUCRS	22	7	32%	15	68%
30	UFPEL	124	59	48%	65	52%

4. Conclusão

Através das técnicas aplicadas, foi possível analisar um grande volume de dados e visualizar a representatividade dos principais institutos de pesquisa do país, no processo evolutivo da ciência referente a grande área Ciências Exatas e da Terra.

Conclui-se que embora a comunidade científica esteja cada vez mais receptiva as redes de colaboração, o percentual de publicações entre pesquisadores da mesma instituição ainda é relativamente alto, impedindo o intercambio de informações, dificultando a proliferação do conhecimento e consequentemente a evolução científica da grande área. Há também uma relação direta de dependência de algumas instituições na consolidação da rede, caracterizando-a em termos como imatura.

Toda análise realizada, tem como proposito colaborar para o desenvolvimento de futuros trabalhos que venham analisar de forma qualitativa as informações apresentadas ou coloca-las a prova.

Agradecimentos: Os autores agradecem a FAPEMIG e o CEFET-MG pelo auxílio na pesquisa.

Referências

- BEAVER, Donald B.; ROSEN, Robert . Studies in Scientific Collaboration: part I: the professional origins of scientific co-authorship. **Scientometrics**, v. 1, p. 64-84, 1978.
- DIAS, Thiago M. R.; MOITA, Gray F. A Method for Identification of Collaborations in Large Scientific Databases. **Em questão**. Porto Alegre, v. 21. N. 2, mai./jul. 2015.
- DIAS, Thiago. M. R. ; MOITA, Gray F ; DIAS, Patrícia M.; MOREIRA, Tales H. J. . Identificação e Caracterização de Redes Científicas de Dados Curriculares. iSys: **Revista Brasileira de Sistemas de Informação**, Rio de Janeiro , v. 07, p. 05-18, 2014.
- DIAS, Thiago. M. R. ; MOITA, Gray F.; Dias, Patrícia M.; Moreira, Tales H. J.; FERREIRA, Leandro. R. . Modelagem e Caracterização de Redes Científicas: Um Estudo Sobre a Plataforma Lattes. **Anais do BRASNAM - II Brazilian Workshop on Social Network Analysis and Mining**. Porto Alegre: SBC, v. 01. p. 10-20, 2013.
- GOFFMAN, William; WARREN, Kenneth S. **Scientific information systems and the principle of selectivity**. New York: Praeger, 1980. p. 127.
- MOREIRA, Tales H. J. ; DIAS, Thiago. M. R. ; FERREIRA, Leandro. R. ; MOITA, Gray F. Genealogia Científica: Uma Análise Hierárquica de pesquisadores orientadores. **Anais do XI Simpósio de Mecânica Computacional (SIMMEC)**. Juiz de Fora: UFJF, v. 1. p. 136-147, 2014.
- NUDELMAN, Arthur. E.; LANDERS, Clifford. E. The failure of 100 divided by 3 to equal 33 1/3. **The American Sociologist**, v. 7, n. 9, 1972.