

EXTRAÇÃO DE DADOS DE PRODUÇÃO CIENTÍFICA DOS ANAIS DO SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DO IFNMG (2009 A 2021)

FILHO, A.F.C.¹; DIAS, P.G.¹; SANTOS, F.R.S.¹; OLIVEIRA, I.V.S.¹; CARVALHO, M.A.D.²

¹Discente do curso de Ciência da Computação do IFNMG – campus Montes Claros; ²Docente do IFNMG – campus Montes Claros.

Palavras chaves: SIC; resumos expandidos; histórico; algoritmo.

Introdução

Em 2009, ocorreu no *campus* Januária – IFNMG, o Seminário de Inovação, Ciência e Tecnologia, o qual atualmente é intitulado Seminário de Iniciação Científica (SIC). O objetivo desse evento é fomentar o desenvolvimento científico não apenas do norte de Minas Gerais, mas também de todo território nacional. O SIC permite que estudantes, professores e pesquisadores de todo o Brasil apresentem suas pesquisas para a comunidade, e provê especialmente a oportunidade de indivíduos ainda inexperientes no trabalho científico aprimorarem suas competências. Esse seminário possui mais de uma década de existência e já publicou diversos textos de diferentes temas por diferentes *campi* do IFNMG e outras instituições (INSTITUTO FEDERAL DO NORTE DE MINAS GERAIS, 2022).

Surge naturalmente a curiosidade de se analisar sua evolução ao longo dos anos. O presente trabalho busca explorar informações pertinentes como: a quantidade de artigos publicados por edição, temática dos textos e as instituições às quais os artigos estão associados. Dessa forma, este texto se trata de uma pesquisa descritiva e exploratória sobre o SIC, e a justificativa para tal é verificar seu crescimento, foco temático e medir a contribuição de cada instituição no evento. Com essas informações disponíveis se pode, principalmente, identificar potenciais fatores deficientes no SIC e aplicar ações afirmativas para supri-las.

Metodologia

A fonte primária desta pesquisa são os anais publicados do SIC (de 2009 a 2021), os quais estão disponíveis no *site* do IFNMG. Assim, todos esses arquivos foram baixados manualmente e devidamente organizados por edição.

Em 12 anos, o SIC acumulou milhares de artigos publicados, por isso a extração manual dos dados relevantes é muito dificultosa. Tendo isso em vista, optou-se pela realização dessa tarefa por meios informáticos; foi criado um programa de computador capaz de ler arquivos brutos de texto contendo artigos do SIC e então gerar uma lista de itens contendo as informações úteis extraídas de cada artigo. O algoritmo utilizado para o desenvolvimento do programa foi baseado em busca por palavras-chave. Por exemplo, um artigo é automaticamente categorizado em Ciências Agrárias se a contagem de palavras típicas a essa área do conhecimento for superior à contagem de palavras típicas a qualquer outra área considerada.

O programa então analisou todos os artigos publicados, palavra por palavra, e fez a contagem de palavras-chave pré-determinadas. Foram ignoradas quaisquer classificações prévias que o artigo porventura tenha recebido de seus autores ou pela comissão do SIC. Por fim, a lista de itens gerada

pelo programa foi exportada para um sistema de banco de dados, o qual permite que dados requeridos sejam selecionados de uma maneira fácil e rápida. Além disso, permite que eles sejam recuperados em formato CVS (arquivo simples para tabelas de dados). Após todo o processo de refinamento, obteve-se gráficos a partir de tais arquivos CVS por meio de *softwares* de planilhas eletrônicas.

Resultados e discussão

De 2009 a 2021, foram contabilizados 2.538 artigos publicados pelo SIC. A edição de 2009 (primeira) contém 53 publicações e até 2019 (pré-pandemia de Covid-19), o seminário obteve crescimento médio de 78 artigos/ano. A pandemia de Covid-19 não apenas impossibilitou a edição de 2020 do SIC, como também prejudicou o desenvolvimento de trabalhos para a edição de 2021, ano em que o SIC ocorreu de forma não presencial e o número de artigos caiu 353 unidades em relação à última edição (2019). A quantidade exata de artigos em cada ano está expressa na Fig. 1.

Os artigos foram classificados nestas áreas do conhecimento: Ciências Exatas; Ciências Agrárias; Ciências Humanas; Ciências Biológicas; Ciências da Saúde; Engenharias; Linguística, Letras e Artes. Ciências Agrárias é a área do conhecimento mais explorada com 1241 (48,9%) trabalhos publicados. Isso é consistente com o fato de que a maioria dos *campi* do IFNMG possui cursos envolvendo esse ramo da ciência, o que condiz com a demanda tecnológica e científica da região do norte de Minas Gerais (ASSOCIAÇÃO COMERCIAL E EMPRESARIAL DE MINAS GERAIS, 2021). A segunda área mais explorada foi Ciências Humanas com 648 (25,5%) e, em terceira colocação, Ciências Exatas com 241 (9,5%) dos artigos publicados. As outras áreas possuem, em média, 102 (4,0%) publicações. A distribuição completa de temas está expressa na Fig. 2.

Levando em consideração o IFNMG, a instituição que mais possui artigos associados é o *campus* Janaúria com 689 (27,1%) publicações, um empate técnico com o *campus* Salinas, que possui 683 (26,9%) publicações. Em terceira colocação, está o *campus* Montes Claros, com 285 (11,3%) publicações. Os *campi* Janaúba, Porteirinha, Diamantina e Teófilo Otoni são os que menos possuem publicações associadas, apenas 28 (1,0%), em média. Instituições não pertencentes ao IFNMG compõem uma porção razoável, 140 (5,5%) publicações associadas identificadas. A quantidade de artigos por cada instituição está expressa na Fig. 3.

Considerações finais

De modo geral, o SIC cresceu consideravelmente de 2009 a 2021. Apesar dos efeitos negativos da pandemia entre 2020 e 2021, espera-se que o número médio de artigos volte a crescer a partir da décima edição do SIC em 2022.

Quanto à distribuição dos temas, uma vez que Ciências Agrárias é intensamente mais explorado do que as outras áreas, os autores desta pesquisa acreditam que a comissão do SIC deveria incentivar a realização de trabalhos sobre temas menos presentes; isso pode ser feito através de um maior encorajamento à participação do estudante do Ensino Médio que, com seus estudos interdisciplinares, contribuiria para a diversificação do conteúdo científico do SIC.

Finalmente, deve-se ressaltar a participação desigual dos *campi* do IFNMG. Embora isso possa ser majoritariamente explicado pela diferença entre a quantidade de integrantes de cada *campus* (docentes ou discentes), é útil expressar que os docentes dos *campi* com menos contribuições no SIC podem criar maiores estímulos a seus discentes para participarem e, assim, cooperem para a redução da gritante disparidade apresentada pela Fig. 3, a qual, possivelmente, reduzir-se-á à medida que todos os *campi* de menor porte se desenvolvam.

Referências

- ASSOCIAÇÃO COMERCIAL E EMPRESARIAL DE MINAS GERAIS. **Minas Gerais por regiões**. 2021. Disponível em: <<https://acminas.com.br/minasguide/pt/minas-gerais-por-regioes/>>. Acesso em: 12 de abr. de 2022.
- INSTITUTO FEDERAL DO NORTE DE MINAS GERAIS. **Anais do seminário de iniciação científica do IFNMG**. 2012. Disponível em: <<http://www.ifnmg.edu.br/seminarios-sic>>. Acesso em: 18 de mar. de 2022.
- INSTITUTO FEDERAL DO NORTE DE MINAS GERAIS. **X seminário de iniciação científica e II seminário de pós-graduação do IFNMG**. 2022. Disponível em: <<https://eventos.ifalmenara.com.br/sic2022>>. Acesso em: 18 de mar. de 2022.

ANEXO I

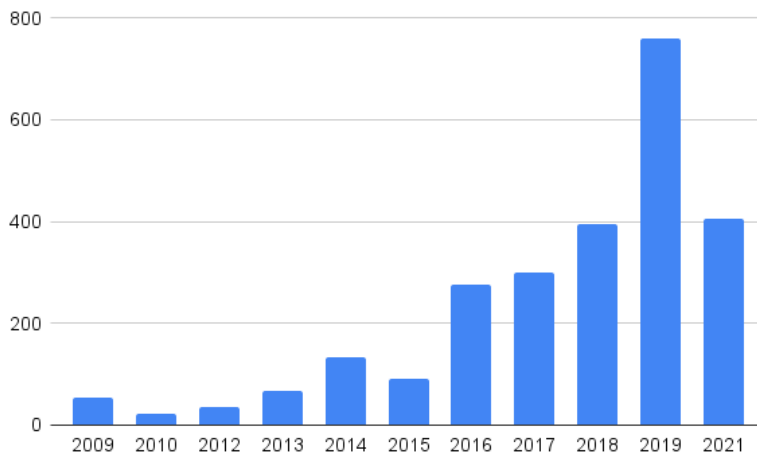


Figura 1. Número de artigos por ano. Fonte: Dados da pesquisa (2022).

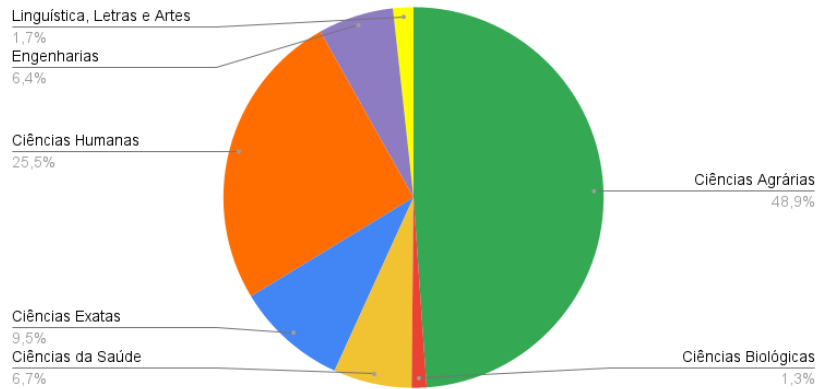


Figura 2. Número de artigos por área de conhecimento. Fonte: Dados da pesquisa (2022).

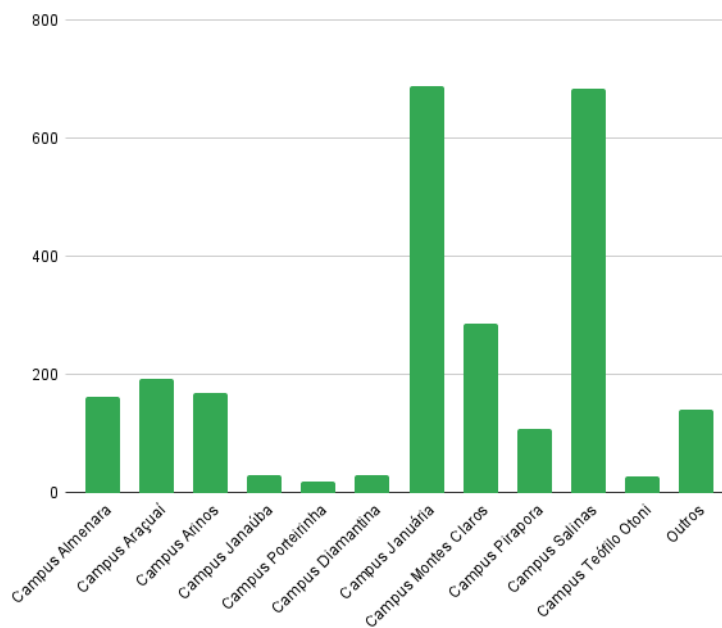


Figura 3. Número de artigos por instituição. Fonte: Dados da pesquisa (2022).