

# Panorama sobre Participação das Mulheres no IHC

Êrica Peters do Carmo<sup>1</sup>, Ediana da Silva de Souza<sup>1</sup>, Arthur Felipe Herdt Schuelter<sup>1</sup>,  
Milene Selbach Silveira<sup>2</sup>, Simone Diniz Junqueira Barbosa<sup>3</sup>,  
Rebeca Schroeder<sup>1</sup>, Isabela Gasparini<sup>1</sup>

<sup>1</sup>Depto de Ciência da Computação - Universidade do Estado de Santa Catarina (UDESC)

<sup>2</sup>Escola Politécnica - Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

<sup>3</sup>Depto de Informática - Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)

{ericapetersc, edianadasilvadesouza, arthur.schuelter}@gmail.com,

milene.silveira@pucrs.br, simone@inf.puc-rio.br,

rebeca.schroeder@udesc.br, isabela.gasparini@udesc.br

**Abstract.** *Studies indicate low participation of women in Computing, causing entities to promote specific actions to try to modify this scenario. Specifically, concerning the Human-Computer Interaction community, which, among its fronts, discusses aspects of interdisciplinarity and diversity related to the construction of computational artifacts, it is important to analyze, in a more rigorous way, the panorama of the participation of women. Thus, this paper presents research about the presence of women in this community, based on the publications of the Brazilian Symposium on Human Factors in Computing Systems. As a result, we can see a balance in the number of authors in the publications, especially in the last 5 years. However, when looking specifically at issues such as the genre of the first author and the Brazilian states in which the author works, it is possible to observe differences in this distribution.*

**Resumo.** *Estudos apontam baixa participação de mulheres na Computação, fazendo com que entidades promovam ações específicas para tentar modificar este cenário. Especificamente em relação à comunidade de Interação Humano-Computador que, dentre suas frentes, discute aspectos de interdisciplinaridade e diversidade relacionados à construção de artefatos computacionais, é importante analisar, de forma mais rigorosa, o panorama de atuação das mulheres. Deste modo, este artigo apresenta uma pesquisa sobre a presença das mulheres nesta comunidade, a partir das publicações do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais. Como resultado, podemos verificar um equilíbrio no número de autores e autoras nas publicações, especialmente nos últimos 5 anos. Porém, ao observar especificamente questões como o gênero do primeiro autor e estados brasileiros em que o autor atua, é possível observar diferenças nesta distribuição.*

## 1. Introdução

A Organização das Nações Unidas apresenta 17 objetivos de desenvolvimento sustentável para transformar o mundo. Destes, o objetivo 5 – *Gender Equality* – trata da igualdade

de gênero e, dentre os seus sub-objetivos, está o de “aprimorar o uso da tecnologia capacitadora, em particular a tecnologia da informação e comunicação, para promover o empoderamento das mulheres” [ONU 2020].

Para promover o empoderamento das mulheres por meio de tecnologias de informação e comunicação, é importante que mulheres façam parte das equipes de design e desenvolvimento dessas tecnologias [Monteiro 2019]. Entretanto, em termos de formação acadêmica, e considerando apenas a realidade brasileira, é possível observar que estamos muito longe desse cenário. Como exemplo, ao analisarmos os cursos de graduação da área de Computação no Brasil, a partir de dados disponibilizados pela Sociedade Brasileira de Computação (SBC), referentes aos anos de 2001 a 2017, podemos observar a evolução do número de concluintes do gênero masculino e a estabilidade do (baixo) número de concluintes do gênero feminino [Nunes 2017].

Neste cenário, entidades da área de Computação, como a própria SBC e a *Association for Computing Machinery* (ACM), promovem ações específicas para fomentar a participação feminina na área, como o programa nacional Meninas Digitais e o ACM-W (*Women in Computing*), incluindo desde a divulgação da área a meninas do Ensino Fundamental até a defesa do envolvimento de mulheres em todos os aspectos da área de Computação. No sentido de contribuir para as pesquisas sobre este tema, nosso foco está na análise da participação de mulheres na comunidade de Interação Humano-Computador (IHC) no Brasil, especificamente em relação à pesquisa acadêmica. A área de IHC, segundo pesquisa realizada na comunidade brasileira, está entre as *top 5* áreas de interesse com maior proporção de mulheres, considerando-se os associados da SBC, sendo ainda a área com maior razão Feminino/Masculino (0,444) dentre as diferentes áreas da SBC [Ribeiro et al. 2019]. Dentre suas frentes, esta área também discute aspectos de interdisciplinaridade e diversidade relacionados à construção de artefatos computacionais.

Assim, para analisar a presença das mulheres nesta comunidade, tomamos como base as publicações do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais, principal evento da SBC na área de IHC. Dentre os resultados obtidos, temos, por edição do evento, a distribuição de autores por gênero, a organização destes autores (individualmente ou em grupos), a ordem de autoria (análise de primeiro e último autor) e a distribuição regional dos autores, particularmente, das autoras mulheres.

Destacamos que, como as análises deste trabalho são resultados de uma investigação com base nos nomes dos autores dos artigos do Simpósio, sabemos que a análise de gênero é limitada e não representa a todos. Neste sentido, Jaroszewski et al. [Jaroszewski et al. 2018] destacam que pessoas trans e não binárias ficam muitas vezes invisíveis em pesquisas e que a identidade de gênero deve ser um assunto de investigação acadêmica, social e política. Reconhecemos esta limitação atual do trabalho e pretendemos abordar a temática de forma mais ampla e robusta em trabalhos posteriores.

A fim de apresentar a pesquisa realizada, este artigo está organizado da seguinte forma: a Seção 2 descreve o processo metodológico seguido para realização desta pesquisa, a Seção 3 analisa os resultados obtidos e discute a análise destes resultados e, por fim, a Seção 4 apresenta as considerações finais sobre a pesquisa apresentada.

## 2. Processo Metodológico

Este trabalho analisou os padrões de gênero dos autores que publicaram artigos no Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (doravante chamado de IHC), considerando as edições que ocorreram entre os anos de 2006 a 2019. Justifica-se o período coletado visto que somente em 2006 o IHC passou a ser indexado pela Biblioteca Digital da ACM.<sup>1</sup>

A extração de dados teve como objetivo coletar os metadados referentes a todas as publicações das edições do IHC de 2006 a 2019 (optamos por considerar os artigos publicados até 2019 pois a extração de dados foi realizada durante o ano de 2020). Como a base da DBLP<sup>2</sup> possui indexação dos artigos publicados nestas edições, ela foi utilizada para encontrar as URLs de cada artigo. Uma vez que os *links* dos artigos foram coletados, um *web scraper* foi desenvolvido para acessar essas páginas e coletar os metadados referentes às publicações e seus autores, tais como título do artigo, resumo, palavras-chave, autores e filiações. Após extraídos, os dados foram convertidos e armazenados em arquivos CSV. Ao total, foram coletados e armazenados os metadados de 777 documentos.

Após a extração foi realizada a verificação dos dados coletados, conferindo o número de documentos indexados na base ACM com o número de documentos coletados. Após a confirmação de que o *dataset* criado e a base possuíam o mesmo número de documentos, passou-se para a fase de processamento dos dados e aplicação de APIs.

A primeira API utilizada foi a API GenderBr,<sup>3</sup> que identifica o gênero dos nomes próprios brasileiros utilizando dados do Censo 2010 do Instituto Brasileiro de Geografia e Estatística (IBGE). Para cada nome próprio enviado à API, esta calcula as proporções de homens e mulheres com esse nome, tendo como base os dados do IBGE, e retorna como resposta o gênero relacionado ao nome caso alguma das proporções (masculina ou feminina) seja maior que 90%. Caso nenhuma das proporções atenda este valor de corte, o nome é classificado como *missing*. Isto ocorre principalmente com nomes próprios que podem ser utilizados por ambos os sexos ou nomes estrangeiros. Em relação aos dados deste trabalho, dos 777 documentos analisados foram extraídos um total de 2457 nomes de autores indexados. Destes, 118 (cerca de 4,8%) não puderam ser classificados pela API. Desta forma, após remover as duplicidades destes nomes, foi realizado um processo manual para a classificá-los com base na busca pelos autores em suas páginas pessoais e currículos. Ao final deste processo, todos os 2457 autores foram categorizados por gênero de acordo com seus nomes.

Diversos autores de múltiplos artigos apareciam com nomes escritos de formas diferentes, *e.g.*, sem ou com acentos, com nomes intermediários omitidos, apenas suas iniciais ou por extenso. Para unificar os nomes e considerar apenas o nome completo e com acentos, seguimos duas estratégias em paralelo: (i) calculamos a distância de edição de Levenshtein<sup>4</sup>[Levenshtein 1966] e (ii) agrupamos os autores com primeiro e último nomes em comum. Com base nesses dados, inspecionamos manualmente os autores e seus grupos e substituímos cada nome abreviado ou sem acento pelo nome completo

---

<sup>1</sup>ACM Digital Library: <https://dl.acm.org/>

<sup>2</sup>DBLP (Digital Bibliography & Library Project): <https://dblp.uni-trier.de/>

<sup>3</sup><https://github.com/meirelesff/genderBR>

<sup>4</sup>Esta distância é calculada com base no número mínimo de operações sobre caracteres para transformar uma cadeia de caracteres em outra.

correspondente. Com este procedimento, reduzimos o conjunto inicial de 2457 para 1549 autores.

Para a obtenção do estado de atuação de cada autor, foi desenvolvido um *script* em Python para identificar o nome ou a sigla do estado brasileiro presente no campo de dados relacionado à instituição do autor. A partir dos metadados dos 777 documentos, foram extraídos 1099 diferentes nomes de instituições. Destes, 844 (cerca de 76,8%) tiveram seus estados brasileiros reconhecidos pelo *script*. As instituições restantes não possuíam informações sobre seus respectivos estados no campo de dados. Assim, foi realizado um processo manual de classificação a partir da busca de seus endereços.

É importante notar que ao longo dos anos um mesmo autor pode ter estado vinculado a diferentes instituições. Neste caso, para as análises relacionadas ao seu local de atuação, cada autor foi contado de acordo com os estados de suas instituições vinculadas. Ao final deste processo, 118 autores tiveram suas instituições identificadas como estrangeiras e classificadas de acordo com seus respectivos países e 36 autores não tiveram seus estados identificados, pois não havia informações sobre suas instituições nos metadados das publicações. Um total de 1432 autores foram classificados de acordo com seus estados brasileiros de atuação.

### 3. Análises e Resultados

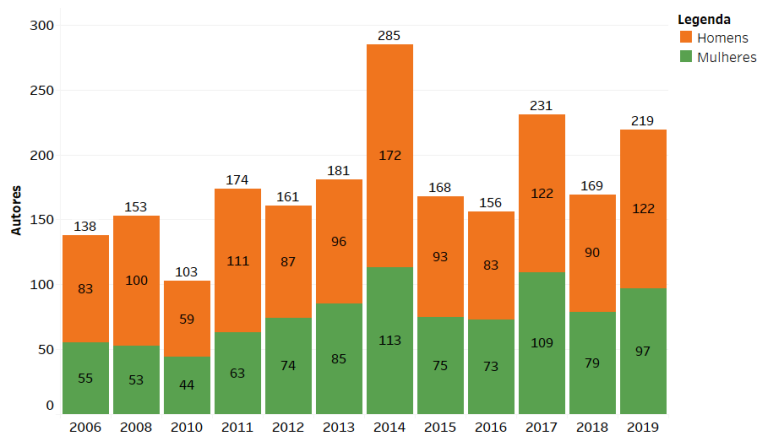
A Tabela 1 apresenta uma visão geral sobre os dados coletados das publicações no IHC entre 2006 e 2019. Nota-se que o número de autores por ano contabiliza cada autor uma única vez, mesmo que ele tenha publicado mais de um artigo no mesmo ano. Da mesma forma, o número total de autores apresentado na última linha da tabela contabiliza cada autor apenas uma única vez, mesmo que ele tenha tido publicações em várias edições do IHC.

**Tabela 1. Visão geral – Dados do IHC por ano**

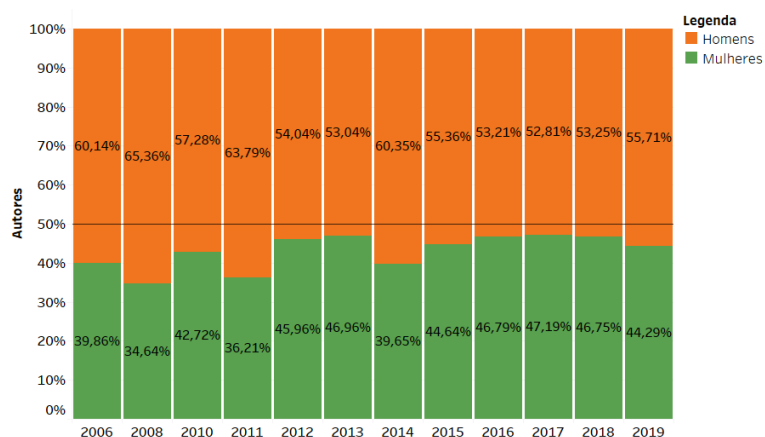
Ano	Número de Artigos	Número de Autores	Média de Autores por Artigo
2006	47	138	3,11
2008	62	153	2,35
2010	46	103	3,17
2011	67	174	2,18
2012	67	161	2,18
2013	63	181	2,32
2014	108	285	1,35
2015	60	168	2,43
2016	57	156	2,56
2017	74	231	1,97
2018	54	169	2,70
2019	72	219	2,03
Total	777	1549*	3,16

A primeira análise realizada diz respeito à distribuição de homens e mulheres entre os autores em cada ano. A Figura 1 apresenta o número de autores de cada gênero por ano. É possível notar que apesar das diferenças entre o total de autores em cada edição do IHC, o número de autores homens sempre é maior do que o número de autoras mulheres. A maior diferença pode ser observada em 2008, com uma proporção de 1,89 autores homens para cada autora mulher, enquanto a menor diferença pode ser observada em 2013, com

1,13 autores homens para cada autora mulher. De forma a complementar esta análise, a Figura 2 apresenta a distribuição percentual de forma normalizada de autores homens e mulheres em cada ano. Apesar de a presença de homens ser maior que a de mulheres em todas as edições do IHC, é possível perceber que a diferença entre os gêneros diminui ao longo dos anos, principalmente quando observados os últimos cinco.



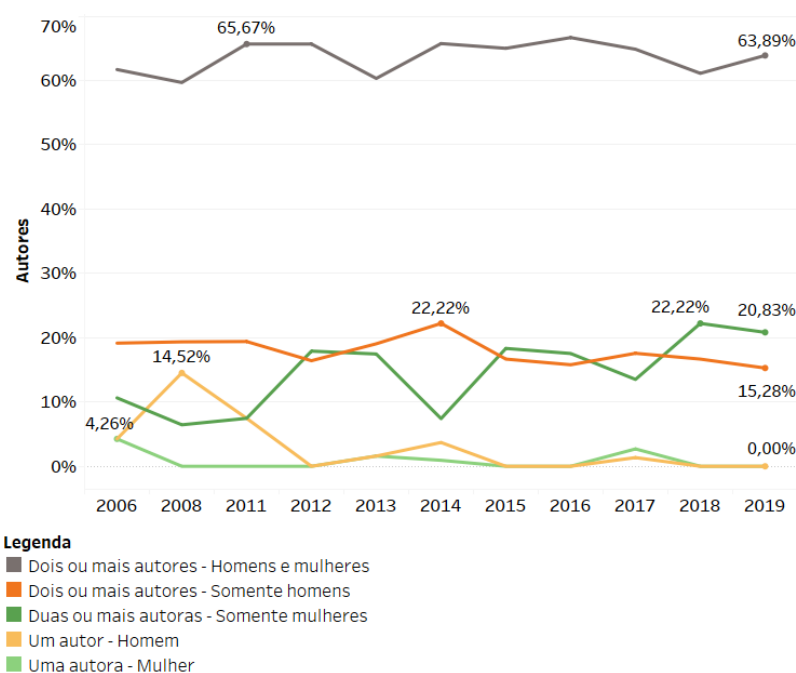
**Figura 1. Número absoluto de autores de cada gênero por ano (edição)**



**Figura 2. Porcentagem de cada gênero por ano (edição)**

Outra contribuição para o entendimento da proporção entre autores homens e mulheres no IHC é observar como os autores se organizam, individualmente ou em grupo, nas publicações realizadas ao longo das edições (Figura 3).

Um ponto observado nestes padrões é o fato de que apenas uma pequena porção das publicações totais são feitas por um único autor, independente do seu gênero. Além disso, publicações feitas por um grupo formado somente por homens ou somente por mulheres também não são maioria, nunca representando mais que 23% dos trabalhos publicados em um ano. Em todas as edições do IHC, pelo menos 60% dos artigos foram escritos por dois ou mais autores, em grupos formados tanto por homens quanto por mulheres. Este é um fator positivo não só pela questão de gênero mas também devido ao número de autores trabalhando juntos, pois uma área tão interdisciplinar precisa que a pesquisa científica seja colaborativa e busque diferentes visões para a resolução de problemas.

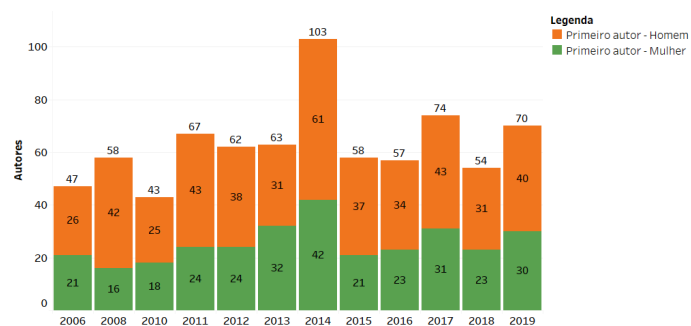


**Figura 3. Número e gênero dos autores de cada artigo por ano (edição)**

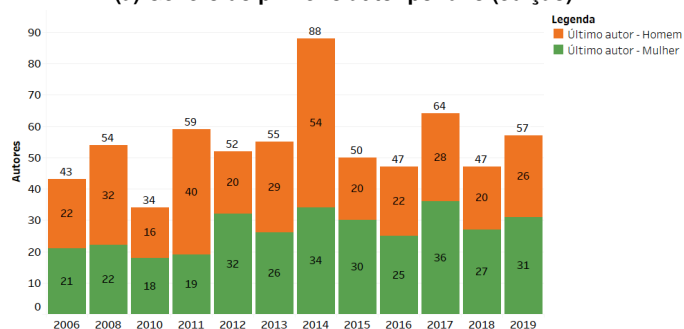
As duas análises apresentadas na Figura 4 buscam investigar a questão do gênero em relação aos alunos orientados e professores orientadores presentes entre os autores de cada publicação. Apesar de não haver uma regra imposta quanto à ordem dos autores em publicações científicas, há um entendimento da comunidade, na maioria dos artigos, em apresentar os orientados como primeiros autores do artigo e os orientadores por último. Assim, utilizou-se o padrão de ordem dos nomes dos autores, onde o primeiro nome é o do orientado e o último o do orientador.

Considerando o primeiro autor de cada artigo como o aluno orientado, é possível visualizar uma grande diferença entre o número de orientados homens e mulheres. Com exceção de 2013, em todos os anos o número de homens foi superior ao de mulheres. Isto pode ser considerado um dos reflexos da baixa presença feminina entre os estudantes de cursos superiores relacionados a Computação. De maneira distinta, ao admitir o último autor de cada artigo como o professor orientador, observamos uma distribuição de gênero que varia de acordo com cada edição. Enquanto anos como 2011 e 2014 tiveram muito mais orientadores homens do que mulheres, nos últimos cinco anos nota-se que as publicações no geral têm mais orientadoras mulheres do que homens. Conforme abordado anteriormente, uma das possíveis causas desta mudança é o aumento da presença feminina entre os autores do IHC tendendo ao equilíbrio da distribuição dos gêneros.

Deve-se lembrar, também, tanto para a análise do orientando quanto para a do orientador, que como os nomes destes autores estão sendo contados uma única vez para cada análise (primeiro ou último autor), não é levado em conta o quanto cada um deles produz. Ou seja, uma orientadora que publica diversos artigos em uma mesma edição não aumenta proporcionalmente o número de mulheres orientadoras nesta edição, pois ela é contada como apenas uma autora. Constata-se então que, apesar de podermos afirmar que existem mais orientados ou orientadores de um gênero do que de outro em determinado



(a) Gênero do primeiro autor por ano (edição)



(b) Gênero do último autor por ano (edição)

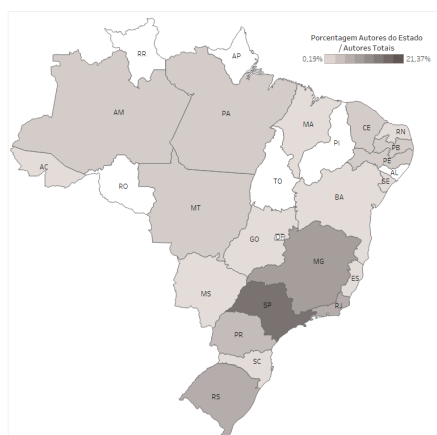
**Figura 4. Gênero dos autores de acordo com a ordem de aparição**

ano, não podemos concluir que a maior parte das publicações nesta edição é feita por este mesmo gênero.

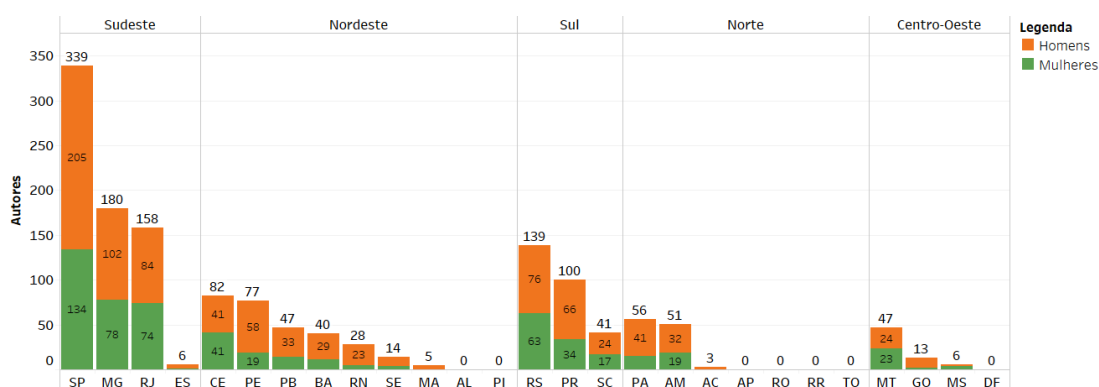
Em adição, é interessante entender como cada estado contribui em relação ao número de autores do IHC e como estes estão divididos entre homens e mulheres. Em relação a essas duas análises, uma alteração na contabilização dos autores precisou ser feita. Percebeu-se que ao longo dos anos alguns autores mudaram de instituição, alterando seus estados de atuação. Assim, para eliminar qualquer viés em relação a estados específicos, nas próximas análises cada autor foi contabilizado de acordo com o número de estados das instituições a que ele esteve vinculado durante os anos em que publicou no IHC. Sendo assim, se um mesmo autor publicou dois artigos enquanto estava vinculado a duas instituições em dois estados diferentes, ele será contabilizado duas vezes. Nota-se ainda que, dos 1549 autores do IHC, 118 estavam vinculados a instituições estrangeiras e 36 autores não preencheram as informações sobre suas instituições. Já que as análises a seguir abordam o cenário nacional, esses dois grupos foram desconsiderados.

A análise da contribuição de autores por estado está representada na Figura 5. Nela, é possível observar a porcentagem que o total dos autores de cada estado representa no número total de autores que já publicaram no IHC. Os estados na cor branca são aqueles cujas instituições não estão vinculadas a nenhum autor. Nota-se que a região brasileira com maior número de autores é a região Sudeste; e a região com menor número de autores é a região Norte.

Já em relação à distribuição de gênero entre os estados, observa-se que, de maneira geral, o número de mulheres é menor que o de homens. Isto está ilustrado na Figura 6. Um dos pontos a notar é que há dois estados, Acre (AC) e Maranhão (MA) em que existem au-



**Figura 5. Autores por estado brasileiro**



**Figura 6. Gênero dos autores por estado brasileiro**

tores homens mas não autoras mulheres, o contrário não ocorre em nenhum outro estado. Além disso, o Ceará (CE) é o único estado onde cada gênero representa exatamente 50% do total. Observa-se ainda que nenhum estado possui mais mulheres do que homens. Dos vinte estados com autores que publicaram algum artigo no IHC, a porcentagem média de autores homens e mulheres é de 61,03% e 38,97%, respectivamente.

Por fim, a Figura 7 apresenta a porcentagem das mulheres por estado (mapa à esquerda) e região (mapa à direita) do Brasil em relação ao total de mulheres que publicaram artigos entre 2006 e 2019. Nota-se que, da mesma forma que São Paulo é o estado que mais contribui com autores de ambos os gêneros, é também o estado com maior presença de mulheres. Em relação aos estados que não apresentam nenhuma autora, além dos que não apresentam trabalhos, entram também na lista Acre (AC) e Maranhão (MA).

Sobre as regiões brasileiras, é possível perceber uma mudança em relação as contribuições de autores de ambos os gêneros. A segunda e terceira região que mais contribuem no IHC, observando todos os autores (conforme Figura 6), são Nordeste e Sul, e, ao observar somente as autoras mulheres, a ordem se inverte (conforme Figura 7) – as regiões Sul e Nordeste, ocupam, respectivamente, o segundo e terceiro lugar de contribuições por autoras.

Finalmente, atenta-se para o fato de que apenas três regiões (Sudeste, Sul e Nor-





diferenças na proporção de autores homens e mulheres, a presença feminina vem crescendo nos últimos anos e há uma tendência ao equilíbrio entre os gêneros. Ao observar a maneira como os autores estão divididos (individualmente ou em grupos), constata-se que a maior parte das publicações são feitas por grupos formados por dois ou mais autores com a presença tanto de homens quanto de mulheres. Já em relação à distribuição de autores pelo território nacional, observa-se que as regiões Sudeste, Nordeste e Sul são as que mais contribuem com autores de ambos os gêneros. Especificamente em relação à presença feminina, nove estados brasileiros ainda não estão vinculados a nenhuma autora, e as regiões Norte e Centro-Oeste representam apenas 11,29% do total de mulheres que publicaram no IHC.

Destaca-se, então, como um dos desafios desta comunidade, a proposta e realização de ações que aumentem a participação das mulheres, e que estas representem mais amplamente os diferentes estados brasileiros. Este trabalho apresenta algumas limitações. A primeira limitação, conforme anteriormente citado, está relacionada à identificação de gênero estar condicionada aos nomes dos autores. Sabemos que a análise de gênero envolve muitos outros aspectos aqui não explorados, e que desejamos no futuro abordar. Acreditamos que novas iniciativas possam surgir para investigar de forma mais abrangente a participação de diversos gêneros. Outra limitação é o fato de que, para extrair o estado de atuação de cada autor (por artigo), consideramos neste trabalho apenas a primeira instituição do autor. Além disso, a escolha de contabilizar o número absoluto de autores, sem levar em consideração a quantidade de publicações de cada um, faz com que possamos analisar de forma acertada a proporção entre os gêneros, mas também faz com que a proporção de produção destes autores seja perdida.

## Referências

- Jaroszewski, S., Lottridge, D., Haimson, O. L., and Quehl, K. (2018). “Genderfluid” or “Attack Helicopter”: Responsible HCI Research Practice with Non-Binary Gender Variation in Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–15, New York, NY, USA. Association for Computing Machinery.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8):707–710.
- Monteiro, M. (2019). *Ruined by Design: How Designers Destroyed the World, and What We Can Do to Fix It*. Mule Books, California.
- Nunes, D. J. (2017). Educação Superior em Computação Estatísticas – 2017. *Sociedade Brasileira de Computação - SBC*. <https://www.sbc.org.br/documentos-da-sbc/summary/133-estatisticas/1200-pdf-png-educacao-superior-em-computacao-estatisticas-2017>.
- ONU (2020). #Envision2030 Goal 5: Gender Equality. <https://www.un.org/development/desa/disabilities/envision2030-goal5.html>.
- Ribeiro, K., Azevedo, J., Maciel, C., and Bim, S. (2019). Uma análise de gênero a partir de dados da Sociedade Brasileira de Computação. In *Anais do XIII Women in Information Technology (WIT)*, pages 159–163, Porto Alegre, RS, Brasil. SBC.