

Uma década de interações entre eventos e pesquisadores do CSBC: Um estudo meta-científico

Luiz Paulo Carvalho¹, Silas Lima Filho¹, Michele A. Brandão²,
Jonice Oliveira¹, Flávia Maria Santoro³, Mônica Ferreira da Silva¹

¹UFRJ – Rio de Janeiro, RJ – Brasil

²IFMG/UFMG – Belo Horizonte, MG – Brasil

³UERJ – Rio de Janeiro, RJ – Brasil

luiz.paulo.carvalho@ppgi.ufrj.br, silaslfilho@ppgi.ufrj.br

michele.brandao@ifmg.edu.br, jonice@dcc.ufrj.br,

flavia@ime.uerj.br, mfsilvamaail@gmail.com

Abstract. *CSBC is the most significant scientific event dedicated to computing in Latin America, integrating Brazilian computing and dealing with the direction of the area in the country. In one of its satellite events, BraSNAM, we find researches about the maturity of the network(s) or community(ies), structuring and combining data that are potential sources of knowledge or feedback for practical actions. This paper analyzes data from the last decade, 2013 – 2022, of publications of the longest-running CSBC events through statistical methods and social network analysis. The results show no patterns between the events but reveal that each community has specific characteristics.*

Resumo. *O CSBC é o maior evento científico dedicado à computação da América Latina, integrando a computação brasileira e tratando dos rumos da área no país. Em um de seus eventos satélites, BraSNAM, encontramos investigações sobre a maturidade da(s) rede(s) ou comunidade(s), estruturando e combinando dados e informações, potenciais fontes de conhecimento ou retroalimentação para ações práticas. Este trabalho apresenta uma análise de dados da última década, 2013 – 2022, de publicações dos eventos mais longevos do CSBC por meio do uso de métodos estatísticos e de análise de redes sociais. Os resultados mostram ausência de padrões entre os diferentes eventos, mas revelam que cada comunidade possui características específicas.*

1. Introdução

Como instituição social, a ciência é composta por diversos elementos e relações [Lakatos e Marconi 2019]. Ela é operada por meio de pessoas [Latour 1987], acadêmicos-cientistas que efetivamente a institucionalizam. A complexidade dos sistemas que compõem a ciência é proporcional às quantidades e qualidades dos elementos, relações e pessoas envolvidas [Lefèvre 2005]. Nesse ensejo, ocorrem movimentações e disposições sociais na intenção de investigar acadêmica-cientificamente as respectivas dinâmicas científicas, seus aspectos sociais, culturais ou históricos, contemplados pela meta-ciência [Ioannidis 2018, Peterson e Panofsky 2020].

Nesse contexto, o CSBC é o maior evento científico dedicado à computação da América Latina, realizado desde 1980, com a finalidade de integrar a comunidade de computação do Brasil e pensar nos futuros rumos da área no país, através de seus eventos e dinâmicas ¹. O CSBC hospeda alguns dos eventos mais longevos e maduros da computação brasileira, e.g., Seminário Integrado de Software e Hardware (SEMISH), em sua 50ª edição em 2023. O CSBC também é responsável por permitir a criação de eventos, que ao amadurecerem, tornam-se independentes.

O Workshop *Brazilian Social Network Analysis and Mining* (BraSNAM) é um evento acadêmico-científico do CSBC especialmente dedicado à mineração e Análise de Redes Sociais (ARS), apresentando diversas comunicações de caráter meta-científico com ênfase analítico descritiva [Marconi e Lakatos 2017, Wazlawick 2014], e.g., [Lobato et al. 2021] apresenta a trajetória de dez anos do BraSNAM, a partir de ARS. Sobre o cenário computacional brasileiro, encontramos outras publicações no BraSNAM [Caparelli e Digiampietri 2018, Boaventura et al. 2014] ou em outros eventos [Digiampietri et al. 2014, Santana e Braga 2020] do CSBC.

Investigações analítico descritivas desenvolvem a maturidade meta-científica de uma comunidade, estruturando e combinando dados e informações, potenciais fontes de conhecimento ou retroalimentação para ações práticas. Contribuem com uma perspectiva sobre a cultura científica [Wang 2018], e.g., [Hosseini et al. 2022] apresenta uma discussão ética relacionada à quantidade de autores em publicações, uma característica cultural intrínseca à prática da comunidade científica. Já [Barabási et al. 2002], a partir de ARS, investiga a evolução das colaborações científicas.

Conforme dados, que englobam aspectos sociais e relacionais, crescem em abundância e complexidade, surge a necessidade de representá-los e comunicá-los utilizando outras abordagens além de listas ou textos corridos. Nesse cenário, as redes sociais apresentam uma possibilidade, convergindo visualização e análise, em diversas dimensões e complexidades [Correa e Ma 2011].

A meta-ciência é um movimento científico social crescente, como prática e área de conhecimento [Peterson e Panofsky 2020]. Enquadramos este trabalho em uma das direções do movimento, de estudos quantitativos sobre a ciência através da ciência de dados [Peterson e Panofsky 2020]. Desviamos do conjunto atual de fenômenos negativos nesta área [Ioannidis 2018], e.g., crise de replicabilidade; apresentando dados e informações descritivos para posterior análise composta ou crítica, aprofundando o contexto do CSBC e informações sobre si.

Assim, como escopo similar de trabalhos predecessores supracitados, questões ou problemas de pesquisa são de interesse secundário; nossa motivação central e primária tem o objetivo de **apresentar uma análise descritiva baseada em ARS e estatística sobre os metadados da última década (2013 – 2022) de publicações dos eventos mais longevos do CSBC, envolvendo dados diversos disponíveis, possíveis e válidos**. Por exemplo, [Lobato et al. 2021] questionam: “Como a comunidade do BraSNAM evoluiu até a sua primeira década de existência?”, o que poderíamos mimetizar; entretanto, o maior valor resultante está na análise-descrição e em suas qualidades, sendo a pergunta de pesquisa facultativa nesta categoria de pesquisas [Marconi e Lakatos 2017], i.e., inde-

¹<https://csbc.sbc.org.br/2023/csbc-2023/> [acesso em 04-04-2023]

pendente da resposta à pergunta, o objetivo gerou conhecimentos avaliados como significativamente positivos. Em resumo, a abordagem adotada neste trabalho é quantitativa descritiva [Wazlawick 2014], com abrangência censitária e sem juízo de valor ou moral. A partir de um procedimento metódico e replicável, aplicamos métodos estatísticos e de ARS com esforço de apresentar o maior e melhor resultado possível, disponível e válido.

O trabalho está estruturado da seguinte forma, a Seção 2 expõe os materiais e métodos, assim como detalhes pontuais sobre o procedimento, e.g., justificativa arrazoada da triagem dos eventos a serem analisados; a Seção 3 apresenta os resultados; e a Seção 4 encaminha as considerações finais.

2. Materiais e métodos

Esta seção detalha os recursos técnicos, os dados e o método, e respectiva instanciamento.

2.1. Sobre o CSBC, os metadados e os eventos escolhidos

Além do BraSNAM, neste trabalho apresentamos uma análise de eventos longevos, maduros e significativos do CSBC, no período de uma década, entre 2013 e 2022. O primeiro desafio foi selecionar os eventos do CSBC para analisar, balanceando completude e integridade. Como a proposta deste trabalho é censitária, analítica descritiva e sem generalização, descartamos uma abordagem amostral, i.e., as informações e conhecimentos dispostos aqui só podem corresponder, com máximo de completude e fidelidade possíveis, ao conjunto analisado. Consideramos os eventos do CSBC com dez edições ou mais e com maior abrangência de dados disponíveis.

Seguimos uma série de critérios e justificativas racionais para inclusão e exclusão de eventos. Foram coletados os metadados de todas as publicações disponíveis entre 2013 – 2022 do SEMISH; BraSNAM; WIT (*Women in Information Technology*); WEI (Workshop sobre Educação em Computação); BreSCI (*Brazilian e-Science Workshop*); WCAMA (Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais); WPerformance (WPerf. – Workshop em Desempenho de Sistemas Computacionais e de Comunicação); SBCUP (Simpósio Brasileiro de Computação Ubíqua e Pervasiva). Outras análises sobre os dados iniciais e maior cobertura de eventos, aqui desconsiderados, e.g., WICS (Workshop sobre as Implicações da Computação na Sociedade); são fortes potenciais de continuidade em trabalhos futuros.

Também foram excluídos os eventos que, em 2023, apresentam 9 edições ou menos, e.g., WICS; os eventos com mais de 10 edições entre 2013 – 2022 que desacoplaram do CSBC, e.g., WIM (Workshop de Informática Médica); eventos que fomos incapazes de coletar integralmente os dados entre 2013 – 2022, e.g., WCGE (Workshop de Computação Aplicada em Governo Eletrônico) em sua décima edição em 2022; eventos com dinâmica epistêmica-prática não tradicional, e.g., JAI (Jornada de Atualização em Informática) e CTD (Concurso de Teses e Dissertações). Trabalhos futuros podem cobrir estes eventos ou outros simpósios brasileiros similares, externos ao CSBC.

Finalmente, coletamos dados disponíveis e possíveis do conjunto de dados: Ano, Evento, Edição, Título, Pessoa autora, Sexo, Instituição 1, UF Inst. 1, Idioma, Resumo, Palavras-chave². A partir destes, analisamos o possível e cabível pelas limitações.

²<https://zenodo.org/record/7977462> [acesso 04-04-2023]

2.2. Recursos, fonte de dados e método

Os seguintes sistemas foram utilizados³: (i) *Web Scraper* e *Octoparse* apresentam versões gratuitas e foram utilizados para extração de dados baseados em estruturas estáticas ou dinâmicas, de forma automatizada e com parametrização manual; (ii) *NetworkX* e *Gephi*, gratuitos e abertos usados para criação, manipulação e estudo de estruturas, dinâmicas e funções de redes complexas; (iii) *Microsoft Excel*, versão comercial, e *Google Sheets*, gratuito, usados para tabelas e planilhas para estruturação, armazenagem e compartilhamento de dados, com funcionalidades aritméticas e estatísticas, representação gráfica de dados e combinações complexas. Utilizamos as funcionalidades gratuitas do sistema online *WordClouds* para geração das nuvens de palavras.

Dois repositórios foram acessados como fonte dos dados durante o mês de fevereiro de 2023, a biblioteca digital SBC SOL⁴ e a coleção de anais do CSBC de 2013, obtida externamente. A qualidade técnica de estruturação, arquitetura e disposição dos metadados da SBC SOL permitiu uma extração facilitada e simples. Foram extraídos manualmente os dados da coleção de anais do CSBC de 2013, cobrindo os indisponíveis na SBC SOL no momento da extração automatizada (WCAMA, WEI e WPerf.).

O método apresenta as seguintes etapas: (i) **extração**; (ii) **tratamento dos dados**; e (iii) **análises**. As etapas (i) e (ii) estão detalhadas em [Lima Filho et al. 2023], aqui nos restringimos às informações pertinentes.

Na etapa de **extração**, para cada evento, foram automaticamente extraídos dados sobre as publicações, como autor, ano de publicação, título, dentre outros. Entretanto, foi necessária a extração manual das instituições de algumas publicações, ausentes na SBC SOL, e.g., publicações do BraSNAM de 2018 e 2017. Também foram observadas inconsistências ou erros em alguns dados, ocorrendo em quantidade quase insignificante (menos de 1% de todos os registros) [Lima Filho et al. 2023].

Relacionado ao **tratamento de dados**, foi realizada a desambiguação das autorias, um desafio tradicional para esta categoria de trabalhos [Digiampietri et al. 2015], feita parcialmente de forma manual para garantia de qualidade dos dados. Para as análises de resumos e palavras-chave, traduzimos quaisquer conteúdos em inglês para português brasileiro. O tratamento de dados se estendeu e resolveu diversos dados de instituições que se encontravam inconsistentes ou absurdos. Por limitações de espaço, realizamos uma quantidade limitada de análises.

Em relação à etapa de **análise**, objetivo primário deste trabalho, foram realizadas análises quantitativas e estatísticas gerais e aplicadas técnicas de ARS. Em ambos os casos, o CSBC foi tratado como um todo e cada evento separadamente. Considerado o excesso de dados e informações gerados, disponibilizamos os mesmos online⁵.

3. Resultados e discussão

Nesta seção apresentamos as análises quantitativas (Seção 3.1) e baseadas em ARS (Seção 3.2) da colaboração entre pesquisadores.

³<https://webscraper.io/> – <https://www.octoparse.com/> – <https://networkx.org/> – <https://gephi.org/> – <https://en.wikipedia.org/wiki/Spreadsheet> – <https://www.wordclouds.com/> [acesso 04-04-2023]

⁴<https://sol.sbc.org.br/> [acesso 04-04-2023]

⁵https://4658.short.gy/CSBC_ARS_2023 [acesso 04-04-2023]

3.1. Análises quantitativas gerais

A Tabela 1 apresenta os quantitativos gerais proporcionais do CSBC e mais oito eventos do CSBC (SEMISH, BraSNAM, WCAMA, WEI, SBCUP, BreSci, WPerf. e WIT). Há uma discrepância entre o quantitativo de publicações e autorias entre os eventos, excluídas mensagens de abertura e similares, e.g., comparativamente, uma quantidade elevada de publicações (428) e autores diferentes (1.204) no WEI. Outras análises culminam em informações comparativas expressivas, a predominância de uma instituição específica não determina predominância da regionalidade dos autores, e.g., WPerf. com a UFPA como instituição mais relevante e maior quantidade de autorias do RJ. Apenas no WIT a participação de mulheres supera a de homens, com a maior discrepância de gênero entre todos os eventos. O SBCUP tem a maior concentração de autores recorrentes. As participações de autores das regiões Sul e Sudeste são maioria no CSBC, o quantitativo de instituições de outras regiões apresentam participação balanceada, principalmente UFPA. Publicações em inglês são raras no CSBC, ocorrendo em sua maioria no BreSci.

Em primeiro momento, as informações básicas da análise quantitativa geral já expõem certos fenômenos. O WEI apresenta superioridade quantitativa, influenciando significativamente as análises gerais do CSBC; o WIT, mesmo realizando a sua primeira chamada de trabalhos em 2016 (7 anos), apresenta maior quantidade de publicações se comparado com outros, e.g., WCAMA (10 anos). O CSBC tem uma proporção de autor por publicação de 38%, o BraSNAM é o único evento acima, com 40%; o SEMISH apresenta o menor valor, com 31%; i.e., são menos autores e mais publicações no primeiro, e o inverso no segundo.

Analisamos as quantidades de autorias por publicações e eventos. Há um comportamento usual da quantidade de autorias por publicação em eventos, a maioria apresenta três autores por publicação, com exceção do BreSci e WPerf., com quatro (28% em ambos). No CSBC como um todo, $\approx 95\%$ publicações apresentam entre dois e oito autores, decrescente conforme este valor aumenta. Apenas no WPerf. há ausência de publicações com um autor apenas. Como valor atípico (*outlier*), uma publicação de 2013 do WCAMA apresenta 21 autores. BreSci acumula a maior quantidade de publicações com mais autores, $\approx 10\%$ apresentam entre oito e doze.

Sobre a quantidade de autores, a maioria expressiva publicou apenas uma vez no CSBC, sem reincidência. $\approx 78\%$ dos autores estão associados a apenas uma publicação; $\approx 12\%$, duas publicações; $\approx 4\%$, três; $\approx 2\%$ em quatro; e as demais quantidades acumuladas somam $\approx 3\%$. Em geral, menos de 1% dos autores estão associados a 10 autorias ou mais. Nesta categoria, Adenauer Yamin é a ocorrência com valor atípico, apresentando co-autoria em 42 publicações, como exposto na Tabela 1. Influenciando significativamente o componente gigante do CSBC, SBCUP e do SEMISH, como exposto na Seção 3.2. Adenauer Yamin e André Aquino são os únicos autores presentes em autorias de todos os anos analisados.

Também sintetizamos as autorias mais prolíficas por evento, e no CSBC, na Tabela 1. Analisamos a concentração de autores com a maior quantidade de autorias pela quantidade de eventos. Com a maior quantidade, André Aquino, com 17 autorias concentradas (SBCUP), i.e., apresenta autoria **apenas** no SBCUP, 17 vezes. Em dois eventos, Gustavo Guedes, com 19 autorias (BraSNAM e BreSci). Em três eventos, Adenauer Yamin (SBCUP, SEMISH e WCAMA); em quatro eventos, Daniel de Oliveira (WPerf.,

Tabela 1. Análises quantitativas gerais dos eventos

	Qtd. pub.	Qtd. autorias	Inst. #1	Inst. #2	UF #1 ***	UF #2 ***	Id. (pt-br)	Id. (en)
SEMISH	179	572	UFPA	UFRJ	RS	PA	139	40
Qtd.	10,3%	12,6%	9,5%	7,4%	15,9%	11,2%	78,0%	22,0%
SBCUP	193	532	UNISINOS	UFC/UFPEl	RS	MG	162	31
Qtd.	11,1%	11,8%	8,0%	7,5%*	30,8%	9,3%	84,0%	16,0%
BraSNAM	240	594	UFMG	UFRJ	MG	RJ	215	51
Qtd.	13,8%	13,1%	19,1%	9,4%	29,6%	19,7%	90%	10%
BreSci	154	462	LNCC	CEFET-RJ	RJ	SP	103	51
Qtd.	8,9%	10,2%	12,8%	8,9%	43,8%	13,1%	67,0%	33,0%
WCAMA	132	446	UFPA	UFRPE	RS	PA	114	18
Qtd.	7,6%	9,9%	8,1%	6,5%	17,6%	13,3%	86%	14%
WPerf.	181	482	UFPA	UFRJ	RJ	SP	148	33
Qtd.	10,4%	10,7%	7,0%	5,9%	12,7%	12,4%	82,0%	18,0%
WEI	428	1204	UFMA	UFPB	PE	RS	401	27
Qtd.	24,6%	26,6%	5,2%	5,0%	10,6%	9,0%	94,0%	6,0%
WIT	232	685	UnB	UFC/UTPR	MG	PR	225	7
Qtd.	13,3%	15,1%	8,0%	5,7%*	8,3%	8,2%	97,0%	3,0%
CSBC	1739	4525	UFRJ	UFPA	RJ	RS	1507	232
Qtd.	100,0%	100,0%	4,5%	4,4%	12,6%	11,1%	86,7%	13,3%
	Part. (M) ***	Part. (F) ***	Autor #1	Autor #2	Autor #3	Autor #4	Autor #5	Autor #6 **
SEMISH	540	152	A. Yamin	M. Scuffo	A. M. Pernas	J. L. Lopes	T. Batista	C. de Barbosa
Qtd.	75,0%	21,0%	15	6	6	6	5	5
SBCUP	631	133	A. Yamin	A. Aquino	C. da Costa	C. Geyer	A. M. Pernas	R. Righi
Qtd.	78,0%	16,0%	24	17	15	14	13	12
BraSNAM	621	219	G. Guedes	L. Digiampietri	A. P. da Silva	J. Oliveira	H. Pereira	F. Benevenuto
Qtd.	71,0%	25,0%	15	14	13	11	11	10
BreSci	455	179	D. de Oliveira	K. Ocaña	E. Ogasawara	F. Porto	M. Mattoso	L. Digiampietri
Qtd.	70,0%	28,0%	18	11	10	8	7	7
WCAMA	395	131	D. G. Gomes	M. de Aguiar	D. F. Adamatti	C. Baptista	G. Gonçalves	V. de Medeiros
Qtd.	68,0%	22,6%	8	6	6	6	5	5
WPerf.	919	197	F. A. Silva	D. Menasché	D. R. Figueiredo	A. B. Vieira	P. Navaux	
Qtd.	75,6%	16,2%	10	9	7	6	6	
WEI	938	534	L. Rivero	C. von Wangenheim	R. Bittencourt	P. Scaico	A. Charão	
Qtd.	59,9%	34,1%	9	9	9	8	8	
WIT	131	826	C. Maciel	A. Araújo	M. Holanda	S. A. Bim	K. Ribeiro	
Qtd.	13,1%	82,8%	14	14	13	12	11	
CSBC	1102	541	A. Yamin	D. de Oliveira	L. Digiampietri	A. M. Pernas	G. Guedes	D. G. Gomes
Qtd.	63,4%	31,1%	42	24	22	20	19	18

* Estas instituições estão empatadas nesta posição, com valores iguais. l ** Quando sexta posição é dividida entre mais de uma pessoa autora (2 ou 3), para evitar um viés implícito sobre a exposição associada ao posicionamento e evitar uma tabela extensa, decidimos omitir os dados neste caso. l *** Nestes campos há ocorrência de dados desconhecidos ou indefinidos.

Qtd. pub.: Quantidade de publicação por evento, seguida da proporção em relação ao CSBC. **Qtd. Autores:** Quantidade de autorias por evento (não autores), seguida da proporção em relação ao CSBC. **Inst. #1 e #2:** Instituições com maior quantitativo de publicações, seguido da proporção no respectivo evento. **UF (Unidade Federativa) #1 e #2:** UFs com maior quantitativo de publicações, seguido da proporção no respectivo evento. **Id.:** Quantitativo de publicações, absoluto seguido de proporcional, em português brasileiro (pt-br) ou inglês (en) no evento. **Part.:** Quantidade de participações, absoluto e proporcional, por sexo dos autores, masculino (M) ou feminino (F). **Autor #1 a #6:** Nome dos autores com maior quantidade de autorias (primárias ou co-autorias) por evento, seguido do valor absoluto de autorias no mesmo.

WCAMA, BreSci e BraSNAM). Fabíola G. Nakamura apresenta a autoria mais diversa, presente em seis dos oito eventos analisados, ausente apenas do WPerf. e WCAMA.

Em relação à primeira autoria, Luciano Digiampietri acumula o maior valor, 10 ocorrências como autor primário. Acompanhado de Karen Ribeiro e Maristela Holanda, com 7; Ronney Castro e Rozelma França, com 5; e a partir de 4 a quantidade aumenta gradativamente. ≈90% dos autores apresentam apenas uma autoria; o somatório daquelas com 4 ou mais, menos de 1%.

Para analisar o perfil epistêmico dos autores com base nos eventos e publicações, foram consideradas três categorias para os eventos: (a) ênfase em aspectos técnicos; (b) mistos entre (a) e (c); e (c) abertura a aspectos não-técnicos. Neste sentido, analisando as o conteúdo das publicações, resumos e palavras-chave, classificamos WPerf., BreSci, SBCUP, WCAMA, SEMISH como (a); WIT e WEI como (c); BraSNAM como (b). Em seguida, geramos uma matriz identidade analisando os eventos por autores de cinco pu-

blições ou mais. De 159 autores analisados, $\approx 30\%$ tem participação mista, publicaram em dois ou três eventos (excluindo a combinação (a) + (c)); $\approx 29\%$ estão concentradas em (c); $\approx 31\%$, concentradas em (a); $\approx 10\%$ entre extremos, publicaram em (a) + (c).

Finalmente, construímos nuvens de palavras para os resumos ⁶, sintetizadas após tratamento de dados. No WIT, WEI, WPerf., BraSNAM e BreSci há uma predominância da sua semântica epistêmica primária, respectivamente, “mulheres”, “ensino”, “desempenho”, “redes” e “científicos”; fenômeno ausente nos demais. Há uma ênfase comum e expressiva no termo “resultados”. WPerf., SBCUP, BreSci e SEMISH apresentam termos muito similares, indicando uma proximidade semântica. Percebemos que o WIT mudou visivelmente a relevância da temática de gênero feminino no CSBC, “mulheres” se tornou um termo visível e aparente na nuvem de palavras do CSBC, mesmo sendo um termo tradicionalmente distante da computação. Um trabalho futuro é analisar estes dados em uma perspectiva temporal.

3.2. Análise da Rede Social do CSBC

Esta seção apresenta uma análise das redes sociais de colaboração do CSBC, envolvendo os oito eventos considerados neste artigo. A Figura 1 expõe o componente gigante dos 10 anos de CSBC. Os nós são os autores e existe aresta entre eles quando publicaram em conjunto, ou seja, a rede social de colaboração formada é de coautoria. As arestas são ponderadas, i.e., quanto mais larga uma aresta, mais publicações em conjunto os autores possuem. Quanto maior o nó do grafo, mais arestas (coautorias) um autor possui. As cores das arestas representam os eventos nos quais as publicações se localizam, e.g., arestas azuis claras representam publicações no BraSNAM. Esta é uma representação que enfatiza as relações (coautorias e eventos).

Para melhor compreender características do grafo apresentado na Figura 1, a Tabela 2 apresenta os valores de métricas (número de nós, número de arestas, densidade, diâmetro, tamanho do componente gigante, tamanho e número máximo de cliques e coeficiente de agrupamento). É possível perceber que a rede de colaboração do CSBC possui 4.505 autores (nós) e 11.397 relacionamentos de coautoria (arestas), mas apenas 39,36% (1.773 nós) desses pesquisadores estão no componente gigante e apenas 48,46% das coautorias entre eles, i.e., possuem coautoria uns com os outros. Por exemplo, a publicação com 21 autores do WCAMA de 2013 está fora desta rede, os autores publicaram apenas uma vez e isolados dos demais.

O WIT e o BraSNAM são os eventos que apresentam a segunda e a terceira maiores redes sociais de coautoria, respectivamente, sendo o WEI a primeira, levando em consideração a quantidade de nós e de arestas. Entretanto, apesar do WIT possuir maior quantidade de coautorias (1.859 arestas) que o BraSNAM (1.333 arestas), o diâmetro do BraSNAM (o valor é 7) é menor que o do WIT (o valor é 8). Ou seja, o BraSNAM possui um grau de separação entre dois autores menor que o WIT, o que pode indicar uma maior interação entre pesquisadores no BraSNAM. Esse resultado pode então ser reforçado pela visualização do componente gigante da rede de coautoria do BraSNAM na Figura 2.

O baixo valor do diâmetro das redes sociais do SEMISH e WCAMA chamam atenção, 2 e 3 respectivamente. Isso indica uma forte coautoria entre os pesquisadores

⁶Limitamos as nuvens de palavras ao repositório complementar, dada a limitação de espaço.

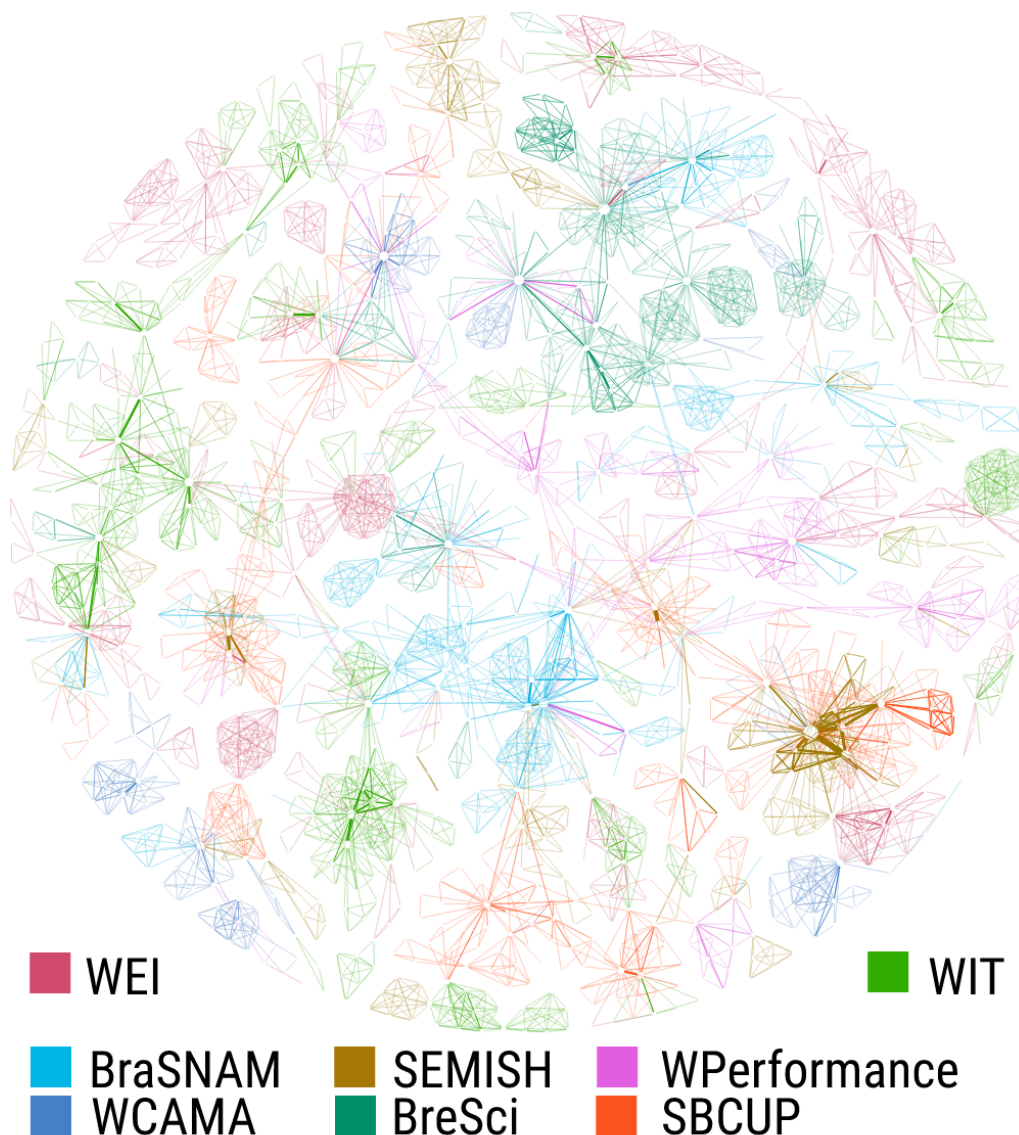


Figura 1. Componente gigante da rede social de coautoria de 10 anos do CSBC.

Tabela 2. Métricas de redes sociais do CSBC e seus eventos.

Métrica	CSBC	BraSNAM	BreSci	SBCUP	SEMISH	WCAMA	WEI	WIT	WPerf.
NumNosGrafo	4.505	588	461	532	569	446	1.196	680	482
NumNosCompGigante	1.773	141	132	100	42	23	96	125	104
NumArestasGrafo	11.397	1.333	1.238	1.304	1.238	1.230	2.566	1.859	892
NumArestasCompGigante	5.523	431	480	313	161	73	238	370	246
Diâmetro	24	7	6	5	2	3	10	8	8
DensidadeGrafo	0,00112	0,00772	0,011675	0,00923	0,00766	0,01239	0,00359	0,00805	0,00769
DensidadeCompGigante	0,00351	0,04366	0,05552	0,06323	0,18699	0,28854	0,05219	0,04774	0,04593
TamAbsolutoCompGigante	1.773	141	132	100	42	23	96	125	104
TamRelativoCompGigante	0,39365	0,24020	0,28696	0,18832	0,07394	0,05169	0,08033	0,18409	0,21622
NumMaxCliquesCompGigante	716	63	49	46	15	7	39	49	41
TamMaxCliquesCompGigante	14	11	12	9	9	8	7	9	8
NumMaxCliquesGrafo	1.608	220	141	182	167	128	396	206	166
TamMaxCliquesGrafo	21	11	12	10	9	21	13	14	8
CoefAgrupCompGigante	0,83994	0,80684	0,85076	0,86133	0,87194	0,89529	0,81078	0,83340	0,83459
CoefAgrupGrafo	0,82495	0,80754	0,86348	0,89288	0,87428	0,86739	0,82275	0,86483	0,78725

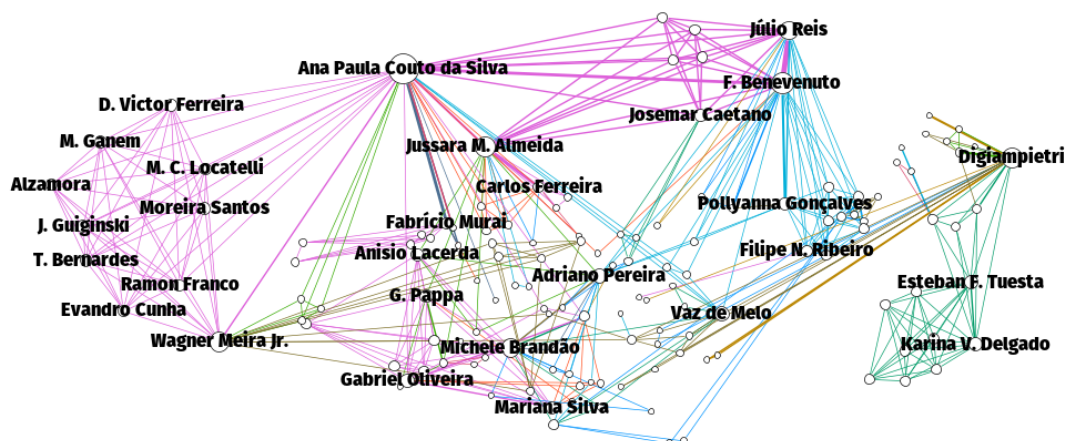


Figura 2. Componente gigante da rede de coautoria de 10 anos do BraSNAM.

dessas comunidades e, possivelmente, são comunidades mais fechadas quando comparadas aos outros eventos. Esse resultado é reforçado pelo valor alto do coeficiente de agrupamento dessas duas redes, 0,874 para o SEMISH e 0,867 para o WCAMA.

Uma das maneiras de se analisar a “saúde” de uma rede social é através do componente gigante. Ter uma parcela significativa de pessoas no componente gigante é algo positivo, pois significa que tais indivíduos podem acessar – direta ou indiretamente – uns aos outros, além de participarem do principal fluxo de informação. Neste sentido, temos os maiores componentes gigantes (em tamanho absoluto, o número de nós) o BraSNAM (141), o BreSCI (132) e o WIT (125). Considerando a proporção entre os nós do componente gigante e a quantidade de nós da rede (tamanho relativo do componente gigante), o BreSCI muda para a primeira posição, com 28% de seus autores participando do componente gigante, seguido do BraSNAM (24%) e WIT (empatado com o SBCUP).

Tabela 3. Matriz do quantitativo de autores em comum entre eventos

Evento/Evento	SEMISH	WIT	WEI	WPerf.	WCAMA	BreSci	BraSNAM	SBCUP
SEMISH		18	25	28	14	16	23	58
WIT	18		65	4	4	6	13	10
WEI	25	65		21	12	17	23	22
WPerf.	28	4	21		6	16	18	34
WCAMA	14	4	12	6		13	7	15
BreSci	16	6	17	16	13		30	12
BraSNAM	23	13	23	18	7	30		9
SBCUP	58	10	22	34	15	12	9	

A Tabela 3 expõe a quantidade de autores em interseção entre os eventos. Em combinação com a Tabela 1, podemos perceber a influência de certas redes sociais em relação a outras. Há forte influência entre autores do SBCUP e SEMISH, acompanhando os dados na Tabela 1, conduzidos por Adenauer Yamin e Ana M. Pernas. O WIT e o WEI possuem influência significativa entre suas comunidades. Como também pode ser visto graficamente na Figura 1, o WCAMA apresenta tanto as menores interseções com outros eventos, como suas sub-redes são isoladas.

4. Considerações finais

Nesta Seção encaminhamos as considerações finais deste trabalho, iniciando por um tópico central, aspectos éticos na Seção 4.1, seguindo pela conclusão na Seção 4.2

4.1. Aspectos éticos

Um dilema ético gerou um debate entre nós: “manter os nomes dos autores ou seguir com anonimato?”. Além da análise dos eventos, do CSBC e das comunidades, reforçamos que são os autores que efetivamente concretizam, sustentam e agenciam o CSBC [Latour 1987, Lefèvre 2005], sendo elementos essenciais à natureza do evento. Cabe reiterar que as informações e contribuições aqui presentes servem apenas para o escopo espacial, os oito eventos, e temporal, dez anos, delimitados, e.g., se alguém ou publicou entre 2005 – 2010 ou publicou no WICS, está fora desta análise.

Moralmente, é ausente uma justificativa ou fundamentação que objetivamente determine um anonimato, pelas diretrizes institucionais governamentais que regem a ética em pesquisa brasileira [Brasil 2016] ou questões éticas gerais de pesquisa [ANPEd 2019, Bos 2020, London 2022]. Lidamos com dados abertos, como disponíveis abertamente na SBC SOL; lidamos com dados de promoção e publicidade de atividade-fim, o que positivamente indica a materialidade do valor do trabalho da pessoa acadêmica-cientista; desligados de qualquer dado pessoal sensível ou potencialmente prejudicial ou consequencialmente negativo às partes; esta é uma pesquisa secundária e trata especificamente de dados meta-científicos, é uma prática comum que pesquisas nesta categoria tratem de autorias, e.g., “neste tópico, este é o autor de maior produtividade e impacto”.

Principalmente, descartamos qualquer juízo moral ou de valor às análises individuais, respeitando o caráter descritivo censitário inicial, e.g., um autor acumula a maior quantidade de autorias, tanto no CSBC no geral como em eventos específicos. Este é um fato objetivo da materialidade concreta. Se disséssemos “Este é o melhor autor do CSBC”, estaríamos emitindo valor e condução moral, ausente neste trabalho.

Adicionalmente, aos autores interessados, principalmente carreiristas acadêmicos-cientistas, podem utilizar destes dados ou informações para sua vantagem, e.g., em memoriais de promoção de carreira. Por outro lado, **eticamente** debatemos uma série de possíveis consequências negativas ou prejudiciais materiais e concretas, excluindo idealismo moralista baseado em ética relativista (e.g., “não quero” ou “não gosto”), e fomos incapazes de pontuar justificativas significativas o bastante para seguir com o anonimato. Sendo assim, mantivemos os nomes e os dados abertos disponíveis associados. Considerados os valores e a natureza do valor do trabalho acadêmicos-científicos [London 2022, Bos 2020, Latour 1987], a intenção de uma suposta “privacidade” ou “confidencialidade”, apesar de justificativa racional, pode ser moralmente questionável.

4.2. Conclusão

Neste trabalho apresentamos análises quantitativas gerais e ARS sobre uma década dos eventos mais longevos do CSBC, através de suas publicações e seus metadados. O CSBC é um evento significativo para ciência computacional brasileira, e através de suas dinâmicas podemos montar parte do quebra-cabeça desta área analisando seus dados, gerando informações e conhecimentos e contribuindo meta-cientificamente com o amadurecimento,

memória e cultura do cenário. Além das análises sobre o BraSNAM, ocasionais no próprio evento [Lobato et al. 2021, Digiampietri et al. 2017], estendemos a composição a outros eventos e à combinação destes, expondo a maioria das interações entre eventos e pesquisadores no CSBC, amadurecidos e já bem-estabelecidos. Desconhecemos trabalhos similares ou fielmente próximos a este, em magnitude de abrangência, completude ou profundidade.

Além das contribuições analítico descritivas aqui apresentadas, algumas limitações e ameaças à validade são pertinentes [Marconi e Lakatos 2017, Wazlawick 2014]: (i) limitados aos eventos e intervalo de tempo cobertos, representando a realidade parcial do CSBC, sem que isto restrinja sua complementação, o que é bem-vindo; (ii) a qualidade dos dados inicialmente presentes na SBC SOL está fora do nosso controle, apesar do tratamento extensivo dos elementos extraídos; (iii) a margem de erro na etapa de tratamento dos dados, que apesar de minuciosa é sujeita a falhas, confiamos que ocorram em menos de 1% dos registros totais (≈ 65). Demais análises podem ser abordadas em trabalhos futuros, quantitativas ou qualitativas, por dimensões ou perspectivas variadas, e.g., outras abordagens ou métricas de ARS.

5. Agradecimentos

Este trabalho foi realizado com apoio da CAPES – Brasil – Código de Financiamento 001; e parcialmente pelo programa *Oracle for Research* (número do prêmio CPQ-2160239).

Referências

- [ANPEd 2019] ANPEd (2019). *Ética e pesquisa em educação: subsídios – volume 1*. volume 1. ANPEd, Rio de Janeiro, RJ.
- [Barabási et al. 2002] Barabási, A., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., e Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3):590–614.
- [Boaventura et al. 2014] Boaventura, M., Boson, K., Silva, A., Veloso, A., e Junior, W. M. (2014). Caracterização temporal das redes de colaboração científica nas universidades brasileiras: Anos 2000-2013. In *Anais do III BraSNAM*, pp. 9–20, Porto Alegre, RS, Brasil. SBC.
- [Bos 2020] Bos, J. (2020). *Research Ethics for Students in the Social Sciences*. Springer Cham, 1st edition.
- [Brasil 2016] Brasil (2016). Ministério da saúde. RESOLUÇÃO Nº 510, DE 07 DE ABRIL DE 2016. Disponível em: https://4658.short.gy/CEP_2016 [acesso 27/02/2023].
- [Caparelli e Digiampietri 2018] Caparelli, L. e Digiampietri, L. A. (2018). Combinando análise bibliométrica e análise de redes sociais para a avaliação de grupos acadêmicos. In *Anais do VII BraSNAM*, Porto Alegre, RS, Brasil. SBC.
- [Correa e Ma 2011] Correa, C. D. e Ma, K.-L. (2011). *Visualizing Social Networks*, pp. 307–326. Springer US, Boston, MA.
- [Digiampietri et al. 2014] Digiampietri, L., Alves, C., Trucolo, C., Delgado, K., e Mugnaini, R. (2014). Análise da rede de relacionamentos dos doutores brasileiros. In *Anais do VIII BreSci*, pp. 21–28, Porto Alegre, RS, Brasil. SBC.

- [Digiampietri et al. 2015] Digiampietri, L., Linden, R., e Barbosa, L. (2015). Desambiguação de nomes em redes sociais acadêmicas: Um estudo de caso usando dblp. In *Anais do IV BraSNAM*, Porto Alegre, RS, Brasil. SBC.
- [Digiampietri et al. 2017] Digiampietri, L., Mugnaini, R., Pérez-Alcázar, J., Delgado, K., Tuesta, E., e Mena-Chalco, J. (2017). Análise da evolução, impacto e formação de redes nos cinco anos do brasnam. In *Anais do VI BraSNAM*, Porto Alegre, RS, Brasil. SBC.
- [Hosseini et al. 2022] Hosseini, M., Lewis, J., Zwart, H., e Gordijn, B. (2022). Evolution of the social network of scientific collaborations. *Science and Engineering Ethics*, 28(3):590–614.
- [Ioannidis 2018] Ioannidis, J. P. A. (2018). Meta-research: Why research on research matters. *PLoS Biol*, 16(3).
- [Lakatos e Marconi 2019] Lakatos, E. M. e Marconi, M. (2019). *Sociologia Geral*. Atlas, 8th edition.
- [Latour 1987] Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard university press.
- [Lefèvre 2005] Lefèvre, W. (2005). Science as Labor. *Perspectives on Science*, 13(2):194–225.
- [Lima Filho et al. 2023] Lima Filho, S., Carvalho, L. P., Suzano, J., Brandão, M., Oliveira, J., e Santoro, F. M. (2023). CSBCSet: Um conjunto de dados para uma década de CSBC, seus eventos e publicações. In *Anais do XII BraSNAM*, Porto Alegre, RS, Brasil.
- [Lobato et al. 2021] Lobato, F., Sousa, G., e Jr., A. J. (2021). Brasnam em perspectiva: uma análise da sua trajetória até os 10 anos de existência. In *Anais do X BraSNAM*, pp. 217–228, Porto Alegre, RS, Brasil. SBC.
- [London 2022] London, A. J. (2022). *For the Common Good: Philosophical Foundations of Research Ethics*. Oxford University Press, 1st edition.
- [Marconi e Lakatos 2017] Marconi, M. d. A. e Lakatos, E. M. (2017). *Fundamentos de Metodologia Científica*. Atlas, São Paulo, SP, 8ª edition.
- [Peterson e Panofsky 2020] Peterson, D. e Panofsky, A. (2020). Metascience as a scientific social movement.
- [Santana e Braga 2020] Santana, T. e Braga, A. (2020). Uma análise cienciométrica das publicações do congresso da sociedade brasileira de computação na perspectiva das mulheres na computação. In *Anais do XIV WIT*, pp. 279–283, Porto Alegre, RS, Brasil. SBC.
- [Wang 2018] Wang, C. (2018). Scientific culture and the construction of a world leader in science and technology. *Cultures of Science*, 1(1):1–13.
- [Wazlawick 2014] Wazlawick, R. S. (2014). *Metodologia de Pesquisa para Ciência da Computação*. Elsevier, São Paulo, SP, 2ª edition.