

**INSTITUTO FEDERAL GOIANO – CAMPUS CERES  
BACHARELADO EM SISTEMAS DE INFORMAÇÃO  
THALIA SANTOS DE SANTANA**

**UM PANORAMA DAS PESQUISAS CIENTÍFICAS NA ÁREA DE CIÊNCIA DA  
COMPUTAÇÃO: UMA PERSPECTIVA FEMININA**

**CERES – GO  
2019**

**THALIA SANTOS DE SANTANA**

**UM PANORAMA DAS PESQUISAS CIENTÍFICAS NA ÁREA DE CIÊNCIA DA  
COMPUTAÇÃO: UMA PERSPECTIVA FEMININA**

Trabalho de conclusão de curso apresentado ao curso de Bacharelado em Sistemas de Informação do Campus Ceres do Instituto Federal Goiano, como requisito parcial para obtenção do título de Bacharela em Sistemas de Informação sob orientação do Prof. Me. Adriano Honorato Braga.

**CERES – GO  
2019**

Sistema desenvolvido pelo ICMC/USP  
Dados Internacionais de Catalogação na Publicação (CIP)  
**Sistema Integrado de Bibliotecas - Instituto Federal Goiano**

S232p Santana, Thalia Santos de  
UM PANORAMA DAS PESQUISAS CIENTÍFICAS NA ÁREA DE  
CIÊNCIA DA COMPUTAÇÃO: UMA PERSPECTIVA FEMININA /  
Thalia Santos de Santana; orientador Adriano Honorato  
Braga. -- Ceres, 2019.  
29 p.

Monografia ( em Bacharelado em Sistemas de  
Informação) -- Instituto Federal Goiano, Campus  
Ceres, 2019.

1. Cienciometria. 2. Computer Science Research.  
3. Mulheres na Computação. 4. Produção científica. 5.  
CSBC. I. Braga, Adriano Honorato, orient. II. Título.



**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR PRODUÇÕES TÉCNICO-CIENTÍFICAS NO REPOSITÓRIO INSTITUCIONAL DO IF GOIANO**

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia Goiano, a disponibilizar gratuitamente o documento no Repositório Institucional do IF Goiano (RIIF Goiano), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IF Goiano.

**Identificação da Produção Técnico-Científica**

- |  |   |
|--|---|
| <input type="checkbox"/> Tese                                  | <input type="checkbox"/> Artigo Científico              |
| <input type="checkbox"/> Dissertação                           | <input type="checkbox"/> Capítulo de Livro              |
| <input type="checkbox"/> Monografia - Especialização           | <input type="checkbox"/> Livro                          |
| <input checked="" type="checkbox"/> TCC - Graduação            | <input type="checkbox"/> Trabalho Apresentado em Evento |
| <input type="checkbox"/> Produto Técnico e Educacional - Tipo: |   |

Nome Completo do Autor: Thalia Santos de Santana

Matrícula: 2016103202030141

Título do Trabalho: UM PANORAMA DAS PESQUISAS CIENTÍFICAS NA ÁREA DE CIÊNCIA DA COMPUTAÇÃO: UMA PERSPECTIVA FEMININA

**Restrições de Acesso ao Documento**

Documento confidencial:  Não  Sim, justifique: \_\_\_\_\_

Informe a data que poderá ser disponibilizado no RIIF Goiano: 18/12/19

O documento está sujeito a registro de patente?  Sim  Não

O documento pode vir a ser publicado como livro?  Sim  Não

**DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA**

O/A referido/a autor/a declara que:

- o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia Goiano os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia Goiano.

Goias 18/12/19  
Local Data

Thalia Santos de Santana  
Assinatura do Autor e/ou Detentor dos Direitos Autorais

Ciente e de acordo:


Adriano Honorato Braga  
Assinatura do(a) orientador(a)

## ATA DE DEFESA DE TRABALHO DE CURSO

Aos vinte e um dias do mês de novembro do ano de dois mil e dezenove, realizou-se a defesa de Trabalho de Curso da acadêmica **Thalia Santos de Santana**, do Curso de Bacharelado em Sistemas de Informação, matrícula 2016103202030141, cujo título é “**Um panorama das pesquisas científicas na área de Ciência da Computação: uma perspectiva feminina**”. A defesa iniciou-se às 21 horas e 00 minutos, finalizando-se às 22 horas e 15 minutos. A banca examinadora considerou o trabalho aprovado com média 10,0 no trabalho escrito, média 10,0 no trabalho oral, apresentando assim média aritmética final 10,0 de pontos, estando a estudante apta para fins de conclusão do Trabalho de Curso.

Após atender às considerações da banca e respeitando o prazo disposto em calendário acadêmico, a estudante deverá fazer a submissão da versão corrigida em formato digital (.pdf) no Repositório Institucional do IF Goiano – RIIF, acompanhado do Termo Ciência e Autorização Eletrônico (TCAE), devidamente assinado pelo autor e orientador.

Os integrantes da banca examinadora assinam a presente.



*Assinatura Presidente da Banca*

Adriano Honorato Braga



*Assinatura Membro 1 Banca Examinadora*

Aliny Karla da Cunha



*Assinatura Membro 2 Banca Examinadora*

Livia Mancine Coelho de Campos

Para as Meninas Digitais de todo o mundo: Sonhem alto, acreditem em si pr3prias, sejam perseverantes e nunca se esqueçam, que voc3s podem ser o que quiserem.

## **AGRADECIMENTOS**

Agradeço a Deus, pelo dom da vida e por seus privilégios. Privilégios de poder perpassar cada um dos inesquecíveis momentos de minha graduação com muitos sorrisos (mesmo que de desespero) e sempre ao lado de pessoas que admiro. Acredito muito que tudo há um motivo e sou eternamente grata pelos encontros que me levaram à ser parte integrante da 1ª Turma do Curso de Bacharelado em Sistemas de Informação do Campus Ceres do IF Goiano, meu lar e a origem de todo esse sonho. Tudo que sou, devo a esta amada Instituição e nunca serei capaz de agradecer ou retribuir o quanto transformou minha vida.

À minha mãe Maria, meu pai Orlando e amados irmãos Thaila, Orlando Filho e Rangel. Minha querida família, que com seu jeito único, acreditou em mim e em meus sonhos, fazendo o que estava a seu alcance para que eu pudesse estudar e fazer todos os incontáveis projetos e atividades pelas quais sou apaixonada. Além disso, agradeço ao meu namorado Lucas, que me incentivou e esteve se orgulhando de minhas realizações, mesmo que isso significasse minha ausência.

Agradecimentos especiais ao meu amigo e orientador, Adriano Honorato Braga. Nem mesmo as consultas mais complexas poderiam sequer responder o quanto sou grata por ter sido sua aluna e orientanda. Foram mais que aprendizados científicos, foram para toda uma vida. Obrigada por ser sempre quem acredita em mim (mesmo que eu ainda não saiba disso), por todas as suas palavras de calma em meio às minhas angústias e por ver luz, nas pessoas, no universo e inclusive, em mim mesma. São pessoas como você, que sem dúvidas, fazem o mundo um lugar melhor e tenho imenso orgulho de dizer que tive a ti como orientador do meu trabalho de curso.

À minha finalmente professora, Ramayane Bonacin Braga, que foi a primeira pessoa com quem conversei sobre este tema, quando nem eu mesma cogitaria em sonhá-lo. Não sabíamos, mas sinto que naquele dia meu coração já compreendia que era isso que iria pesquisar e foi, felizmente, culpa sua. O meu mais sincero obrigada, além de todos os desabafos que partilhamos e que fizeram com que eu chegasse nesta reta final tão contente e grata. Gratidão também por ter partilhado

toda essa jornada ao lado da maravilhosa equipe Meninas Digitais no Cerrado, projeto que levo e levarei comigo para todo o sempre.

Gostaria de deixar também um agradecimento à todas e todos os professores que fizeram parte deste caminho. Meu encanto com a educação paira sobre todas as aulas e conhecimento que compartilhamos juntos. Em nome do professor Marcos Moraes, que também conheceu meus medos e incertezas neste estudo (mas que não verá “ao vivo” o desfecho deste trabalho), agradeço pelos aprendizados e reafirmo, que é a educação, a ferramenta mais poderosa para transformamos o mundo.

E quase em último mas não menos importante, aos amigos que fiz durante esta trajetória. Sem vocês não seria tão bonito e muito menos, tão emocionante! Gostaria de citar todos, mas mesmo correndo o risco de pecar e esquecer alguém, agradeço principalmente aos colegas Alexandre, Lara, Gustavo, Luciano e Ianka, sendo os três primeiros com quem eu sempre me achava no direito de debater os problemas do meu trabalho em busca de respostas e os demais, quem me faziam rir diante de todas as adversidades. Agradeço por não terem me bloqueado, e claro, pelas valiosas contribuições.

Afinal, à todos vocês, que ouviram, ouviram e ouviram falar desse famoso TC. Este trabalho é nosso e sou muito feliz por saber que de algum jeito estivemos conectados. Obrigada pela empatia e sororidade partilhada. Avante!



“Quanto mais estudo, mais sinto que minha mente  
nisso é insaciável”.

Ada Augusta King, Condessa de Lovelace

## RESUMO

É notória a desproporcionalidade de gênero em campos ligados à tecnologia da informação. Mesmo em um cenário de esforço para valorização de modelos femininos de êxito em computação, as mulheres são minoria no mundo acadêmico. Com reflexos de vieses históricos, ocasionou-se também impacto no desenvolvimento da pesquisa científica na área. A cienciometria é o ramo do conhecimento capaz de estabelecer índices para a produção intelectual, além de contribuir para verificação de lacunas e tendências de pesquisas. Neste contexto, ela pode ser utilizada inclusive para caracterização da diversidade de gênero entre seus cientistas na pesquisa em ciência da computação (CSR – em inglês, *computer science research*). O presente trabalho realizou um estudo acerca da presença feminina em eventos do Congresso da Sociedade Brasileira da Computação (CSBC), o qual é considerado uma das conferências de maior relevância em computação no país. Os dados foram obtidos por intermédio de *web scraping* (do inglês – raspagem de dados), compreendendo o período de 2017 a 2019, abstraindo as seguintes variáveis: nome dos autores, título dos trabalhos, ano de publicação e eventos. Por meio da base construída, os resultados apontaram predominância do público masculino nos eventos do CSBC, entretanto, quando avaliado a presença feminina mediante os artigos publicados, as mulheres estão contidas em quase 60% dos trabalhos, mesmo constituindo apenas 31% do quantitativo de autores, colocando em evidência a relevância do papel de autoras na construção e progresso científico em computação.

**Palavras-chave:** Cienciometria. *Computer Science Research*. Mulheres na Computação. Produção científica. CSBC.

## ABSTRACT

Gender disproportionality in fields related to Information Technology is notorious. Even in a scenario of endeavor to value female models of successful computing, as women are minimal in the academic world. With historical reflections, occasionally also affects the development of scientific research in the area. Scientometrics is the field of knowledge capable of recognize indexes of intellectual production, besides contributing to the verification of gaps and research trends. In this context, it can even be used to characterize gender diversity among its scientists in computer science research (CSR). The present work was carried out in a study about the female presence in events of the Congress of the Brazilian Computer Society (CSBC), that is considered one of the major concerns of relevance in the country. Information were tracked by data scraping, covering the period from 2017 to 2019, presenting the following variables: name of the authors, title of the paper, year of publication and events. Based on the constructed base, the results indicate a predominance of the male public at CSBC events. However, when the female presence was evaluated using the published articles, as women are contained in almost 60% of the articles, even constituting only 31% of the quantitative by highlighting the relevance of the role of authors in the construction and scientific progress in computing.

**Keywords:** Sciencimetric. Computer Science Research. Women in Computer Science. Scientific production. CSBC.

## LISTA DE ILUSTRAÇÕES

- Figura 1 – Gráfico de barras quanto ao gênero de autoras/es e ordem de autoria.. 19
- Figura 2 – Gráfico de barras do quantitativo de publicações por ano e evento.....20
- Figura 3 – Gráfico de barras do percentual de autoras por ano e evento.....22

## LISTA DE TABELAS

Tabela 1 – Busca primária e quantitativo de resultados.....	6
Tabela 2 – Trabalhos referência e <i>strings</i> com principais palavras.....	7
Tabela 3 – Quantidade de publicações encontradas em cada base pelo título.....	7
Tabela 4 – Eventos do CSBC por tipo, ano e edição.....	14
Tabela 5 – Eventos, quantitativo de autoras/es por gênero e razão F/M.....	21

## SUMÁRIO

INTRODUÇÃO.....	1
REVISÃO DE LITERATURA.....	4
MATERIAL E MÉTODOS.....	10
RESULTADOS E DISCUSSÃO.....	19
CONSIDERAÇÕES FINAIS.....	23
REFERÊNCIAS.....	26

## INTRODUÇÃO

Análises sobre diversidade e desequilíbrio de gênero na pesquisa e a consequente participação de mulheres na popularização científica, são uma importante questão para a equidade e entendimento sobre a sub-representação de minorias em Ciência da Computação (CC) (AGARWAL; MITTAL; SUREKA, 2017). Considerando as produções na área, somente cerca de 21% dos pesquisadores que são ativos em conferências de renome são mulheres, contrapondo-se aos 79% homens, dentre o indexador *DBLP Computer Science Bibliography*<sup>1</sup>, um repositório de grande relevância na computação internacionalmente (AGARWAL et al., 2016). Ainda, conforme os autores, o Brasil tem um número maior de mulheres contribuintes em conferências de alto nível (37%), quando comparado com Japão e Alemanha, tornando o desequilíbrio de gênero inespecífico a países em desenvolvimento.

Avaliando-se todas as áreas de produção científica, as mulheres brasileiras são responsáveis por 72% dos artigos publicados no país. O resultado é proveniente de estudos recentes da Organização dos Estados Ibero-americanos (OEI), onde Albornoz et al. (2018) avaliaram entre 2014 e 2017, as produções em que pesquisadoras se apresentassem seja como primeira autora, seja como coautora. Apesar desse relatório destacar o Brasil com maior participação de mulheres quando comparado com os 18 países pesquisados, áreas como engenharias possuem apenas 32% de mulheres nas publicações, enquanto na medicina são 56%. Contudo, mesmo assinando a maior parte dos artigos, o quantitativo total de pesquisadoras (49%) é menor do que do público masculino (ALBORNOZ et al., 2018).

Nesse contexto brasileiro, é perceptível a baixa proporção feminina na comunidade científica em CC (MONARD; FORTES, 2013). Os índices possuem influência de elementos históricos, refletidos muitas vezes pela invisibilização de mulheres na informática. Pioneiras como Ada Lovelace, a primeira pessoa a construir um algoritmo e considerada a primeira programadora; Grace Hopper, desenvolvedora do primeiro compilador e colaboradora da criação da linguagem de

---

1 <https://dblp.uni-trier.de/>

programação COBOL (*Common Business Oriented Language*) e as garotas do ENIAC (*Electronic Numerical Integrator and Computer*), programadoras de um dos primeiros computadores digitais/eletrônicos do mundo; são exemplos da presença feminina na computação que contribuíram para sua concepção enquanto ciência (SCHWARTZ et al., 2006).

Apesar disso, modelos femininos do passado frequentemente não são conhecidos pelas atuais gerações (HUSSAIN et al., 2015). Ao mesmo tempo, a pequena taxa de estudantes mulheres em cursos ligados a tecnologia corrobora para a ausência das mesmas na pesquisa científica. Os dados mais atuais da Sociedade Brasileira de Computação (SBC), demonstram o percentual de 13,95% discentes do gênero feminino no ensino superior em cursos relacionados a computação no Brasil, totalizando 44.812 mulheres e 276.308 homens matriculados no ano de 2017, de acordo com informações do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e censo do referido ano (SBC, 2017).

Arruda et al. (2009) ao avaliar o currículo Lattes<sup>2</sup> de docentes de pós-graduação em CC no país, verificaram a existência, em novembro de 2006, de 201 pessoas do sexo feminino e 685 do sexo masculino dentre os pesquisadores, totalizando 23% mulheres – sendo um valor muito próximo ao avaliado no estudo de Agarwal et al. (2016). Além disso, por temáticas de pesquisa, foi verificado estatisticamente que as mulheres têm proporção maior em áreas que possuem em comum o componente humano, como inteligência artificial, computadores na educação e interface homem-máquina. Entretanto, afastam-se daquelas de componente tecnológico, a exemplo de arquitetura de computadores e *hardware*, bem como redes e sistemas distribuídos. Inclusive, nestes campos de pesquisa mais “femininos”, mulheres correspondem apenas a 29,7% e são mais produtivas que os homens, com quantitativo de 5,3 publicações de revistas por pesquisadores e 6,0 por pesquisadoras (ARRUDA et al., 2009).

Sendo assim, o presente trabalho objetivou diagnosticar produções científicas publicadas na área de CC e realizar uma análise quantitativa por meio dos dados catalogados sob a perspectiva de gênero. A pesquisa buscou averiguar a existência de possíveis brechas por eventos entre a presença de autoras e autores, além de

---

2 <http://lattes.cnpq.br/>



colaborar com os estudos brasileiros acerca do cenário de diversidade de gênero na academia, apresentando comparações entre eventos de uma das maiores conferências científicas no ramo. Os dados foram minerados na *Web* por meio das publicações dos anais das três últimas edições (2017-2019) do Congresso da Sociedade Brasileira de Computação (CSBC). Além da introdução, o estudo apresenta as sessões de revisão literatura, materiais e métodos empregados, resultados e discussões e considerações finais.

## REVISÃO DE LITERATURA

Em meio a ciência, costumeiramente inventos e demais produções de grande destaque omitiram a presença feminina como promotoras e colaboras de tais descobertas. Na cultura científica, hábitos foram desenvolvidos em uma oposição à participação de mulheres (SCHIEBINGER, 2008), a qual corroborou para um apagamento histórico, mesmo com significativas contribuições, a exemplo da Química, Física, Biologia ou mesmo, na tecnologia.

Para Conceição e Teixeira (2018, p. 1), foi construído um “território hegemonicamente masculino” em uma dicotomia (CORDEIRO, 2013): aos homens, campos como a política, negócios e ciência, e para as mulheres, atividades domésticas, a enfermagem e a educação. Há cerca de cinco décadas surgiram as primeiras publicações científicas em relação a problematização da escassez de mulheres em carreiras de Ciência & Tecnologia (C&T), por meio das quais Rossi (1965) questionou a participação feminina na área, desde políticas de incentivo até relações sociais e estigmas para a realização da mulher, a exemplo do casamento, maternidade e do próprio trabalho.

Estudos recentes de Codeço e Dias (2018) apontam que apenas um terço de discentes em Ciências, Matemática e Tecnologia mundialmente, são do sexo feminino, sendo ainda mais evidente a desproporção quando analisado posições acadêmicas avançadas. Khan e Luxton-Reilly (2016) na Universidade de Auckland (Nova Zelândia), demonstraram que entre 2005 e 2015, 49% dos bacharéis em ciências foram mulheres, mas apesar de ser um bom número no cenário geral, ao verificar-se no espectro da computação com as mesmas variáveis de pesquisa, somente 14% correspondiam ao público feminino.

Monardes e Fortes (2013) realizaram análises de estudantes concluintes em computação durante o período de 2001 a 2011 no Brasil. Nos anos iniciais de abrangência da pesquisa, as mulheres totalizavam proporção entre o público feminino e masculino de 47% e nos anos finais do período, somente 21%, indicando uma tendência de diminuição de formandas na área. Já quando analisado o gênero dentre as titulações de mestres e doutores, dados de 2016 do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), evidenciam que o quantitativo de

homens em CC (cerca de 75%) ultrapassa o de mulheres em ambos os níveis (ALMEIDA; ZANLORENSSI, 2018).

Por meio desse contexto, Cunha et al. (2014) correlacionam o desequilíbrio de gênero em carreiras científicas ao processo educacional. Enquanto meninos são estimulados por meio de computadores, carros, máquinas e ferramentas, meninas são tendenciadas para o cuidado, como na educação, saúde e bem-estar, implicando assim em seus interesses futuros. Os estudos dos autores ainda incitam discussões sobre estereótipos sociais, que acabam promovendo a concepção errônea de que determinadas áreas de atuação relacionam-se a aptidões diferentes para mulheres e homens. Resultados de pesquisas com estudantes brasileiros de ensino médio indicaram que, em maioria e independente do sexo, não há a intenção de adentrarem em profissões para tornarem-se cientistas. O agravante consiste em uma análise realizada por gênero, que considerando uma média entre todas as regiões brasileiras, apenas 18,6% do total de meninas pretendem ser cientistas – número ainda menor que o de meninos (26,9%) (CUNHA et al., 2014).

Assim, além de estratégias de ingresso, permanência e motivação de mulheres, levantamentos acerca de publicações científicas vêm sendo realizados a fim de entender questões de desigualdade de gênero, inclusive por intermédio da participação de mulheres como autoras de trabalhos de pesquisa (AGARWAL et al., 2016). Apesar de várias barreiras enfrentadas para promoção do sucesso feminino em ramos com estereótipos masculinizados, mulheres são bem-sucedidas em campos como tecnologia (SMEDING, 2012), com reflexos também em suas produções científicas.

Todavia, verifica-se ainda a elevada sub-representação feminina gerada tanto por uma discriminação direta quanto por meio de comportamentos subconscientes que perpetuam estigmas da área (BEAUBOUF; ZHANG, 2011). O mesmo ocorre conforme o aumento da escala de prestígio científico, definido como “*leaking pipeline*” (ARRUDA et al., 2009) dentro de Ciência, Tecnologia, Engenharia e Matemática (STEM) como também no enfoque da pesquisa em ciência da computação (CSR – em inglês, *computer science research*), terminologia utilizada para distinguir estudos com enfoque na pesquisa científica na área.

Para avaliação de indicadores inerentes à publicações acadêmicas, é comum o uso de cienciometria e bibliometria como metodologia de extração de dados mensuráveis na literatura para análises estatísticas de trabalhos científicos (AGARWAL et al., 2016). A cienciometria é definida como o estudo de aspectos quantitativos da ciência (TAGUE-SUTCLIFFE, 1992), possibilitando realizar a avaliação e quantificação do progresso científico. Além disso, é capaz de prover subsídios na academia e sociedade em geral quanto aos tópicos de interesse, contribuindo para a verificação de lacunas, inclusive quanto à diversidade de gênero.

Diante dos trabalhos encontrados a partir do cenário de mulheres pesquisadoras na ciência de forma geral e em computação, considerou-se a realização de uma revisão sistemática. Foi feita uma busca por publicações considerando as bases científicas ACM *Digital Library*<sup>3</sup>, IEEE *Xplore Library*<sup>4</sup> e SciELO<sup>5</sup> (Tabela 1), usando *strings* de busca em abril de 2019, por meio do título dos trabalhos como também realizado por Louzada et al. (2014).

**Tabela 1 – Busca primária e quantitativo de resultados.**

Base	String buscada	Quantidade
ACM <i>Digital Library</i>	<i>acmdlTitle:((woman +OR +women +OR +gender) +AND +("computer science" +research))</i>	7
IEEE <i>Xplore Library</i>	<i>((("Document Title":woman) OR "Document Title":women) OR "Document Title":gender) AND "Document Title":"computer science") AND "Document Title":research)</i>	2
SciELO	<i>(ti:(woman)) OR (ti:(women)) OR (ti:(gender)) AND (ti:("computer science")) AND (ti:(research))</i>	0

Fonte: Própria (2019).

Dentre os trabalhos retornados, foram investigados títulos, resumos e por fim, texto completo. Na ACM *Digital Library*, os artigos possuíam relação com a sub-representação feminina, pesquisas para o interesse na área e recrutamento, encontrando um trabalho de relevância em CSR. Quanto à IEEE *Xplore*, ambas as publicações apresentaram pesquisas com mulheres universitárias com estratégias para envolver mulheres na computação, inclusive em salas de aula, mas não no

3 <https://dl.acm.org/>

4 <https://ieeexplore.ieee.org/Xplore/home.jsp>

5 <https://scielo.org/>

âmbito de pesquisa científica, tornando a *string* de busca muito específica.

Desta forma, dada a inexistência na literatura de diversos trabalhos de relevância sobre o objeto de estudo, passou-se a considerar Arruda et al. (2009) e Agarwal et al. (2016) como artigos de referência. Inclusive, Arruda et al. (2009) em sua publicação citam o tema como uma lacuna de pesquisa para a área. Sendo assim, para continuidade deste estudo, primeiramente foi criada uma nuvem de palavras para cada um dos trabalhos com o objetivo de visualizar as três principais palavras citadas com maior frequência no texto, conforme apresentado na Tabela 2.

**Tabela 2 – Trabalhos referência e *strings* com principais palavras.**

<b>Trabalho</b>	<b>Nuvem de Palavras</b>
Arruda et al. (2009)	<i>science</i> AND <i>research</i> AND <i>computer</i>
Agarwal et al. (2016)	<i>women</i> AND <i>authors</i> AND <i>CSR</i>

Fonte: Própria (2019).

Assim, foram geradas novas *strings* de busca com base nos trabalhos de referência. O termo CSR foi buscado em abreviação e sem abreviação (“*computer science research*”), visto que poderia ocorrer ambiguidade nas buscas pela sigla ao possuir também outro significado inclusive na computação. Não houve nenhum outro tipo de restrição, a não ser pelo título dos trabalhos, conforme resultado na Tabela 3.

**Tabela 3 – Quantidade de publicações encontradas em cada base pelo título.**

<b>Base</b>	<b><i>Strings</i> buscadas</b>	<b>Quantidade</b>
ACM Digital Library	<i>acmdlTitle:(+science +research +computer)</i>	805
	<i>acmdlTitle:(+women +authors +CSR)</i>	0
	<i>((("Document Title":science) AND "Document Title":research) AND "Document Title":computer)</i>	80
IEEE Xplore Library *	<i>((("Document Title":women) AND "Document Title":authors) AND "Document Title":CSR)</i>	0
	<i>(ti:(science)) AND (ti:(research)) AND (ti:(computer))</i>	3
SciELO **	<i>(ti:(women)) AND (ti:(authors)) AND (ti:(CSR))</i>	0

Fonte: Própria (2019).

\* A base IEEE Xplore disponibiliza campos de restrição de busca como “*documente title*” e “*publication title*”. Optou-se para a pesquisa o uso do campo “*document title*”, conforme Louzada et al. (2014).

\*\* Na base SciELO as buscas também foram feitas com as palavras em português, mas foram obtidos os mesmos resultados.

Como critérios de inclusão, em meio a leitura dos resumos e dos trabalhos completos, delimitou-se: i) trabalhos com abordagem de mulheres pesquisadoras em computação e tecnologia, ii) trabalhos ligados a alguma relação com gênero ou sexo em pesquisas e/ou produção científica em computação, iii) trabalhos pertinentes a evolução de subáreas da computação. Os demais temas que não tiveram relação explícita com os critérios de inclusão foram considerados como critérios de exclusão.

Foi perceptível que a combinação das palavras mais usadas no texto de Agarwal et al. (2016) não foram relevantes para nenhuma base. Já no primeiro conjunto de descritores, dentre os 805 trabalhos encontrados na ACM *Digital Library*, verificou-se que para essa base, a *string* usada mostrou-se muito abrangente. Retornaram temas como pesquisa em educação na CC bem como experimentações de forma geral, mas raramente de mulheres pesquisadoras e produtoras de publicações científicas no ramo. Também retornaram estudos duplicados, sendo escolhidos apenas 17 trabalhos após consideração dos critérios de inclusão e exclusão. Quanto a IEEE foram pré-selecionados 13 estudos e nenhum na SciELO.

Por fim, sete publicações tiveram alguma relação com o objetivo desta pesquisa. Na ACM *Digital Library* três estudos foram considerados, sendo um deles artigo referência e resultado da busca primária, enquanto na IEEE *Xplore Library*, selecionou-se quatro trabalhos após a eliminação de produções duplicadas. Como método adicional, utilizou-se o levantamento bibliográfico “bola de neve” (*snowball*), por meio da leitura e seleção das referências dos artigos definidos como relevantes, corroborando para o arcabouço teórico de pesquisas analisadas.

Banshal et al. (2014), Banshal et al. (2015), Effendy e Yap (2017), Noosrikong, Ngamsuriyaroj e Ayudhya (2017) e Yang et al. (2017) investigaram dados de produções científicas com avaliações quanto a abrangência em subáreas da computação, mediante identificação de campos de publicação mais prolíficos. Contudo, foi verificado dentre os estudos selecionados, que apenas Agarwal, Mittal e Sureka (2017) abrangeram correlações de gênero ao dissertar acerca da disparidade e sub-representação de cientistas minoritários, além de como já mencionado em informática por Agarwal et al. (2016).

Em meio ao referencial dos artigos de relevância, cinco trabalhos foram considerados. Destes, apenas dois realizaram avaliações em áreas de

conhecimento específicas, de modo que Gallivan e Benbunan-Fich (2006) e Vela, Caceres e Cavero (2012) examinaram a produtividade entre gêneros, respectivamente, em Sistemas de Informação e Engenharia de *Software*. Gurer e Camp (2002) foram responsáveis pela única revisão de literatura sistemática encontrada na busca, a qual fez uso do banco de dados *on-line* da ACM *Women in Computing*<sup>6</sup> (ACM-W). Cohoon, Nigai e Kaye (2011) aproximam-se dos autores anteriores ao trabalhar dentro da ACM *Digital Library*, enquanto Cavero et al. (2015) avaliaram outro grande repositório de publicações em computação, DBLP.

Deste modo, apesar de existirem diversas pesquisas que avaliam áreas/eventos específicos e/ou de modo geral, CSR, em determinados países, não foram identificados estudos que analisam mulheres autoras em trabalhos científicos no Brasil com enfoque em eventos distintos, ligados a uma única conferência. O CSBC é um congresso referenciado e com uma parcela relevante da comunidade de CC ativa, contudo, até então não foram tecidas análises de forma abrangente aos eventos ligados a esta conferência, ainda mais ao tratar-se de uma das maiores realizações da principal instituição civil brasileira representativa neste campo (RIBEIRO et al., 2019).

---

6 <https://women.acm.org/>

## MATERIAL E MÉTODOS

A presente pesquisa é do tipo exploratória, por meio de abordagem quantitativa. Em primeiro momento, objetivou-se realizar as análises por meio do JEMS, sistema oficial da SBC para submissão de artigos, responsável pelo gerenciamento de todo o processo relacionado ao envio, revisão e notificação de aceites em conferências acadêmicas (JEMS, 2019). O mesmo é disponibilizado aos seus eventos e assim, contém uma base de dados com elevado número de informações inerentes a autoras/es, artigos e conferências (CONTESSA; PALAZZO, 2006).

Considerando que a SBC é a maior sociedade de computação de toda a América do Sul (CSBC, 2019), ela reflete-se como a principal realizadora e apoiadora da execução de eventos técnico-científicos brasileiros na área. A exemplo disso, um de seus feitos mais reconhecidos trata do CSBC, o qual é realizado desde o ano de 1980 e encontra-se em sua 39ª edição, oportunizando troca de experiências no âmbito científico, acadêmico e profissional em CC, com abrangência nacional e internacional (SBC, 2019a). O CSBC é o maior em computação em nível acadêmico na América Latina, além de possuir prestígio junto a comunidades científicas, bem como instituições e órgãos de fomento.

Sendo assim, o congresso consolidou-se estruturado por diversos eventos base e satélites. Atualmente, são reunidos nove eventos da parte fixa e demais eventos esporádicos, que acontecem dentro da própria conferência após submissão e seleção de propostas pelo Comitê Gestor – que também recebe projetos para organização e sede da conferência anualmente. Dentre os eventos base, são eles: Curso de Qualidade (CQ), Seminário de Computação e Mercado (COMPUTEC), Concurso de Teses e Dissertações (CTD), Concurso de Trabalhos de Iniciação Científica (CTIC), Jornadas de Atualização em Informática (JAI), Seminário de Computação na Universidade (SECOMU), Seminário Integrado de *Software* e *Hardware* (SEMISH), Workshop sobre Educação em Computação (WEI) e o *Women in Information Technology* (WIT) (SBC, 2019a).

No Brasil, ressalta-se que para a CC, as avaliações de produção científica Qualis/Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)



trabalha com publicações de periódicos e conferências. Mediante análises e critérios adotados, como a existência de conferências de prestígio consolidadas em CC, artigos completos submetidos e aceitos em eventos de relevância são tão importantes para divulgação científica, quanto os artigos publicados em periódicos indexados de renome, visto passar por um elevado rigor de avaliação por pares também em anais de eventos, inclusive sendo passíveis de avaliação por parâmetros similares à de periódicos (CAPES, 2016).

Ademais, conferências como o CSBC, garantem disseminação rápida dos resultados de pesquisas, visto que a computação perpassa por uma defasagem acelerada, além de ser uma área de conhecimento ainda considerada recente. Além disso, ao possuir um evento brasileiro robusto, apresenta impacto inclusive sobre o índice Qualis, o qual é responsável por prover a estratificação da qualidade de produção intelectual, sendo este congresso angariador de vários eventos com tal classificação (CAPES, 2017).

Após solicitação à SBC para uso desta base de artigos via plataforma JEMS, foi informado que a mesma não dispõe de política de privacidade que permita exportar os dados de suas conferências e ser uma das ações futuras da Diretoria de Eventos e Comissões Especiais, agilizando assim quando disponível, no processo de descoberta de informações. Deste modo, não foi possível trabalhar com os metadados diretamente do sistema JEMS, sendo necessário um novo método para obtenção dos dados das publicações na área no país.

Assim sendo, é da importância do objeto de estudo desta pesquisa os dados da conferência, optando-se por obter as publicações de trabalhos presentes nas páginas eletrônicas do CSBC, como material para análises cienciométricas. Foram considerados artigos publicados entre os anos de 2017 e 2019 por meio de procedimentos de raspagem de dados, comumente conhecidos como *Web Scraping*. Tal técnica permite a aquisição automatizada de dados de páginas *web*, sem a necessidade de cópia manualmente. Assim, informações não estruturadas, geralmente em formato HTML (*HyperText Markup Language* – em português, Linguagem de Marcação de Hipertexto), são transformadas em dados estruturados para análise em bancos de dados ou planilhas eletrônicas, inclusive localmente (VARGIU; URRU, 2013).

A partir de informações disponíveis nos endereços eletrônicos das três últimas edições do CSBC, realizou-se a raspagem e coleta de dados com a linguagem de programação Python<sup>7</sup>, versão 2.7.16, no mês de agosto de 2019. Foi analisada a estrutura das páginas antes de executar o Algoritmo 1, por meio de ferramentas de inspeção presentes em navegadores, como no caso, o Google Chrome. Sabendo-se dos blocos de conteúdo que continham informações dos eventos, foi possível recuperar dados da internet que exprimissem às publicações do CSBC nestes últimos anos, fazendo uso de um *script* com auxílio das bibliotecas *Pandas*, *BeautifulSoup* e *URLOpen*.

A biblioteca *BeautifulSoup* é especificamente utilizada para coleta de dados via *Web Scraping*. Dado o formato das informações *web*, edições anteriores ao ano de 2017 não entraram para raspagem, visto que tais dados não se encontravam em formato HTML ou haviam poucos eventos com trabalhos informados, impossibilitando a coleta automática.

A cada ano, o CSBC disponibiliza os anais do congresso em páginas diferentes, tornando assim necessário um *script* do tipo *scraper* para cada uma destas, dentre os quais foram informados os endereços eletrônicos de cada evento e seus trabalhos, além do objeto HTML a ser mapeado, provendo como resultado um arquivo no formato CSV (*Comma-separated-values* – em português, valores separados por vírgula), contendo evento e sua edição, título e autoras/es. Assim, houve uma unificação dos arquivos e conseqüente, normalização de nomes e siglas dos eventos.

Quanto a seleção dos eventos, foram eliminados em primeira instância aqueles que não possuíram submissão e aceite de trabalho, bem como aqueles que produziram itens diferentes de artigos como resultado, como a JAI, a qual publica um livro ao final da conferência. Desta forma, em 2017 foram retornados 18 eventos com algum tipo de produção científica no CSBC, seja artigos curtos, seja artigos completos. Em 2018, 19 eventos e em 2019, 16 eventos. Além disso, foi realizada a limpeza dos dados coletados, eliminando registros que contiveram prefácios de abertura, limitando-se somente às publicações em específico e seus autores/as.

---

<sup>7</sup> <https://www.python.org/>

## Algoritmo 1 – Web Scraping das publicações de 2017.

```
1  from pandas import pandas as pd
2  from bs4 import BeautifulSoup
3  from urllib import urlopen
4  import csv
5  titulos = []
6  autores = []
7  eventos = []
8  links = ['http://csbc2017.mackenzie.br/anais/eventos/2-etc',
9  'http://csbc2017.mackenzie.br/anais/eventos/2-workshop-do-testbed-fibre',
10 'http://csbc2017.mackenzie.br/anais/eventos/4-encompif',
11 'http://csbc2017.mackenzie.br/anais/eventos/4-wpietfirtf',
12 'http://csbc2017.mackenzie.br/anais/eventos/5-wtrans',
13 'http://csbc2017.mackenzie.br/anais/eventos/6-brasnam',
14 'http://csbc2017.mackenzie.br/anais/eventos/6-desafie',
15 'http://csbc2017.mackenzie.br/anais/eventos/8-wcama',
16 'http://csbc2017.mackenzie.br/anais/eventos/9-sbcup',
17 'http://csbc2017.mackenzie.br/anais/eventos/11-bresci',
18 'http://csbc2017.mackenzie.br/anais/eventos/11-wit',
19 'http://csbc2017.mackenzie.br/anais/eventos/14-sbsc',
20 'http://csbc2017.mackenzie.br/anais/eventos/16-wperformance',
21 'http://csbc2017.mackenzie.br/anais/eventos/17-wim',
22 'http://csbc2017.mackenzie.br/anais/eventos/25-wei',
23 'http://csbc2017.mackenzie.br/anais/eventos/30-ctd',
24 'http://csbc2017.mackenzie.br/anais/eventos/36-ctic',
25 'http://csbc2017.mackenzie.br/anais/eventos/44-semish']
26
27 for link in links:
28     print link
29     html = urlopen(link)
30     soup = BeautifulSoup(html.read(),"html.parser")
31     tags = soup.findAll("div", {"class": "paper__text"})
32     title = soup.find(name = 'h2').get_text().strip()
33     for i in tags:
34         titulo = (i.find('h4', class_='paper__title').get_text()).strip()
35         titulos.append(titulo)
36         autor = (i.find('p', class_='paper__authors').get_text()).strip()
37         autores.append(autor)
38         eventos.append(title)
39
40 import sys
41 reload(sys)
42 sys.setdefaultencoding('utf8')
43 with open("trabalhos.csv", "w") as file:
44     writer = csv.writer(file, delimiter=",")
45     writer.writerow(["titulo", "autor", "evento"])
46     for i in range(0, len(autores)):
47         writer.writerow([titulos[i], autores[i], eventos[i]])
```

Fonte: Própria (2019).

Após limpeza, o conjunto totalizou 919 trabalhos publicados nestes três anos, sendo respectivamente 308, 348 e 263 artigos. Em busca de uma análise ao longo do tempo, manteve-se somente os congressos em comum entre os anos analisados. Por fim, 14 eventos foram elencados (Tabela 4), com 759 trabalhos publicados.

**Tabela 4 – Eventos do CSBC por tipo, ano e edição.**

<b>Sigla</b>	<b>Nome do evento</b>	<b>Tipo</b>	<b>Ano</b>	<b>Edição</b>
BraSNAM	<i>Brazilian Workshop on Social Network Analysis and Mining</i>	Satélite	2017	6°
			2018	7°
			2019	8°
BreSci	<i>Brazilian e-Science Workshop</i>	Satélite	2017	11°
			2018	12°
			2019	13°
CTD	Concurso de Teses e Dissertações da SBC	Base	2017	30°
			2018	31°
			2019	32°
CTIC	Concurso de Trabalhos de Iniciação Científica da SBC	Base	2017	36°
			2018	37°
			2019	38°
ENCompIF	Encontro Nacional de Computação dos Institutos Federais	Satélite	2017	4°
			2018	5°
			2019	6°
ETC	Encontro de Teoria da Computação	Satélite	2017	2°
			2018	3°
			2019	4°
SBCUP	Simpósio Brasileiro de Computação Ubíqua e Pervasiva	Satélite	2017	9°
			2018	10°
			2019	11°
SEMISH	Seminário Integrado de <i>Software e Hardware</i>	Base	2017	44°
			2018	45°
			2019	46°

WCAMA	Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais	Satélite	2017	8°
			2018	9°
			2019	10°
WEI	Workshop sobre Educação em Computação	Base	2017	25°
			2018	26°
			2019	27°
WIT	<i>Women in Information Technology</i>	Base	2017	11°
			2018	12°
			2019	13°
WPerformance	Workshop em Desempenho de Sistemas Computacionais e de Comunicação	Satélite	2017	16°
			2018	17°
			2019	18°
WPIETF	Workshop Pré-IETF	Satélite	2017	4°
			2018	5°
			2019	6°
WTrans	Workshop de Transparência em Sistemas	Satélite	2017	5°
			2018	6°
			2019	7°

---

Fonte: Própria (2019).

A SBC possui um modelo próprio de submissão de artigos, sendo utilizado em toda a conferência. É interessante destacar que diferentemente de boa parte de congressos e até mesmo da plataforma Lattes, uma das mais utilizadas por pesquisadores brasileiros, os nomes das/os autoras/es são descritos de forma literal e completa, sem abreviações pelo último sobrenome e inicial da letra do primeiro nome, o que contribuiu para identificar o gênero das/os autoras/es.

Nenhum dos eventos supracitados não fornecem informações quanto ao gênero dos componentes da autoria dos artigos. Portanto, foi necessária a utilização de uma biblioteca de classificação automática, para identificação do sexo dos indivíduos pesquisados e conseqüentemente, prover análises quanto a participação feminina no CSBC. Existem variados sistemas que realizam tal predição, para este

trabalho escolheu-se o pacote genderBR<sup>8</sup>, na versão 3.4.4, desenvolvido com base em informações públicas do Censo do Instituto Brasileiro de Geografia e Estatística (IBGE) do ano de 2010 (MEIRELES, 2018).

A maioria das publicações no Brasil constitui de autoria de brasileiras/os, sendo um fator importante a ser considerado para a classificação, que é executada a partir do nome próprio. Deste modo, o genderBR é acoplado ao R, um ambiente para implementação de funções estatísticas, muito utilizado para gerar visualizações e realizar análise de dados (SILVA; PERES; BOSCARIOLI, 2016).

Mesmo que possua um padrão da SBC para identificação de autoras/es em publicação de artigos, ainda foram encontrados nomes com abreviação. Assim, tais identificações foram buscadas manualmente *on-line* e encontrada a nomenclatura completa por meio de demais trabalhos já publicados, ou mesmo por redes de pesquisadores, como Lattes, *Research Gate* e/ou citações por pessoa do Google Acadêmico. Com a base preparada, a classificação foi obtida juntamente do uso da função *clipboard*, a qual pega os dados da área de transferência, criando um *dataframe* – uma espécie de matriz com possibilidade de valores de tipo de dados diferentes entre si (Algoritmo 2).

#### **Algoritmo 2 – Classificação do genderBR aos autores e autoras do CSBC.**

```
1 library(genderBR)
2 x <- read.table(file = "clipboard", sep = "\t", header=FALSE)
3 sexo <- get_gender(x$V1)
4 write.table(sexo, "clipboard", sep="\t", row.names=FALSE, col.names=FALSE)
```

Fonte: Própria (2019).

O Algoritmo 2 retorna classificação entre “*Female*” (Feminino), “*Male*” (Masculino) ou NA (Não Identificado). O genderBR considera como NA os nomes com proporções abaixo de 90%, geralmente quando o nome pode ser usado para ambos os gêneros ou mesmo, para aqueles que não foram encontrados na base do IBGE. Cerca de 4,4% dos indivíduos não foram identificados e também passaram por uma classificação manual. Por fim, o conjunto total de autores conteve 2.723 registros, prevalecendo a ordem de autoria entre 1° e 10° autor para cada artigo.

---

8 <https://cran.r-project.org/web/packages/genderBR/index.html>

Como última etapa de construção da base, foi criado um novo campo para armazenamento de nomes de autores corrigidos, sendo este processo denominado como desambiguação, como descrito por Duarte, Moura e Moro (2019). Isso foi feito para agrupamento de alguns registros de autores que referiam-se ao mesmo indivíduo, a exemplo de Thalia Santana e Thalia S. Santana, mantendo-se como regra o nome em sua forma mais completa, também realizado de modo manual.

Após tais procedimentos, os dados foram organizados de acordo com a modelagem relacional de um Diagrama Entidade-Relacionamento (DER), em um banco de dados MySQL<sup>9</sup> versão 8.0.16, por meio do Sistema Gerenciador de Banco de Dados (SGBD) MySQL *Workbench*<sup>10</sup> versão 8.0.15. Os dados passaram por análise por intermédio de consultas (comumente conhecidas em inglês como *queries*) em linguagem SQL (*Structured Query Language* – em português, Linguagem de Consulta Estruturada), um formato de pesquisa padrão para bancos de dados. Posteriormente, utilizou-se da linguagem de programação Python 3.7.3 dentro da plataforma Google *Colaboratory*<sup>11</sup>, juntamente das bibliotecas *Numpy*, *Scipy* e *Matplotlib*, onde foi possível gerar visualizações e análises estatísticas.

O longo processo descrito acima para a criação da base de dados é definido como *Extraction-Transformation-Loading* (ETL – em português, Extrair-Transformar-Carregar). Apesar de relacionar-se frequentemente com a *Business Intelligence* (inteligência de negócios) e o carregamento em um *datawarehouse* (armazém de dados), o ETL é aplicável ao carregamento de quaisquer bases de dados, visto acompanhar atividades importantes para oferecer aos usuários informações limpas, completas e verdadeiras (VASSILIADIS, 2009). Ainda conforme o autor, dados podem possuir inconsistências, como erros ortográficos ou a falta de alguma informação, podendo ser eliminados no âmbito do ETL, que para Kimball e Ross (2013), é composto por quatro áreas principais, sendo: extração, limpeza e conformidade, entrega e gerenciamento de dados úteis.

Levando em conta este panorama, foram levantadas as seguintes questões de pesquisa (QP):

- QP1) Qual é a distribuição de gênero no período analisado?

---

9 <https://www.mysql.com/>

10 <https://www.mysql.com/products/workbench/>

11 <https://colab.research.google.com/>

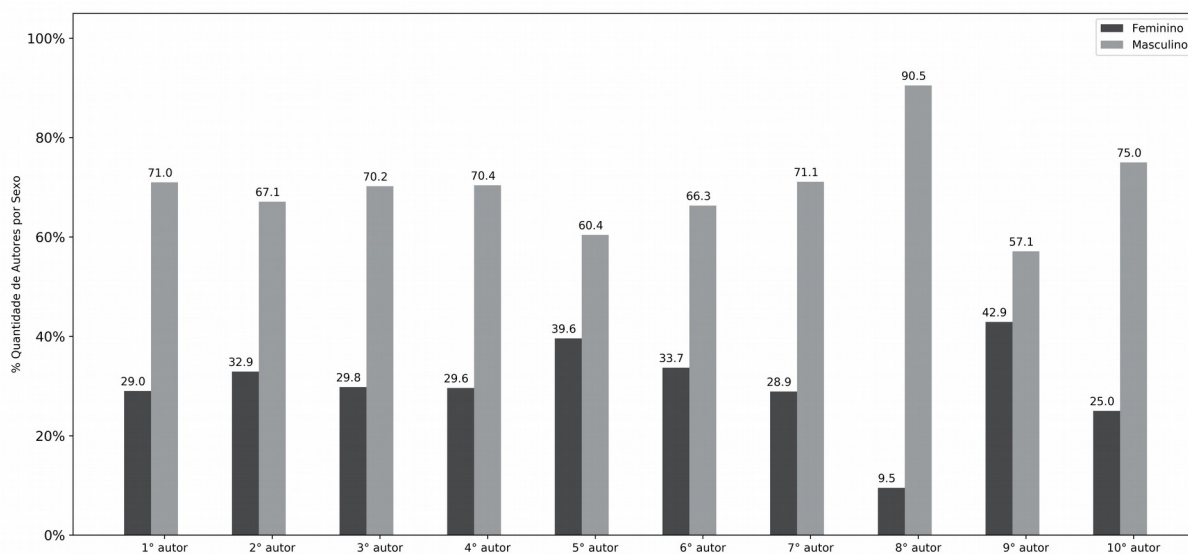
- QP2) Como a participação feminina varia ao longo dos anos em relação ao evento?
- QP3) Como varia o número de trabalhos aceitos por evento e ano?
- QP4) Como configura-se a diferença entre autoras e autores por ordem de autoria?
- QP5) Quantas mulheres estão entre os três autores mais prolíficos?
- QP6) Considerando as/os autoras/es, como é a presença feminina entre as produções científicas?
- QP7) Quais são as diferentes tendências de eventos em relação aos artigos publicados por mulheres?

Portanto, dentre os trabalhos e eventos selecionados, as seguintes variáveis foram identificadas: nome e gênero das/os autoras/es, título, edição e ano das conferências e ordem de autoria. O estudo se concentrou em artigos publicados nos anais do evento, sendo este trabalho guiado pelas QPs, as quais serão respondidas nas próximas sessões.



## RESULTADOS E DISCUSSÃO

Considerando os eventos base e satélite do CSBC, do total de registros da análise por autor, 1.878 pesquisadores são do gênero masculino (69%), em contraponto à 845 autoras do gênero feminino, que compõem um percentual de cerca de 31%. Os três com maior número de publicações são todos do sexo masculino, sendo o primeiro com 12 artigos, e o segundo e terceiro do *ranking* empatados com 11 trabalhos nos três anos do CSBC. Avaliando a quantidade de autores por sexo juntamente da ordem de autoria, 220 mulheres são primeira autora enquanto 539 homens se encontram como autor principal. A situação de desproporcionalidade se repete em todos os níveis de coautoria (Figura 1) verificados nas publicações, visto que em qualquer uma das ordens, o público masculino é maior que o feminino, mesmo não havendo regras de limite de autores por artigo.

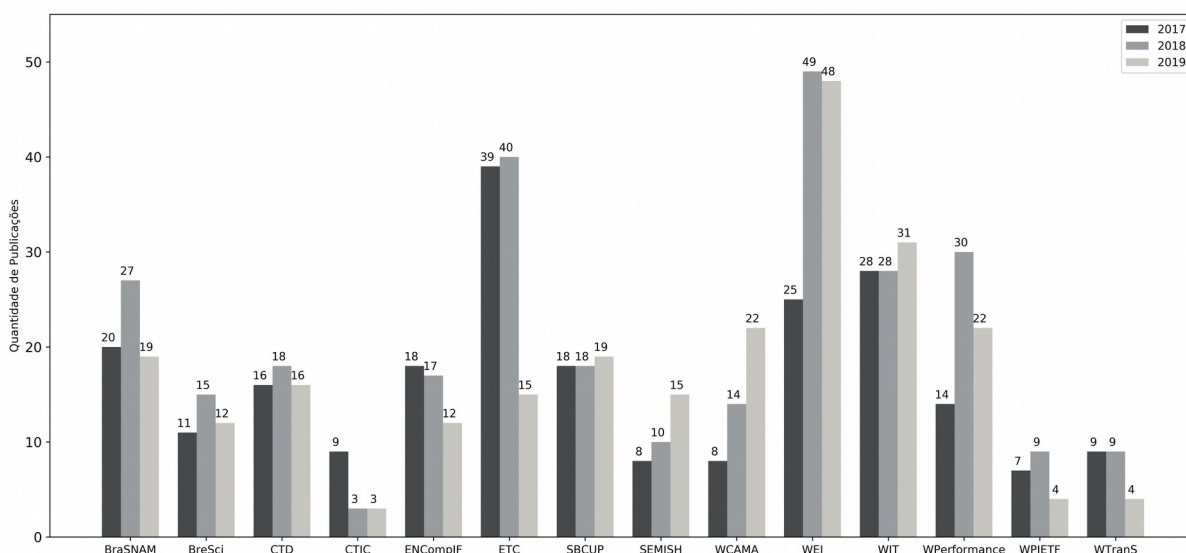


**Figura 1 – Gráfico de barras quanto ao gênero de autoras/es e ordem de autoria.**

Fonte: Própria (2019).

Quanto aos eventos, o WEI é o que possui maior número de trabalhos: 122 publicações. Em seguida, vem o ETC, com 94 produções e por último, o WIT, com 87 trabalhos. Destaca-se que no caso do WEI o número pode ser considerado como esperado, pois trata-se de um evento consolidado, que existe há 27<sup>a</sup> edições e é

enquadrado no índice Qualis B3, além de possuir uma área abrangente. Entretanto, os eventos consecutivos a tal são novos, o ETC está na 4ª edição e ainda é somente evento satélite. Já o WIT, na 13ª edição, apesar de ser um evento base, é o mais novo dentre os desta categoria (que possui eventos com quase meio século, a exemplo do SEMISH), demonstrando que a comunidade científica vem comumente participando de eventos novos, que vão de encontro à temática de pesquisa, mesmo que ainda não possuam reconhecimento do extrato Qualis/CAPES. Analisando ano a ano, o WEI figura nos anos de 2018 e 2019, com maior quantidade de publicações, como melhor pode ser observado na Figura 2.



**Figura 2 – Gráfico de barras do quantitativo de publicações por ano e evento.**

Fonte: Própria (2019).

O WIT dentre todos os eventos promovidos no CSBC, é o que possui maior número de mulheres como autoras principais, com 80 mulheres e 7 homens. Sendo o evento justamente destinado a temática de mulheres na computação, é esperado o elevado índice de pesquisadoras. Entretanto, eventos como o CTD, onde premia-se teses e dissertações da computação consideradas como de relevância, somente uma mulher foi encontrada como primeira autora – reflexos do baixo percentual de mestres e doutoras em CC (ALMEIDA; ZANLORENSSI, 2018).

Ainda no âmbito de autoras/es principais, em 2017, 2018 e 2019, respectivamente, publicaram 58, 89 e 73 mulheres, em relação à 172, 198 e 169

homens, o que corresponde a uma porcentagem de 25,2%, 31% e 30,2%. Quanto à presença feminina nos artigos analisados, um total de 449 publicações contém mulheres dentre as/os autoras/es, representando um percentual de cerca de 59,1%. Assim, mesmo em menor número, as cientistas estão contidas em mais da metade dos trabalhos que são publicados no CSBC. Já em relação ao público masculino, estão presentes em 699 artigos.

Para analisar a proporção de gênero dos autores de cada evento, além do total por sexo e porcentagem, calculou-se a razão feminino/masculino (F/M) de pessoas em cada um destes (Tabela 5), conforme efetuado por Ribeiro et al. (2019). Tanto em número absoluto quanto em proporção, os eventos “mais femininos” foram o WIT, WEI e BreSci. Todos com viés multidisciplinar, respectivamente, de forma geral tratam de questões envoltas à mulher na computação e projetos de incentivo, informática na educação e e-Ciência (em inglês – *e-Science*, relacionada a aplicações de apoio computacional para os processos de experimentação científica). Já com menor proporção, o CTD, WPIETF e o WPerformance obtiveram razão F/M mais baixa, sendo estes, em ordem, relacionados com trabalhos finais da pós-graduação, padrões para a Internet e desempenho de sistemas computacionais.

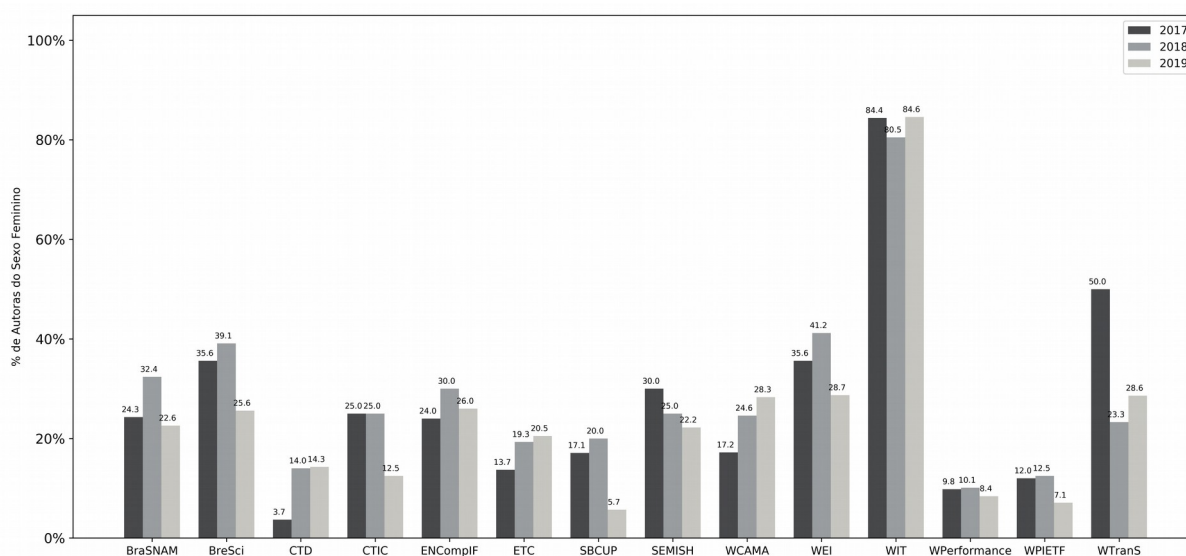
**Tabela 5 – Eventos, quantitativo de autoras/es por gênero e razão F/M.**

<b>Evento</b>	<b>Geral</b>	<b>Feminino (F)</b>	<b>Masculino (M)</b>	<b>%F</b>	<b>%M</b>	<b>Razão F/M</b>
WIT	339	282	57	83,2	16,8	4,9474
WEI	452	158	294	35,0	65,0	0,5374
BreSci	153	53	100	34,6	65,4	0,5300
WTranS	66	22	44	33,3	66,7	0,5000
BraSNAM	241	66	175	27,4	72,6	0,3771
ENCompIF	185	49	136	26,5	73,5	0,3603
WCAMA	192	49	143	25,5	74,5	0,3427
SEMISH	133	33	100	24,8	75,2	0,3300
CTIC	44	10	34	22,7	77,3	0,2941
ETC	260	45	215	17,3	82,7	0,2093
SBCUP	232	34	198	14,7	85,3	0,1717
CTD	112	13	99	11,6	88,4	0,1313
WPIETF	71	8	63	11,3	88,7	0,1270
WPerformance	243	23	220	9,5	90,5	0,1045

Fonte: Própria (2019).

Em prol de verificar a existência ou não de diferença estatisticamente significativa entre os dados apresentados na Tabela 5, utilizou-se o método ANOVA de fator único, que permite comparação entre médias de populações distintas, sendo efetuada entre os cinco grupos de maior (WEI, BreSci, WTranS, BraSNAM e ENCompIF) e menor proporção de autoras (ETC, SBCUP, CTD, WPIETF e WPerformance). Mesmo sendo o WIT com maior razão entre o público feminino, foi necessário desconsiderá-lo para obter uma visão mais realista entre os eventos e suas áreas. Com isso, obteve-se os valores de  $f\text{-ratio} = 53.809631$  e  $p\text{-value} = 0.000081$  – significativo para  $p < 0.05$ , demonstrando que há diferença significativa (hipótese alternativa) na distribuição por gênero em eventos do CSBC, em relação à proporção feminino/masculino, rejeitando a hipótese nula de médias iguais.

De modo mais específico, também foi verificado o percentual de autoras em cada um dos 14 eventos durante os três anos, conforme apresentado na Figura 3. De 2017 para 2018, é visto um considerável aumento de cientistas mulheres, contudo, houve queda no percentual como autoras em nove eventos em 2019. Nos cinco em que a porcentagem esteve mais elevada, em relação ao ano anterior (2018), como o CTD, ETC, WCAMA, WIT e WTranS, a taxa de aumento foi em boa parte mínima, sendo respectivamente, de 0,3%, 1,2%, 3,7%, 4,1% e 5,3%.



**Figura 3 – Gráfico de barras do percentual de autoras por ano e evento.**

Fonte: Própria (2019).

## CONSIDERAÇÕES FINAIS

A cienciometria vem possibilitando realizar avaliações cada vez mais profundas em torno do saber científico, o qual perpassa o caminho da produção da ciência por intermédio da difusão em periódicos e conferências. Ao passo que contribui para esta consolidação de conhecimento, é importante que tal mecanismo avalie os reflexos também no viés de gênero, a fim de possibilitar uma compreensão dos indivíduos que possuem impacto direto na construção da ciência.

Por meio deste estudo, acredita-se que haja contribuição com o corpo de pesquisas brasileiras com recorte de gênero na computação, já que os dados disponíveis frequentemente são dependentes de pesquisas isoladas e insuficientes para visualizações abrangentes (RIBEIRO et al., 2019). Desta maneira, apesar de na literatura possuírem pesquisas que avaliem especificamente alguns dos eventos presentes no CSBC sob a ótica cienciométrica, não foi encontrado nenhum estudo que investigue em específico todos os eventos desta conferência, ainda mais na perspectiva de grupos minoritários dentro do contexto de CSR.

Considerando as QPs citadas anteriormente, quanto a QP1, as mulheres ainda estão em menor número (31%) no quesito de produção científica, categorizando a pesquisa em CC como um campo masculino. Conforme o relatório “*Gender in the Global Research Landscape*” da Elsevier (ALLAGNAT et al., 2017), só há equilíbrio de gênero quando as mulheres compõem de 40-60% do total de pesquisadores. Os autores ainda apontaram que dentre 27 áreas de conhecimento, a computação no Brasil figura com 7% dos cientistas, sendo somente 2% mulheres (quase 1/3 do sexo masculino), tornando a projeção do CSBC um pouco acima do analisado internacionalmente pela Elsevier.

Ademais, comparado aos recentes estudos de Ribeiro et al. (2019) este valor supera o de 21,87% do gênero feminino de associadas da SBC, e mantêm-se quase similar ao encontrado por Duarte, Moura e Moro (DUARTE; MOURA; MORO, 2019) de cerca de 30% de participantes de comitês de programa em eventos de computação brasileiros. Contudo, ultrapassa quase em dobro o percentual de discentes matriculadas na área no ensino superior (cerca de 14%), demonstrando que em suma, mesmo em menor quantitativo, uma porcentagem considerável de

mulheres vem marcando presença nas publicações e delineando conhecimento científico no CSBC – principalmente quando comparado à demais indicadores.

Em relação à QP2, ainda é discrepante a quantidade de homens e mulheres por evento, não havendo em nenhum dos analisados, números de paridade próximos a 50% para ambos os gêneros. O mesmo ocorre quando visto de modo específico ano a ano, onde apesar de um crescente aumento do percentual feminino em 2018, o índice sofreu em maioria, brusca queda no ano seguinte, impossibilitando um crescimento gradativo e linear. Na QP3, conseqüentemente, oito dos eventos da conferência também sofreram decaimento no número de trabalhos aprovados para 2019 e no que refere-se a ordem de autoria (QP4), não há paridade começando pelos autores principais, com proporção F/M de 0,4081.

Considerando os autores de trabalhos mais ativos (QP5), foi verificado que nenhum destes tratou-se de uma mulher, avaliando-se os três com maior número de artigos publicados. O autor mais prolífico distribuiu suas publicações entre o SBCUP (computação ubíqua e pervasiva) e SEMISH (aspectos gerais de *software* e *hardware*), aqui com destaque às publicações do primeiro, sendo um evento com pouco mais de uma década de existência. A autora com maior número de publicações, situa-se na quarta posição com artigos em quatro eventos distintos (BraSNAM, WIT, CTIC e WEI), empatada em quantitativo com outro autor do sexo masculino. Entretanto, na QP6, para a presença feminina, é notório que mesmo em menor número, as mulheres são parte representativa da maioria das produções catalogadas, demonstrando o proeminente papel feminino na difusão e efetivação de saberes científicos.

Para a QP7, mulheres participantes do CSBC tem tendência de concentrarem-se em eventos que possuem características multidisciplinares. No caso de Informática na Educação, é considerado como um dos campos no país tipicamente definido como “feminino” (DUARTE; MOURA; MORO, 2019). Segundo estas autoras, o fator interdisciplinar pode ser visto como favorável, haja vista a possibilidade de atuar em variados ramos ligados a computação. Sob a perspectiva do quantitativo, os eventos com menor número de mulheres tiveram decréscimo do ano de 2018 em relação a 2019, mostrando uma tendência de afastamento feminino de áreas que tem como fator o componente tecnológico e/ou níveis mais altos de

titulação no caso do CTD, garantindo maior participação de mulheres em eventos direcionados ao público ou aqueles relacionados a fatores humanos e sociais.

Durante o desenvolvimento deste, alguns entraves foram perpassados, a exemplo de limitada literatura com enfoque no tema de estudo e falta de disponibilidade de dados de plataformas de submissão de artigos e/ou quaisquer fontes similares para respectivo fim, o que dificulta a compreensão de um panorama mais fidedigno a área. Não obstante, mesmo os anais do CSBC estando em endereços distintos quando da realização deste trabalho, a própria SBC lançou no ano de 2019 como resposta a demanda de seus associados e interessados, a biblioteca SBCOpenLib (SOL), como repositório para inserção de publicações no âmbito da sociedade e eventos apoiados (SBC, 2019b), que devidamente populada futuramente, auxiliará aos interessados neste processo, mantendo em um só local todas as publicações dos eventos por ela realizados.

Outrossim, espera-se que os dados analisados possam vir a fornecer subsídios para a SBC quanto ao cenário de pesquisas e produções bibliográficas brasileiras. Sendo portanto, de encontro quanto ao fomento de atividades em prol de equidade de gênero, bem como manutenção e destaque de ações apoiadas como o já cancelado Programa Meninas Digitais, para incentivo de meninas na área desde o ensino fundamental, o qual é institucionalizado pela referida sociedade científica e visto como de interesse nacional pela comunidade ligada à CC (MENINAS DIGITAIS, 2019).

Para trabalhos futuros, é importante aumentar o período de tempo da análise a fim de verificar efetivamente a evolução da participação feminina no CSBC por evento. Ademais, ressalta-se também a necessidade do acesso aos anais e aos resumos (*abstracts*) para a execução de algoritmos de classificação para se descobrir tendências de pesquisa dos mesmos, comparado ao gênero e evento. Desta maneira, será possível correlacionar a evolução por subáreas com a presença de mulheres. Com o aumento do número de registros, adicionalmente, submeter os dados a regras de associação (AGRAWAL; IMIELIŃSKI; SWAMI, 1993), com vistas a avaliar o quanto a presença de determinados itens na base de dados implicam em demais registros distintos, verificando-se as subáreas da CC que mais tem evoluído mediante as publicações das cientistas.

## REFERÊNCIAS

- AGARWAL, S. et al. Women in computer science: What is the Bibliography Data Telling Us? **ACM Computers & Society (SIGCAS)**, v. 46, n. 1, p. 7–19, 2016.
- AGARWAL, S.; MITTAL, N.; SUREKA. Minority Ethnic Groups in Computer Science Research. What is the Bibliography Data Telling Us? **ACM SIGCAS Newsletter Volume 47 Issue 2**, v. 47, n. 2, p. 102–111, 2017.
- AGRAWAL, R.; IMIELIŃSKI, T.; SWAMI, A. Mining association rules between sets of items in large databases. **ACM SIGMOD RECORD**, v. 22, n. 2, p. 207–216, 1993.
- ALBORNOZ, M. et al. **Las brechas de género en la producción científica Iberoamericana**. Observatorio Iberoamericano de la Ciencia, la Tecnología y la Sociedad de la Organización de Estados Iberoamericanos (OCTS-OEI). Papeles del Observatorio nº 09 - Octubre de 2018.
- ALLAGNAT, L. et al. **Gender in the Global Research Landscape**. Elsevier, 2017.
- ALMEIDA, R.; ZANLORENSSI, G. **Qual o gênero e a idade de mestres e doutores no Brasil**. Disponível em: <<https://www.nexojournal.com.br/grafico/2018/05/23/Qual-o-genero-e-a-idade-de-mestres-e-doutores-no-Brasil>>. Acesso em: 20 abr. 2019.
- ARRUDA, D. et al. Brazilian computer science research: Gender and regional distributions. **Scientometrics**, v. 79, n. 3, p. 651–665, 2009.
- BANSHAL, S. K. et al. Mapping Computer Science research in Bangladesh. **SKIMA 2014 - 8th International Conference on Software, Knowledge, Information Management and Applications**, 2014.
- BANSHAL, S. K. et al. Computer science research in India: A scientometric study. **12th IEEE International Conference Electronics, Energy, Environment, Communication, Computer, Control: (E3-C3), INDICON**, 2015.
- BEAUBOUEF, T.; ZHANG, W. Where are the women computer science students? **Journal of Computing Sciences in Colleges**, p. 14–20, 2011.
- CAPES. **Considerações sobre Qualis: Ciência da Computação**. Disponível em: <[https://www.capes.gov.br/images/documentos/Qualis\\_periodicos\\_2016/CienciaComputacao2013-2015.pdf](https://www.capes.gov.br/images/documentos/Qualis_periodicos_2016/CienciaComputacao2013-2015.pdf)>. Acesso em: 14 maio. 2019.
- CAPES. **Qualis**. Disponível em: <<https://www.capes.gov.br/acessoainformacao/perguntas-frequentes/avaliacao-da-pos-graduacao/7422-qualis>>. Acesso em: 14 maio. 2019.
- CAVERO, J. M. et al. The evolution of female authorship in computing research. **Scientometrics**, v. 103, n. 1, p. 85–100, 2015.



- CODEÇO, C. T.; DIAS, C. M. Mulheres na ciência. **Cadernos de Saúde Pública**, v. 34, n. 10, p. 2–3, 2018.
- COHOON, J. M.; NIGAI, S.; KAYE, J. Gender and computing conference papers. **Communications of the ACM**, v. 54, n. 8, p. 72, 2011.
- CONCEIÇÃO, J. M. DA; TEIXEIRA, M. DO R. F. Mulheres na ciência: um estudo da presença feminina no contexto internacional. **#Tear: Revista de Educação, Ciência e Tecnologia**, v. 7, n. 1, p. 1–18, 2018.
- CONTESSA, D. F.; PALAZZO, J. DE O. M. **Biblioteca Digital da SBC: Disponibilizando os Artigos do JEMS Através de um Provedor de Dados Compatível com OAI**. II Workshop de Bibliotecas Digitais. **Anais...** Florianópolis: 2006.
- CORDEIRO, M. D. **Mulheres e ciência**. 2º Simpósio de Integração Científica e Tecnológica do Sul Catarinense – SICT-Sul. **Anais...**2013.
- CSBC. **Inscrições**. Disponível em: <<http://csbc2019.sbc.org.br/inscricoes/>>. Acesso em: 11 out. 2019.
- CUNHA, M. B. et al. As mulheres na ciência: O interesse das estudantes brasileiras pela carreira científica. **Educacion Quimica**, v. 25, n. 4, p. 407–417, 2014.
- DUARTE, B. D.; MOURA, A. F. C.; MORO, M. M. **Mulheres na Computação: Análises por Sub-Áreas**. Women in Information Technology (WIT) - CSBC. **Anais...**2019.
- EFFENDY, S.; YAP, R. H. C. Analysing Trends in Computer Science Research: A Preliminary Study Using The Microsoft Academic Graph. **WWW (Companion Volume)**, p. 1245–1250, 2017.
- GALLIVAN, M. J.; BENBUNAN-FICH, R. Examining the relationship between gender and the research productivity of IS faculty. **SIGMIS-CPR'06**, p. 103–113, 2006.
- GÜRER, D.; CAMP, T. An ACM-W literature review on women in computing. **ACM SIGCSE Bulletin**, v. 34, n. 2, p. 121, 2002.
- HUSSAIN, A. J. et al. An Investigation into Gender Disparities in the Field of Computing. **Proceedings - 2015 International Conference on Developments in eSystems Engineering, DeSE 2015**, p. 20–25, 2015.
- JEMS. **Welcome to JEMS**. Disponível em: <<https://jems.sbc.org.br/jems2/>>. Acesso em: 12 maio. 2019.
- KHAN, N. Z.; LUXTON-REILLY, A. Is Computing for Social Good the Solution to Closing the Gender Gap in Computer Science? **ACSW '16 Proceedings of the Australasian Computer Science Week**, p. 1–5, 2016.

KIMBALL, R.; ROSS, M. **The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling**. Third Edit ed. John Wiley & Sons, 2013.

LOUZADA, C. S. et al. Um mapeamento das publicações sobre o ingresso das mulheres na computação. **40 CLEI/ VI Congreso de la Mujer Latinoamericana en la Computación (LAWCC)**, 2014.

MEIRELES, F. **Predizendo sexo a partir de nomes próprios com o genderBR**. Disponível em: <<https://fmeireles.com/blog/rstats/predizendo-sexo-nomes-proprios-genderbr/>>. Acesso em: 10 out. 2019.

MENINAS DIGITAIS. **Sobre o Programa Meninas Digitais**. Disponível em: <<http://meninas.sbc.org.br/index.php/sobre/>>. Acesso em: 4 nov. 2019.

MONARD, M. C.; FORTES, R. P. DE M. **Uma visão da Participação Feminina nos Cursos de Ciência da Computação no Brasil**. V Congresso de la Mujer Latinoamericana en La Computacion, LAWCC. **Anais...**2013.

NOOSRIKONG, C.; NGAMSURIYAROJ, S.; AYUDHYA, S. P. N. Identifying focus research areas of computer science researchers from publications. **IEEE Region 10 Annual International Conference, Proceedings/TENCON**, p. 811–816, 2017.

RIBEIRO, K. DA S. F. M. et al. **Uma análise de gênero a partir de dados da Sociedade Brasileira de Computação**. Women in Information Technology (WIT) - CSBC. **Anais...**2019.

ROSSI, A. S. Women in Science: Why So Few?: Social and psychological influences restrict women's choice and pursuit of careers in science. **Science**, v. 148, n. 3674, p. 1196–1202, 1965.

SBC. **Educação Superior em Computação Estatísticas – 2017**. Disponível em: <<https://www.sbc.org.br/documentos-da-sbc/summary/133-estatisticas/1200-pdf-png-educacao-superior-em-computacao-estatisticas-2017>>.

SBC. **CSBC - Congresso da Sociedade Brasileira de Computação**. Disponível em: <<https://www.sbc.org.br/eventos/csbc>>. Acesso em: 11 out. 2019a.

SBC. **Conheça a SBC Open Lib, a biblioteca digital da SBC**. Disponível em: <<https://www.sbc.org.br/noticias/10-slideshow-noticias/2134-conheca-a-sbc-open-lib-a-biblioteca-digital-da-sbc>>. Acesso em: 12 out. 2019b.

SCHIEBINGER, L. Mais mulheres na ciência: questões de conhecimento. **História, Ciências, Saúde**, v. 15, p. 269–281, 2008.

SCHWARTZ, J. et al. Mulheres na informática: quais foram as pioneiras? **Cadernos Pagu**, n. 27, p. 255–278, 2006.

SILVA, L. A. DA; PERES, S. M.; BOSCARIOLI, C. **Introdução à Mineração de Dados com aplicações em R**. 1ª edição ed. GEN LTC, 2016.

SMEDING, A. Women in Science, Technology, Engineering, and Mathematics (STEM): An Investigation of Their Implicit Gender Stereotypes and Stereotypes' Connectedness to Math Performance. **Sex Roles**, v. 67, n. 11–12, p. 617–629, 2012.

TAGUE-SUTCLIFFE, J. An introduction to informetrics. **Information processing & management**, v. 28, n. 1, p. 1–13, 1992.

VARGIU, E.; URRU, M. Exploiting web scraping in a collaborative filtering-based approach to web advertising. **Artificial Intelligence Research**, v. 2, p. 44–54, 2013.

VASSILIADIS, P. A Survey of Extract–Transform–Load Technology. **International Journal of Data Warehousing and Mining**, v. 5, n. 3, 2009.

VELA, B.; CÁCERES, P.; CAVERO, J. M. Participation of women in software engineering publications. **Scientometrics**, v. 93, n. 3, p. 661–679, 2012.

YANG, L. et al. Big data analysis for SCIE research output of Chinese computer science. **6th International Conference on Computer Science and Network Technology, ICCSNT 2017**, p. 261–265, 2017.