

Cinco anos da Escola Regional de Engenharia de Software: Uma Análise Bibliométrica

Andréa Sabedra Bordin¹, Ramon Giovani Guth da Silva¹

¹Universidade Federal de Santa Catarina (UFSC) - Campus Araranguá
Santa Catarina, Brasil.

andrea.bordin@ufsc.br, ramomguth@gmail.com

Abstract. *This article presents a bibliometric analysis of the five years of the Regional School of Software Engineering. Through procedures of collection, processing and analysis of data from publications indexed in the SOL library, indicators of productivity, impact and scientific collaboration were obtained. The results reveal that an expressive number of articles, more productive authors and with more frequent collaborative relationships, come from the same institution. It is concluded that the event can be further consolidated by increasing and intensifying the ties of collaboration between the institutions.*

Resumo. *Este artigo apresenta uma análise bibliométrica dos cinco anos da Escola Regional de Engenharia de Software. Por meio de procedimentos de coleta, processamento e análise dos dados das publicações indexadas na biblioteca SOL foram obtidos indicadores de produtividade, impacto e de colaboração científica. Os resultados revelam que um expressivo número de artigos, autores mais produtivos e com relações de colaboração mais frequentes são oriundos de uma mesma instituição. Conclui-se que o evento pode se consolidar mais aumentando e intensificando os laços de colaboração entre as instituições.*

1. Introdução

Eventos científicos do tipo Escola Regional vêm sendo apoiados pela Sociedade Brasileira de Computação (SBC) há décadas. Este tipo de evento se caracteriza por congrega pesquisadores e estudantes da área de Computação em nível de graduação e pós-graduação em um escopo regional para divulgação de suas pesquisas e relatos de experiências.

A Escola Regional de Engenharia de Software (ERES) foi criada com o objetivo de disseminar o conhecimento e as boas práticas em Engenharia de Software (ES) e teve sua primeira edição em 2017 na cidade de Alegrete (RS) sob a organização da Universidade Federal do Pampa (UNIPAMPA) e do Instituto Federal Farroupilha (IFF). De lá para cá foram realizadas mais quatro edições deste evento: a segunda em Dois Vizinhos (PR), organizada pela Universidade Tecnológica Federal do Paraná (UTFPR); a terceira em Rio do Sul (SC), organizada pela Universidade do Estado de Santa Catarina (UDESC) e Instituto Federal Catarinense (IFC); a quarta e a quinta foram realizadas remotamente.

O período de cinco anos de realização consecutiva da ERES como evento científico, com um total de 228 artigos recebidos e 122 artigos aceitos, indica o interesse e os esforços da comunidade de Engenharia de Software da região Sul do país em contribuir com os avanços nesta área de conhecimento. Dessa forma, é um importante marco temporal que suscita alguns questionamentos de natureza bibliométrica.

Análises bibliométricas resultam em indicadores bibliométricos, dentre os quais estão os que medem a produtividade científica e tecnológica, como o número de artigos e livros publicados e patentes registradas; os que buscam analisar o uso e a qualidade dos documentos, baseados principalmente nos estudos de citações e os indicadores de colaboração, que analisam redes sociais e colaborativas estabelecidas entre pesquisadores, organizações ou países [Rehn and Kronman 2008].

O objetivo deste artigo é apresentar uma análise bibliométrica das publicações da ERES em seus últimos cinco anos, na qual são identificados e analisados os autores e instituições que mais publicaram artigos, os artigos de maior impacto, os grupos de autores com maior laços de colaboração científica, dentre outros. A partir desta análise é possível tomar decisões em nível individual e em nível de gestão do evento. Por exemplo, ao conhecer os autores mais produtivos, pesquisadores juniores podem estabelecer contato para orientações futuras, assim como pesquisadores mais experientes podem aumentar e qualificar sua rede de colaboração. Já em nível organizacional, a comissão responsável pode convidá-los para se envolverem na programação e na organização do próprio evento.

O restante deste artigo está organizado da seguinte forma: na Seção 2 é descrita a metodologia e algumas definições do método utilizado; os resultados são apresentados e discutidos na Seção 3 e, por fim, as considerações finais são expostas na Seção 4.

2. Metodologia de pesquisa

O desenvolvimento desta pesquisa passou por três etapas: (1) Coleta de dados, (2) Processamento dos dados e (3) Análise dos resultados. Essas etapas serão detalhadas nas próximas subseções.

2.1. Coleta de dados

Os dados utilizados nesta pesquisa compreendem o título, autores e respectiva instituição de afiliação dos artigos publicados nos Anais das cinco edições da ERES (2017 a 2021). Os dados foram coletados automaticamente através de um algoritmo de *web scraping* executado nas páginas dos Anais que estão disponíveis publicamente na Biblioteca Digital da SBC - SOL¹. Os dados coletados estão disponíveis no repositório Zenodo².

2.2. Processamento dos dados

Após a coleta, os dados foram processados de forma a gerar resultados para as análises bibliométricas. Alguns procedimentos foram realizados manualmente e outros foram realizados de forma automatizada.

Primeiramente os nomes dos autores foram normalizados com o objetivo de se identificar de forma única um autor/pesquisador. Este procedimento foi feito manualmente, o que demandou um tempo considerável em função da análise criteriosa para evitar problemas como a homonímia, quando existem autores diferentes com um mesmo nome, assim como a existência de diversas variações de nome de um mesmo autor.

Os indicadores de produtividade como número de artigos por autor e instituição foram calculados por meio de algoritmos desenvolvidos especificamente para esse fim. Já

¹<https://sol.sbc.org.br/index.php/eres>

²<https://zenodo.org/record/7017873.YztzO3aZPIU>

os indicadores de impacto como número de citações recebidas e qualidade das publicações que citaram os artigos da ERES foram obtidos por meio de pesquisas no *Google Scholar* e na Plataforma Sucupira.

Os indicadores de colaboração científica foram obtidos pelo método de Análise de Redes Sociais (ARS). No método de ARS os autores de artigos científicos e suas relações de coautoria são modelados como um grafo (rede), onde os vértices representam os autores e as arestas representam as relações de coautoria [Newman 2004]. A partir deste grafo é possível calcular várias métricas ou indicadores como grupos de vértices (autores) isolados, nível de conectividade da rede, autores que mais colaboram com outros autores, dentre outros.

Para a geração do grafo de coautoria, todos os autores foram identificados de forma individual (código e nome), assim como foram processadas todas as relações de coautoria entre os autores, conjuntamente com a frequência da relação (número de vezes que dois autores escreveram conjuntamente). Este processo foi realizado de forma automatizada através de um algoritmo que gerou o arquivo da rede de coautoria de todas as edições da ERES.

O arquivo de rede foi utilizado como entrada para o software de análise exploratória de dados Gephi³. Com esta ferramenta foi possível visualizar a rede e calcular várias métricas de ARS, as quais serão explicadas na etapa seguinte.

2.3. Análise dos dados

As análises bibliométricas foram organizadas em dois grupos: as que resultaram em **indicadores de produtividade e de impacto** e as que resultaram em **indicadores de colaboração científica**. Como indicadores de produtividade e impacto foram obtidos o número de artigos por ano, o número de artigos por autor e instituição e o número de citações recebidas.

A análise dos indicadores de colaboração científica foi realizada em três níveis, de acordo com Wasserman e Faust [Wasserman and Faust 1994]: nível de grupo, onde foram observadas as formações de subgrupos de autores da rede; nível individual, onde as propriedades individuais dos vértices (autores) foram analisadas e o nível de rede, onde propriedades gerais da rede foram analisadas.

No nível de análise **de grupo** foram identificados os subgrafos ou componentes da rede. Componentes são subgrafos onde os vértices estão conectados por algum caminho possível, mas desconectados dos demais subgrafos. Componente gigante é o nome dado ao subgrafo que contém o maior número de vértices conectados. [Wasserman and Faust 1994].

No nível de análise **individual**, foram analisadas métricas de centralidade que buscam descrever as propriedades de localização de um vértice (autor) na rede, tais como a Centralidade de Grau e a Centralidade de Intermediação. A Centralidade de Grau (CG) é medida por meio do número de vértices adjacentes a um vértice na rede e reflete o papel do autor em termos de popularidade e atividade. Já a Centralidade de Intermediação (CI) mede o quanto um determinado vértice se encontra entre os vários outros vértices do grafo, ou seja, atribui importância a um vértice em função do fluxo

³Disponível em: <http://gephi.org/>

que passa por ele para interligar outros dois vértices da rede, através do menor caminho possível [Freeman 1978].

No nível de análise **de rede** encontrou-se a Densidade, que mede o estado de completude de um grafo. Quanto mais vértices estão ligados uns aos outros, mais denso ou conectado será o grafo. O Grau Médio de uma rede também é um indicador de conectividade da rede, pois determina o número médio de relações de adjacência de um vértice.

3. Resultados e Discussão

Ao longo dos últimos cinco anos da ERES, foram submetidos 228 artigos, dos quais 122 foram aceitos, representando uma taxa de aceitação de 53,3%. A Figura 1 exibe o Total de Artigos Recebidos (TAR) e o Total de Artigos Aceitos (TAA) a cada edição do evento. Pode-se observar que as duas maiores taxas de aceitação correspondem à edição de 2018, com 68,7% e à última edição do evento (2021), com 72%, que são também as edições com menor número de artigos recebidos. O maior número de submissões ocorreu em 2020, o primeiro ano da pandemia de Covid-19. Neste caso, os motivos podem estar relacionados com a facilidade e os custos que a modalidade remota oferece. No entanto, o mesmo cenário não se repetiu na edição de 2021, que também foi online. Neste caso, o motivo pode estar relacionado com as dificuldades de engajamento dos alunos nas atividades remotas de pesquisa.



Figura 1. Número de artigos recebidos x Número de artigos aceitos

3.1. Indicadores de produtividade e impacto

Os indicadores de produtividade da ERES apresentados nesta seção revelam os autores que mais publicaram, assim como as instituições e os estados que mais tiveram artigos publicados. Em relação aos autores mais produtivos, a Tabela 1 exibe o nome dos autores com os cinco maiores números de publicações, o vínculo institucional (P de Professor e A de Aluno), a instituição de filiação e o número de artigos.

Como pode ser observado, todos os autores são da mesma instituição e cinco deles são professores. Destaca-se também o número expressivo de publicações dos dois primeiros autores, que em média publicaram aproximadamente cinco (5) artigos em cada edição do evento. Todos os professores atuavam no curso de graduação em Engenharia de Software na ocasião das publicações, sendo que os quatro primeiros também atuavam no

Tabela 1. Número de artigos por autor

Autor	IES	Nu. Artigos
Maicon Bernardino (P)	UNIPAMPA	29
Elder de Macedo Rodrigues (P)	UNIPAMPA	25
Fábio Basso (P)	UNIPAMPA	16
João Pablo S. da Silva (P)	UNIPAMPA	10
Andréa Sabedra Bordin (P)	UNIPAMPA	6
Jonnathan Riquelmo (A)	UNIPAMPA	6
Yury Lima (A)	UNIPAMPA	6

Programa de Pós-Graduação de Engenharia de Software (PPGES) profissional em nível *strictu sensu*. Os dois últimos autores foram alunos de graduação.

O número de artigos publicados por instituição de filiação dos autores também foi levantado. Trinta (30) instituições e um (1) pesquisador independente publicaram na ERES. Desses, quinze (15) publicaram um (1) artigo, seis (6) publicaram dois (2) artigos e três (3) instituições publicaram três (3) artigos. As demais são exibidas na Tabela 2, com a distribuição dos artigos ao longo dos anos, pois são as instituições com maior número de artigos publicados. Destaca-se que cada instituição foi contabilizada uma única vez por artigo, independentemente do número de coautores da mesma instituição.

A UNIPAMPA (Campus Alegrete) é a instituição de filiação dos autores que mais publicou na ERES, com quase o triplo de artigos em relação à segunda colocada (UTFPR-Dois Vizinhos). Além disso, manteve frequência em todos os anos, além do maior número de artigos ao longo dos anos, com exceção de 2018. Todas as instituições da Tabela 2 possuem cursos de graduação em ES, exceto a Universidade Estadual de Maringá (UEM). Apenas a UNIPAMPA (Campus Alegrete) possui um Curso de Mestrado Profissional em Engenharia de Software (MPES). PUCRS e UEM mantêm programas de pós-graduação com cursos de mestrado e doutorado, onde uma das linhas de pesquisa é Engenharia de Software. Assim, entende-se que o protagonismo dessas instituições da região Sul deva-se à existência de comunidades fortes de pesquisadores vinculados aos cursos e linhas de pesquisa em ES que fomentam o desenvolvimento da área.

Tabela 2. Número de artigos por instituição e por ano

Instituição	2017	2018	2019	2020	2021	Total
UNIPAMPA (Campus Alegrete)	18	8	12	12	8	58
PUCRS	7		1	1	1	11
UTFPR (Campus Dois Vizinhos)		8	2	7	3	20
UEM				7	1	8
UNIJUÍ				6		6

O levantamento do número de artigos por região geográfica revelou que a região Sul foi responsável por 89,8% dos artigos publicados, seguido pela região Sudeste com 4,4% dos artigos, região Centro-Oeste com 2,9%, região Nordeste com 2,2% e, por final, a região Norte com 0,7%. Esses números confirmam o objetivo de congregar a comunidade científica da região Sul do país que a ERES possui.

Os indicadores de impacto foram levantados com base na análise das citações re-

cebidas. Trinta (30) artigos publicados na ERES receberam ao menos uma (1) citação, representando 24.6% do total de artigos. Destes, três (3) artigos receberam três (3) citações cada, dezesseis (16) artigos receberam duas (2) citações e os demais uma (1) citação, totalizando 52 citações. Do total de citações recebidas, 13 são oriundas de publicações estrangeiras em anais de eventos ou periódicos e foram feitas para artigos da ERES publicados em inglês. A Tabela 3 mostra os três artigos mais citados, conjuntamente com o ano de publicação e a IES de afiliação dos autores. Destes, apenas o segundo artigo pertence a Trilha de Pós-Graduação.

Tabela 3. Artigos com maior número de citações

Artigo	Ano	IES
FERREIRA, M.; COSTA, L.; SOUZA, F. C. Search-based Test Data Generation for Mutation Testing: a tool for Python programs.	2020	UTFPR (Dois Vizinhos)
BELINAZO, G.; FONTOURA, L. Mapeamento Sistemático de Gamificação em Engenharia de Software.	2017	UFSM (Santa Maria)
LOPES, J.; MEDEIROS, G.; FIALHO, D.; BORDIN, A. Resolução de Problemas no Curso de Engenharia de Software: Uma Experiência Envolvendo Extensão e Ensino.	2017	UNIPAMPA (Alegrete)

Para além do aspecto quantitativo, é importante destacar a qualidade ou impacto das conferências e periódicos cujas publicações citaram artigos da ERES. O primeiro e o segundo artigo da Tabela 3 foram citados por publicações com Qualis B1 e B4. Cabe mencionar também outra publicação não listada na tabela, mas que também foi citada por uma publicação Qualis B1, o artigo "A customizable SPL scoping process for SPL reengineering" de autoria de Marchezan, L.; Rodrigues, E.; Bernardino, M.; Basso, F., filiados à UNIPAMPA e publicado em 2019.

3.2. Indicadores de Colaboração Científica

A rede de coautoria de artigos da ERES (2017-2021) é composta por 240 vértices (autores) e 524 arestas (relações de coautoria). A Figura 2 fornece uma visão geral inicial da rede, onde cada autor é representado pela cor da instituição de filiação.

Partindo do nível de análise dos subgrupos de vértices (autores) da rede, pode-se observar a existência de vários subgrupos, os quais são chamados de componentes conectados. Componentes são subgrafos onde os vértices estão conectados por algum caminho possível, mas desconectados dos demais subgrafos [Wasserman and Faust 1994]. Em uma rede de coautoria a presença de mais de um componente na rede indica a existência de grupos que publicam de forma isolada dos demais autores da comunidade.

Na rede da ERES existem 31 componentes conectados, indicando que as relações de coautoria (colaboração científica) entre os autores desta comunidade são fragmentadas, com autores da mesma instituição escrevendo predominantemente com autores da mesma instituição, como pode ser observado pelas cores dos vértices da rede na Figura 2. Três (3) componentes são compostos por um único vértice, indicando que publicaram sozinhos. Nove (9) componentes são compostos por vértices (autores) de instituições diferentes, indicando que ao menos um (1) artigo foi escrito em coautoria com autor de outra instituição. Dezoito (18) componentes são compostos exclusivamente por autores da mesma instituição.

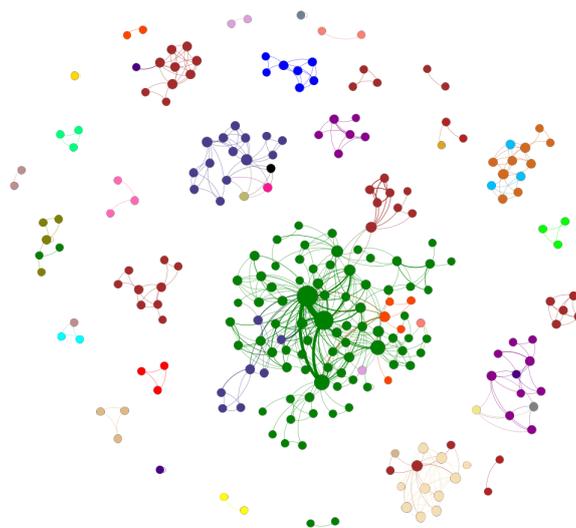


Figura 2. Rede de coautoria da ERES

Ao centro da rede destaca-se o componente com maior número de vértices, chamado de componente gigante. Componente gigante é o nome dado ao sub-grafo que contém o maior número de vértices conectados por algum caminho [Wasserman and Faust 1994]. Este componente é composto por 99 vértices (autores) cuja filiação é predominantemente da mesma instituição. A Figura 3 mostra este componente com os vértices rotulados e coloridos por instituição. Como pode ser observado existe um grande número de autores da UNIPAMPA, seguido por um número menor de autores da UTFPR, PUCRS, UNIJUÍ, UFSM e UFRGS.

Na Figura 3 estão destacados com um tamanho maior os vértices ou autores que mais escreveram em coautoria com outros autores conjuntamente com a frequência de coautoria, assim como destaca a frequência dessas relações de coautoria com um traço mais denso. É notória a frequência das relações de coautoria entre os autores Maicon Bernardino e Elder de Macedo Rodrigues (23 relações de coautoria), Maicon Bernardino e Fábio Basso (12 relações de coautoria) e Elder de Macedo Rodrigues e Fábio Basso (11 relações de coautoria), todos professores da UNIPAMPA.

Na perspectiva de análise individual dos vértices (autores) da rede, algumas métricas de centralidade foram calculadas. A Tabela 4 exibe os autores com as dez (10) maiores Centralidade de Grau Ponderada (CGP), conjuntamente com a Centralidade de Grau (CG) simples, a proporção entre CGP e CG, a instituição de filiação e a indicação do vínculo institucional: Professor (P) ou Aluno (A). A CGP é uma métrica que leva em conta o número de vértices adjacentes a um vértice, conjuntamente com a frequência das relações de adjacência, ou seja, no caso de uma rede de coautoria ela indica o número de autores com quem um autor escreveu acrescido da quantidade de vezes que escreveram juntos. Já a CG calcula apenas o número de vértices adjacentes, ou seja, com quantos autores um autor colaborou escrevendo artigos. Uma CG alta denota uma diversidade de relações de coautoria, enquanto que uma CGP alta indica uma frequência alta de relações de coautoria com o(s) mesmo(s) autor(es).

Na Tabela 4 é possível perceber que todos os autores possuem CGP maior que a

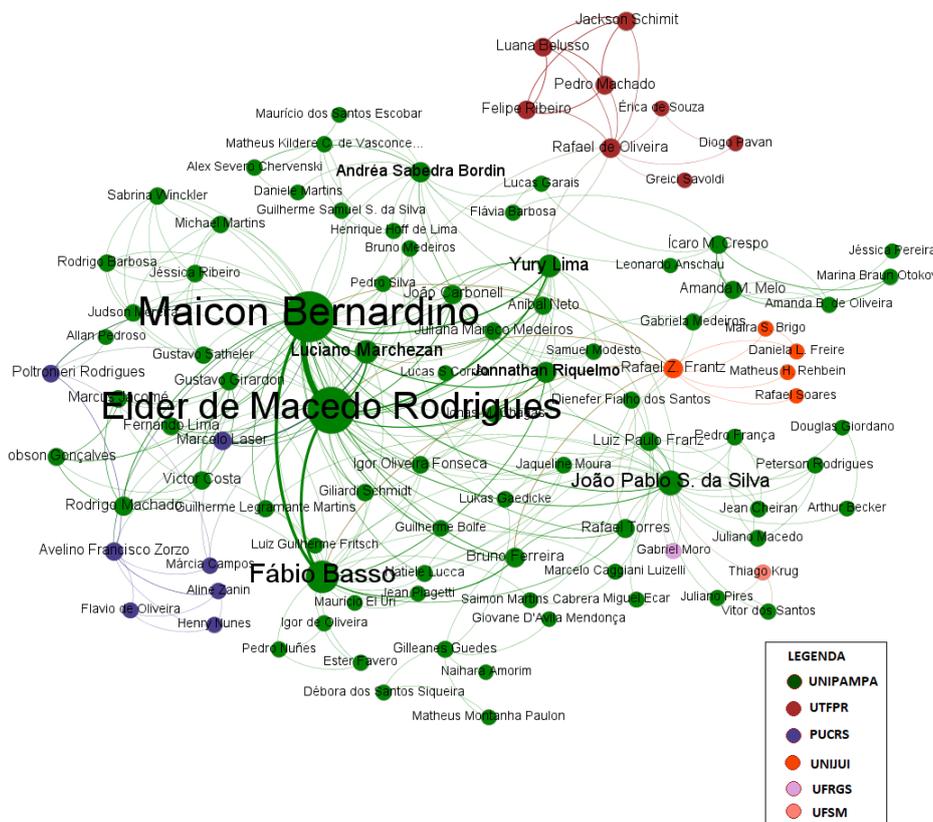


Figura 3. Componente gigante da rede

CG, indicando que escreveram ao menos mais de uma vez com o mesmo autor. Alguns autores se destacam pelo n mero expressivo de coautores, denotado pela CG, como   caso dos cinco primeiros autores. Desses, 4 s o professores da UNIPAMPA, onde est o vinculados a um curso de gradua  o em Engenharia de Software e tamb m a um mestrado profissional nesta mesma  rea. Os tr s primeiros, conjuntamente com o quinto autor, cujo vinculo   de aluno, pertencem ao mesmo grupo de pesquisa e tamb m possuem la os de coautoria frequentes, como pode ser visto na rede da Figura 3. Destaca-se que dos onze (11) autores da UNIPAMPA, apenas o quarto e o oitavo n o participam do mesmo grupo de pesquisa dos autores j  mencionados.

A coluna CGP/CG indica a propor  o de artigos escritos com os mesmos autores dentro do conjunto de coautores de um autor. Esta rela  o permitiu verificar os grupos de coautores com la os mais densos (frequentes) de colabora  o cient fica. Como pode ser observado ainda na Tabela 4 se destacam novamente os tr s primeiros autores, mas tamb m o d cimo autor, que escreveu com somente 5 coautores, mas de forma mais frequente com algum ou alguns deles. Por outro lado, destaca-se a autora da oitava posi  o com o menor CG, indicando que escreveu menos vezes com os seus coautores.

Outra m trica de centralidade abordada nesta pesquisa,   a Centralidade de Intermedia  o (CI). Em uma rede de coautoria, um autor que tem um alto grau de CI pode ser considerado uma “ponte” entre distintos grupos de pesquisa, uma vez que est 

Tabela 4. Autores com maior Centralidade de Grau Ponderada

	Pesquisador	CGP	CG	CGP/CG	Instituição
1	Maicon Bernardino (P)	111	45	2,5	UNIPAMPA
2	Elder de Macedo Rodrigues (P)	98	39	2,5	UNIPAMPA
3	Fábio Basso (P)	54	26	2,1	UNIPAMPA
4	João Pablo S. da Silva (P)	30	24	1,3	UNIPAMPA
5	Luciano Marchezan (A)	25	19	1,3	UNIPAMPA
6	Yury Lima (A)	24	12	2,0	UNIPAMPA
7	Jonnathan Riquelmo (A)	18	9	2,0	UNIPAMPA
8	Andréa Sabedra Bordin (P)	16	14	1,1	UNIPAMPA
9	Rafael de Oliveira (P)	14	10	1,4	UTFPR
10	Bruno Ferreira (A)	14	5	2,8	UNIPAMPA
11	Rafael Z. Frantz (P)	13	9	1,4	UNIJUÍ
12	João Carbonell (A)	13	9	1,4	UNIPAMPA
13	Rodrigo Machado (A)	13	7	1,9	UNIPAMPA

no caminho mais curto entre outros autores (vértices) [Souza et al. 2012]. A Tabela 5 mostra os autores com as maiores CI (com valores normalizados entre 0 e 1.0) na rede de coautoria da ERES. A análise desses valores é melhor contextualizada com a rede exibida na Figura 3, onde é possível visualizar o papel de "ponte" que esses autores exercem na rede. Observa-se que existem vários autores da UNIPAMPA presentes nesta tabela, o que indica a existência de grupos de autores distintos na mesma instituição que se conectam através desses autores com maiores CI. A existência desses grupos pode estar relacionada a interesses de pesquisa distintos na área de ES, dentre outros.

Tabela 5. Autores com maior Centralidade de Intermediação

Pesquisador	CI	Instituição
Maicon Bernardino (P)	0,41	UNIPAMPA
João Pablo S. da Silva (P)	0,24	UNIPAMPA
Elder de Macedo Rodrigues (P)	0,22	UNIPAMPA
Fábio Basso (P)	0,17	UNIPAMPA
Andréa Sabedra Bordin (P)	0,16	UNIPAMPA
Rafael Oliveira (P)	0,13	UTFPR
Rafael Z. Frantz (P)	0,08	UNIJUI
Gilleanes Guedes (P)	0,08	UNIPAMPA

No nível de análise global de rede, obteve-se um valor 0,018 para a Densidade (onde o valor máximo é 1,0), o que indica pouca exploração das possibilidades de colaboração entre os autores. Este valor baixo também é explicado em função do número de componentes (31 componentes) isolados. Já a Densidade do maior componente (componente gigante) é um pouco maior, de 0,058. Neste caso, como a maior parte dos autores desse componente possui vínculo com a mesma instituição (UNIPAMPA) pode-se considerar esse nível de conectividade baixo. Novamente, a existência de grupos com interesses diversos pode tornar difícil a colaboração em pesquisas dentro da mesma instituição. Por outro lado, a métrica Grau Médio da Rede, quando aplicada ao maior componente, resultou em 5,6, indicando que em média cada autor colaborou com esse número de coautores.

4. Conclusões

Esse trabalho apresentou uma análise bibliométrica dos artigos publicados na ERES, por meio de alguns indicadores de produtividade, impacto e colaboração científica. Os indicadores levantados nesta pesquisa fornecem subsídios para entender o comportamento da comunidade de pesquisadores de Engenharia de Software da região Sul do país, principalmente aquela envolvida com pesquisas desenvolvidas no âmbito da graduação, uma vez que 71% dos trabalhos aceitos nesse evento são dessa categoria.

Todos os indicadores revelam uma grande expressividade de autores da UNI-PAMPA, instituição de filiação de 34% dos autores do evento, dentre os quais estão os autores mais produtivos e aqueles com laços de colaboração mais frequentes. É provável que este cenário seja o resultado de um grande e permanente incentivo a publicar neste evento. Entender mais a fundo os aspectos motivadores e as estratégias utilizadas pelos professores orientadores das pesquisas para o engajamento dos estudantes seria importante, uma vez que outras instituições poderiam adotar estratégias semelhantes.

A análise da rede de colaboração científica dos pesquisadores indica a existência de vários grupos publicando de forma isolada. Sendo uma comunidade de pesquisa em processo de consolidação, parece ser interessante que mais laços sejam estabelecidos e fortificados. Existem cada vez mais evidências científicas que pesquisas com qualidade/impacto são feitas de forma colaborativa, em um processo de troca de saberes que tende a qualificar a pesquisa.

Como ameaça à validade deste trabalho destaca-se a normalização dos nomes dos autores, que ocorreu de forma manual e pode estar sujeito a erros. Todos os *datasets* utilizados nesta pesquisa estão disponibilizado de forma a garantir a proveniência e a reprodutibilidade dos resultados. Como trabalhos futuros, pretende-se aprofundar as análises bibliométricas, por exemplo, apresentando as temáticas ou áreas da engenharia de software mais abordadas nas publicações, além da realização de análise de citações dos artigos, que permite desvendar o fluxo de conhecimento desta área.

Referências

- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.
- Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 101(1):5200–5205.
- Rehn, C. and Kronman, U. (2008). *Bibliometric Handbook*, volume 1. Karolinska Institutet University Library.
- Souza, C. d., Barbastefano, R. G., and Lima, L. d. (2012). Redes de colaboração científica na área de química no Brasil: um estudo baseado nas coautorias dos artigos da revista química nova. *Química Nova*, 35(4):671–676.
- Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge university press.