

ESTUDO DA COLABORAÇÃO CIENTÍFICA NO ENANCIB: ASPECTOS DA PARTICIPAÇÃO INTERINSTITUCIONAL NOS GRUPOS DE TRABALHO

Ronnie Fagundes de Brito

Doutor em Engenharia e Gestão do Conhecimento pela Universidade Federal de Santa Catarina (UFSC)

Tecnologista do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Ingrid Torres Schiessl

Bacharela em Biblioteconomia pela Universidade de Brasília (UnB)

Assistente de pesquisa no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Milton Shintaku

Doutor em Ciência da Informação pela Universidade de Brasília (UnB)

Tecnologista do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

INTRODUÇÃO

Desde a criação do primeiro curso de pós-graduação em Ciência da Informação (CI) no Brasil em 1970, ofertado pelo Instituto Brasileiro de Biblioteconomia e Documentação (IBBD), discussões foram encaminhadas para o fortalecimento da área de conhecimento. Tanto que, em 1989 foi criada a Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB), como relata Barreto (2009). Em 1993, por meio da recém-criada ANCIB, ocorre a primeira edição do Encontro Nacional de Pesquisa em Ciência da Informação (ENPeCI), que já no segundo encontro foi denominado de Encontro Nacional de Ciência da Informação e Biblioteconomia (ENANCIB). No primeiro ENPeCI/ENANCIB, ocorrido em Belo Horizonte (MG), existiam sete Grupos Temáticos (GT), apresentando 22 trabalhos em seus anais. Na edição de 2018, ocorrido em Londrina - Paraná, foram 11 GTs com 486 trabalhos publicados nos anais, revelando a evolução deste encontro.

Parece ser consenso entre os profissionais de informação, que o ENANCIB é o principal evento relacionado a pesquisas efetuadas no âmbito da pós-graduação em CI, tanto que, Mugnaini (2015) destaca a importância deste evento para os docentes e discentes dos cursos de pós-graduação em CI. Da mesma forma, Café e Cunha (2007) advogam que o ENANCIB é a oportunidade de interação entre pesquisadores, sendo um evento funda-

mental para a evolução da área de conhecimento. Por isso, torna-se importante conhecer quais são as instituições que mais apresentam os seus resultados de pesquisas e como elas colaboram entre si.

Assim, o presente estudo tem o objetivo de apresentar o cenário de participação institucional nos ENANCIBs de 2011 – início da configuração atual dos GTs – a 2019, por meio dos trabalhos publicados nos anais. Com isso, verificar quais instituições utilizam esse evento como parte das discussões das pesquisas, com proposto por Lievrouw e Carley, (1990) em seu modelo de comunicação científica, no qual a apresentação de resultados de pesquisa em eventos é um processo intermediário na disseminação dos resultados de pesquisa. Os dados coletados estão disponíveis em repositório de dados. (BRITO; SCHIESSL; SHINTAKU, 2020b)

PROCEDIMENTOS METODOLÓGICOS

Para atender ao objetivo do estudo, optou-se pela abordagem mista, no qual os dados são coletados de forma quantitativa e analisados de forma qualitativa, numa abordagem aninhada. (CRESWELL, 2007) Assim, alinham-se a precisão dos dados quantitativos com a profundidade da análise qualitativa, necessária para o estudo, que procura não apenas verificar padrões, mas entendê-los. Quanto ao tipo de pesquisa, o estudo apresenta características descritivas, observadas por Gil (2017) como sendo aquelas que procuram caracterizar uma população ou fenômeno, ou mesmo relações entre variáveis. Para tanto, requer coleta sistemática de dados, de forma a criar bases de dados estruturados, que possibilitam a descrição do objeto de pesquisa.

A pesquisa possui o horizonte temporal transversal, buscando apenas os trabalhos das edições dos anos de 2011 a 2019, por seguirem a configuração dos GTs atuais. A coleta de dados, por sua vez, foi feita de forma automatizada, com apoio de programas e protocolos de comunicação informatizados. Os dados foram extraídos dos anais dos ENANCIBs que estão majoritariamente hospedados em portais implementados com o software Open Conference System (OCS), sistema informatizado que implementa o Open Archives Initiative - Protocol Metadata Harvesting (OAI-PMH). Nesse processo, foram coletados metadados sobre os trabalhos publicados nos anais, baseados no padrão Dublin Core, de forma a criar uma base de dados estruturados, com campos fornecidos pela interface OAI-PMH do OCS, como título, autor, assunto, descrição, e outros.¹ Especificamente, foi usado o campo Author/Creator para a atribuição de autoria institucional, pois os anais seguem o padrão <autor>;<instituição>. Nos casos em que a coleta não ocorreu de forma automática, foi feita manualmente, de forma a ter a base completa.

1 Conforme especificação. Ver: https://www.openarchives.org/OAI/2.0/oai_dc.xsd

O estudo utilizou um *script* desenvolvido em Jupyter Notebook. (BRITO; SCHIESSL; SHINTAKU, 2020a) O qual consiste uma plataforma computacional operacional em web, que possibilita operar com fontes diversas, possibilitando combinações úteis na pesquisa sendo considerada por como uma das ferramentas livres mais promissoras para trabalhar com Ciência Aberta. (PEREZ; GRANGER, 2015; RANDLES et al., 2017) Assim, variáveis como autor, instituição do autor, ano de publicação e GT puderam ser extraídos e processados, fornecendo indicadores para análise.

Além disso, o estudo adapta a Lei de Bradford, que identifica as revistas núcleo de um tópico, para verificação das instituições núcleo em cada GT, comparando a participação das instituições. De forma a verificar quais são as instituições núcleos em cada GT é considerado núcleo do estudo do tema, com índice de 80% dos trabalhos submetidos. A seleção das instituições participantes do estudo deu-se pelas listadas nos programas de pós-graduação em comunicação pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), da área de “comunicação e informação”,² incluindo Ciência da Informação, Comunicação e Museologia, um total de 73 instituições, criando um filtro para a identificação da autoridade, de forma a excluir instituições cujo autor participa do ENANCIB, mas não possuem programa de pós-graduação na área. Dessa forma, instituições como Fundação Getúlio Vargas, Museu Goeldi, Embrapa e Universidade Federal do Maranhão ficaram fora do estudo.

Ao se falar da métrica “participação”, contabiliza “Autores por Trabalhos”, de forma que um trabalho publicado com três autores será contabilizado três vezes, permitindo analisar assim as participações interinstitucionais. Não foi feita distinção entre artigos completos e posters, pois em algumas edições essa informação não pode ser obtida diretamente dos metadados, em outras constava apenas no documento trabalho publicado.

Um problema encontrado foi o fato da interface OAI-PMH do OCS listar todos os documentos incluindo os rejeitados pelos avaliadores. Então, para analisar apenas os trabalhos publicados, inseriu-se uma condição para recuperar os metadados, somente dos registros que possuíam documentos no formato .pdf ou que constavam nos anais.³ Importa ressaltar, que alguns trabalhos diferem quanto à grafia, ou seja, o título registrado no OCS não é igual ao publicado nos anais, portanto, esses trabalhos não foram contabilizados para amostra.

Outro ponto, foram os registros que não se enquadraram nos critérios adotados para a seleção da amostra, isto é, foi atribuído o valor “Não enquadrado” para as seguintes condições: a) a instituição não constava na lista do programa de pós-graduação da Capes; b) não foi possível a coleta da instituição por meio dos metadados e/ou busca no texto do

² Ver: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/programa/quantitativos/quantitativoAreaConhecimento.jsf?areaAvaliacao=31>

³ Edições de 2011 e 2014.

documento. Importa ressaltar, que um trabalho pode conter autores, que são de instituições cadastradas na Capes e de não cadastradas.

RESULTADOS E DISCUSSÕES

A coleta de dados foi executada automaticamente exceto nos anos de 2010 e 2011, feitos manualmente, criando um recorte composto por 3.145 títulos, com 7.038 participações e 3.087 autores, de 41 instituições, que foram publicados durante 9 edições do evento, de 2011 a 2019. Com essa base foi possível verificar certa tendência de aumento no número de trabalhos a cada edição. Durante o período analisado, observou-se um aumento de 200% na quantidade de estudos apresentados no evento, proporcionalmente, verificou-se o crescimento do número de autores.

Uma das Lei de Bradford, também conhecida pela denominação de Lei da Distribuição, tem atuação original na quantidade de publicação de artigos de um determinado assunto, de forma a possibilitar a verificação das revistas núcleos de um tema. Assim, adaptando para o estudo a essa lei, procura-se verificar quais são as instituições mais produtivas em GTs específicos, que neste caso determinam as instituições núcleo nas linhas de pesquisa, ou temas.

Neste sentido, alinhado a Lei de Distribuição, para a definição das instituições núcleo de cada GT, selecionou-se as mais produtivas, representando a 80% dos trabalhos publicados em cada grupo. O Quadro 1 apresenta as instituições mais produtivas em cada GT, seguida pela quantidade de instituições que tiveram artigos publicados. Com isso, procurou-se definir quais são as instituições núcleo em pesquisa e publicação sobre cada linha, conforme a Lei adaptada de Bradford.

Quadro 1: Instituições responsáveis por 80% das publicações em cada GT

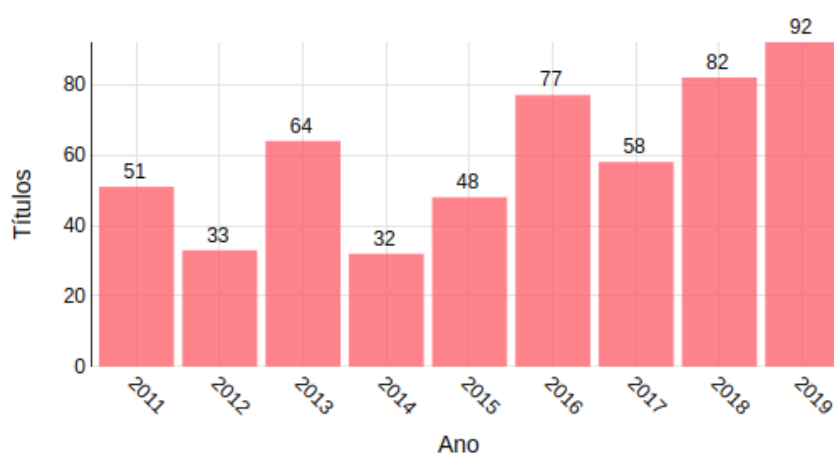
GT	Instituições	Instituições
1	IBICT; UFPB; UFF ; UNIRIO; UFMG; USP; UNB	26
2	UFMG; UNESP; UFF	26
3	UFMG; USP; UNESP; IBICT; UEL	27
4	UFPB; UFMG; UNESP	29
5	IBICT; UFMG; UNIRIO; UFF; UFPB	26
6	UFSC; UFMG; UFBA; UFPB; IBICT; UFF; UNIRIO; USP; UDESC	26
7	UNESP; UFRGS; UNIRIO; UFPE; IBICT	27
8	UNESP; UFPB; USP	29
9	UNIRIO; MAST; IBICT; UFMG; USP; UNB	21
10	UFPE; UNIRIO; UFMG; IBICT; UNESP; UFBA; USP	27
11	FIOCRUZ; UFMG; IBICT; UNESP; UFBA; UNIRIO; USP; UEL; UFAL; UFF; UFPB; UFC; UFSC	24

Fonte: elaborado pelos autores.

Em alguns GT as instituições núcleo são mais representativas, como nos GT-2, GT-4 e GT-8, em que apenas três instituições, pouco mais de 20% das instituições, representam 80% dos trabalhos submetidos. Em outros casos, como no GT-1, GT-3, GT-5, GT-7, GT-9 e GT-10, as instituições núcleo ficam entre cinco e sete instituições. Entretanto, o único grupo que apresenta uma distribuição anormal é o GT-11 com 13 instituições de 24, quebrando totalmente a lei, revelando que o tema de “Informação e Saúde” tem uma dispersão maior.

A colaboração interinstitucional está presente em 537 artigos, correspondendo a cerca de 17% dos trabalhos publicados no ENANCIB. Ao longo do tempo a evolução da colaboração pode visto na Gráfico 1.

Gráfico 1: Quantidade de títulos com autores de mais de uma instituição



Fonte: elaborado pelos autores.

Entre as instituições, nota-se que o Instituto Brasileiro de Informação em Ciência e Tecnologia/ Universidade Federal do Rio de Janeiro (Ibict/UFRJ) e a Universidade Federal do Estado do Rio de Janeiro (Unirio) foram as que mais colaboraram, com 49 títulos ao longo dos nove anos. Quando se analisam as instituições que tem 7 ou mais trabalhos publicados em colaboração no ENANCIB, obtemos o grafo apresentado no Figura 1. Nota-se a Universidade Federal da Paraíba (UFPB), Unirio, Ibict/UFRJ, Fiocruz, Universidade Federal da Pernambuco (UFPE) e Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp) como núcleos na colaboração, correlacionando-se também como núcleos de cada GT.

e informação pela Capes, 56% utilizem o evento para disseminação dos resultados de pesquisa, como proposto por Lievrouw e Carley (1990) em seu modelo de comunicação científica.

Outro aspecto que deve ser observado na análise é quantidade de registros com o valor de autoria “não enquadrada nos critérios utilizados na pesquisa”, correspondendo a 11% das participações e 14% dos trabalhos apresentados nas nove edições, sugerindo novos estudos focados nas instituições que não possuem programa de pós-graduação em CI, mas que participam do ENANCIB. Da mesma forma, como trabalho futuro, é possível realizar a formação de círculos de coautoria (CHAKRABORTY et al., 2015), assim como o detalhamento da colaboração além dos GTs, como por temas e assuntos, com dados obtidos a partir de análises dos títulos e resumos.

Materiais complementares, como o script em Jupyter Notebook e a base de dados resultante, estão disponíveis para replicação e reuso em novas pesquisas, alinhado ao movimento dos dados abertos. Espera-se continuar a análise nas próximas edições do ENANCIB e acompanhar a evolução do evento, de forma a contribuir com as discussões envolvidas neste evento, alinhado a proposta original do ENANCIB em ser o *locus* de discussões sobre a CI.

REFERÊNCIAS

- BRITO, R. F. D.; SCHIESSL, I.; SHINTAKU, M. *A script to harvest publications and process authorship and institutional affiliation of Enancib Conference papers*. Geneva: Zenodo, 2020a. Disponível em: <https://zenodo.org/record/3665752>. Acesso em: 28 maio 2020.
- BRITO, R. F. D.; SCHIESSL, I.; SHINTAKU, M. *Enancib Participation list*. Geneva: Zenodo, 2020b. Disponível em: <https://zenodo.org/record/3665763>. Acesso em: 28 maio 2020.
- CAFÉ, L. M. A.; CUNHA, M. F. V. Editorial das organizadoras do VI ENANCIB. *Encontros Bibli*, Florianópolis, v. 11, n. 1, p. 1, 2007.
- CHAKRABORTY, T.; PATRANABIS, S.; GOYAL, P. et al. On the Formation of Circles in Co-authorship Networks. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 21., 2015, Sydney. *Anais [...]*. Sydney: ACM Press, 2015. p. 109-118. Disponível em: <http://dl.acm.org/citation.cfm?doid=2783258.2783292>. Acesso em: 28 maio 2020.
- CRESWELL, J. *Projeto de pesquisa: métodos qualitativo, quantitativo e misto*. Porto Alegre: Artmed, 2007.
- GIL, A. C. *Como Elaborar Projetos de Pesquisa*. 6. ed. São Paulo: Atlas, 2017.
- HURD, J. M. The transformation of scientific communication: a model for 2020. *Journal of the American Society for Information Science*, New York, v. 51, n. 14, p. 1279-1283, 2000. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/1097-4571%282000%299999%3A9999%3C%3A%3AAID-ASII044%3E3.o.CO%3B2-1>. Acesso em: 28 maio 2020.
- LIEVROUW, L. A.; CARLEY, K. Changing patterns of communication among scientists in an era of “telescience”. *Technology in Society*, New York, v. 12, n. 4, p. 457-477, 1990.

MUGNAINI, R. Ciclo avaliativo de periódicos no Brasil: caminho virtuoso ou colcha de retalhos?. *In: ECONTRO NACIOANL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 16., 2015. João Pessoa. *Anais [...]*. João Pessoa: Universidade Federal da Paraíba, 2015. Disponível em: <http://www.ufpb.br/evento/index.php/ENANCIB2015/ENANCIB2015/paper/view/2984>. Acesso em: 28 maio 2020.

PEREZ, F.; GRANGER, B. E. Project Jupyter: computational narratives as the engine of collaborative data science. *Retrieved September*, [s. l.], v. 11, n. 207, p. 108, 2015.

RANGLES, B. M.; PASQUETTO, I. V.; GOLSHAN, M. S. *et al.* Using the Jupyter Notebook as a Tool for Open Science: an empirical study. *In: JOINT CONFERENCE ON DIGITAL LIBRARIES*, 16., 2017, Toronto. *Anais [...]*. Toronto: IEEE, 2017. p. 1-2. Disponível em: <http://ieeexplore.ieee.org/document/7991618/>. Acesso em: 28 maio 2020.