



A computação presente nas revistas do Centro Universitário da Região da Campanha – Urcamp

The computing present in the journals of the University Center of the Campaign Region – Urcamp

Ana Paula Oliveira Machado¹

Márcia Cristina Domingues Leite²

Fábio Josende Paz³

Resumo: A Ediurcamp, editora inserida no Centro Universitário da Região da Campanha é responsável pela publicação da produção científica, tecnológica e cultural no que diz respeito a Revistas, anais de eventos, relatórios técnicos e livros. Dentre estas, a Revista do Centro de Ciências da Economia e Informática (CCEI) e a Revista da Jornada de Pósgraduação e Pesquisa são dois importantes veículos de comunicação escrita que tem documentado, ao longo dos anos, importantes estudos que visam contribuir nas mais diversas áreas da ciência. Neste sentido, o estudo objetivou quantificar as publicações destas duas revistas nos últimos quatro anos, apontando o foco das pesquisas por ano e por área do conhecimento. Através do estudo, percebeu-se um aumento das publicações voltadas para a computação, principalmente relacionadas a temas voltados para a área de educação. Para isso, após separar os artigos que seriam utilizados na análise, utilizou-se ferramentas para mineração de textos, como Voyant Tools e Sobek para resumir e encontrar possíveis relações entre os artigos utilizados para análise.

Palavras-chave: Revista científica; Computação; Publicações.

Abstract: *Ediurcamp, publisher in the Centro Universitário da Região da Campanha is responsible for the publicity of scientific, technological and cultural production in regard to Magazines, annals of events, technical reports and books. Among these, the Journal of the Center of Sciences of the Economy and Informatics (CCEI) and the Journal of the Postgraduate and Research Journey are two important vehicles of written communication that have documented, over the years, important studies that aim to contribute in the most areas of science. In this sense, the study aimed to quantify the publications of these two journals in the last four years, pointing out the focus of the surveys by year and area of knowledge. Through the study, there was an increase in publications focused on computing mainly related to topics*

1Graduanda em Ciências Contábeis, Centro Universitário da Região da Campanha – Urcamp Bagé.
{anamachado182813@sourcamp.edu.br}

2Bacharel em Sistemas de Informação, Centro Universitário da Região da Campanha – Urcamp Bagé.
{marciacristina2312@gmail.com}

3Mestre em processos industriais – Centro Universitário da Região da Campanha – Urcamp Bagé.
{fabiopaz@urcamp.edu.br}

related to the area of education. For this, after separating the articles that would be used in the analysis, we used tools for text mining, such as Voyant Tools and Sobek to summarize and find possible relationships between the articles used for analysis.

Keywords: *Scientific journal; Computing; Publications.*

1. INTRODUÇÃO

As revistas científicas são, de maneira geral, uma ótima opção para quem deseja divulgar os resultados de suas pesquisas, pois proporcionam à comunidade respostas para problemas através de diversas metodologias, que é o principal objetivo de uma pesquisa (TEIXEIRA, E., 2011). A publicação eletrônica científica, como utilizada na Urcamp, destaca-se, em comparação com a publicação impressa, pois possui um pequeno período de tempo para aprovação e publicação de artigos, custos menores, agilidade no feedback, além de não possuir problemas com o armazenamento de impressões físicas (CUNHA, 1997).

A Revista do Centro de Ciências da Economia e Informática e a Revista da Jornada de Pós-Graduação – Congrega Urcamp, são revistas administradas pela Editora do Centro Universitário da Região da Campanha (EdiUrcamp), cujo propósito é levar para a comunidade artigos originais a nível regional, nacional e internacional. A primeira, é responsável pelas áreas de Administração, Ciências Contábeis, Computação e Engenharia, enquanto a segunda é responsável por todas as áreas que o Centro Universitário engloba.

A área da computação atua desde a elaboração de projetos e programação de ferramentas simples até a criação de softwares mais complexos (BROOKSHEAR, 2005). No entanto, para que se possa oferecer um desenvolvimento ainda maior para uma região, é necessário que além da realização dos projetos ocorra a publicação destes trabalhos, pois assim os resultados chegarão em uma parcela maior da população.

A fim de mensurar a quantidade de publicações científicas na área da computação, este estudo tem como objetivo realizar um levantamento de artigos relacionados a esse tema para identificar os assuntos de maior interesse por parte dos desenvolvedores para posteriormente realizar uma análise sobre eles.

O estudo foi estruturado da seguinte forma: na seção introdutória foi possível nortear o estudo, bem como realizar um breve resumo sobre os assuntos que o constituem e contextualizam. A etapa seguinte foi a revisão bibliográfica que teve como objetivo reunir estudos de outros autores que já escreveram sobre o mesmo assunto, posterior, na seção 3 é explicado os métodos utilizados para fazer a mineração de dados em cima dos artigos

escolhidos previamente. Na seção 4 são apresentados os resultados da mineração e uma discussão sobre os dados encontrados. Por fim, apresenta-se as conclusões e referências do artigo.

2. REFERENCIAL

Antes de iniciar um artigo para publicação em uma revista científica, é necessário compreender o que é ciência, para assim entender e utilizar suas metodologias da melhor forma possível. De acordo com Kerlinger (1979), os objetivos básicos da ciência são explicar fenômenos naturais e criar teorias, compreendendo de que maneira estes se relacionam com os demais fenômenos, psicológicos, sociológicos e educacionais. Além disso, é importante conhecer os problemas, que significa se questionar como as variáveis estão relacionadas. Tendo isso em vista, uma das metodologias utilizadas neste estudo foi o levantamento de dados, que, ainda segundo Kerlinger (1979), corresponde a uma grande metodologia pois possui como destaque a característica de ter a capacidade de inferir dados exatos a partir de amostras, além de ser importante para investigar as relações existentes entre variáveis.

Além disso, Santos (2003) afirma que há uma grande quantidade de periódicos circulando pelo mundo, e um amplo interesse por parte dos especialistas em dados quantitativos. Por ser um conjunto extenso de publicações, há a necessidade da utilização de processos que agilizem a análise desses dados. Complementando, Keim et al., (2002) destaca que todos anos ocorre a geração de um Exabyte de dados, e que técnicas como a mineração de textos são muito importantes para a análise exploratória de dados e grandes bancos de dados. Os dados quantitativos correspondem ainda como importante fonte de visualização de dados, pois levantam questões, ilustram informações e comunicam ideias (Slone, 2009).

No entanto, fazer análise em textos não estruturados é uma tarefa complexa e demorada, nesse sentido, surge a Mineração de textos que é uma área de pesquisa em computação que busca auxiliar nesse processo (VIJAYARANI *et al.*, 2015). A Mineração de

textos segundo Konchady (2006) consiste em explorar e identificar termos relevantes e estabelecer padrões textuais, além de analisar a frequência de aparecimento de termos no domínio a ser analisado, corroborando Vijayarani et al. (2015) acrescentam que as técnicas de Mineração de Texto podem ser utilizadas em diversos tipos de domínios de pesquisa, como processamento de linguagem natural, recuperação de informações, classificação e agrupamento de texto. Um exemplo pode ser

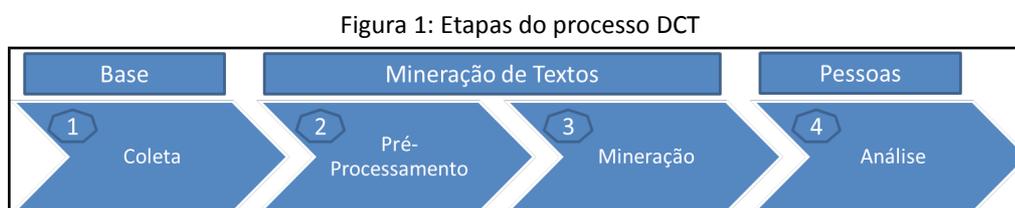
a visualização de dados em formato de nuvem de palavras, que serve para resumir visualmente um texto e como ponto inicial antes de uma análise mais profunda. Romero e Ventura (2007) afirmam que o aumento do uso da mineração no campo educacional é devido às possibilidades que a mineração proporciona conforme o objetivo de quem está fazendo uso de suas práticas.

Alguns estudos (BARION E LAGO (2008); KRASSMANN *et al.*, 2017; Paz e Cazella, 2018) demonstram que a Mineração de Texto (*text mining*), também conhecida como Descoberta de Conhecimento em Textos (DCT) é uma técnica adequada para manipulação automática de grandes coleções de dados textuais desestruturados. Barion e Lago (2008) demonstraram as técnicas utilizadas na mineração de texto, bem como sua funcionalidade e importância. Foi aplicada a técnica de Descoberta de Conhecimento em Banco de Dados, que possui as seguintes etapas:

Dados, Seleção de Dados, Processamento, Transformação, Mineração de Dados, Interpretação/Avaliação.

Krassmann *et al.* (2017) afirma que a aplicação de técnicas de Mineração de Texto demonstrou ser rápida, simples, intuitiva e eficiente para analisar um volume considerável de textos acadêmicos, o qual foram extraídas informações relevantes para o domínio investigado. Paz e Cazella (2018) apresentaram os resultados da mineração de textos de artigos científicos e descreverem os resultados e relações encontrados, referentes aos assuntos e autores da área pesquisada. Para isso, usaram a Descoberta de Conhecimento em Textos (DCT) cujo processo passa pelas seguintes etapas: Corpus, Pré-processamento, Mineração de Textos, PósProcessamento e Conhecimento.

Na Figura 1, visualizam-se as etapas do processo de Descoberta de Conhecimento em Textos (DCT) adaptado do estudo de Tan (1999).



Fonte: Adaptado de Tan (1999)

Na etapa de Coleta, consiste na formação da base de documentos ou Corpus para o estudo, o pré-processamento ou limpeza dos dados tem o intuito de remover ruídos e preparar os

dados para a mineração, ou seja, retirar caracteres indesejados, correção ortográfica e morfológica, também a realização da análise semântica e normalização dos vocabulários, na etapa da Mineração são aplicados métodos ou tarefas de mineração através de ferramentas e pôr fim a interpretação dos resultados obtidos (TAN, 1999; SCHIESSL E BRASCHER, 2011).

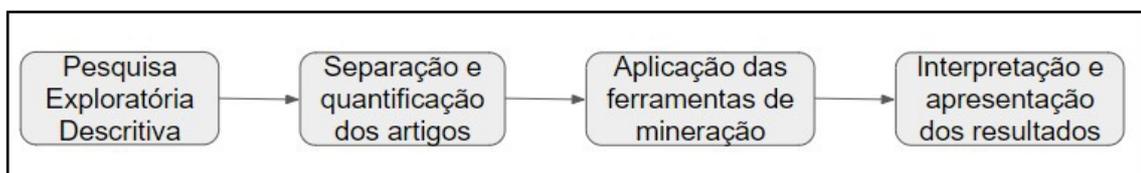
3. METODOLOGIA

Este estudo foi conduzido através de uma pesquisa exploratória e descritiva. Inicialmente, no processo de coleta de dados, buscou-se no site de cada uma das revistas pelas publicações dos anos de 2014 a 2017. Em seguida, coletou-se todos os artigos referentes aos anos citados, nas áreas de Administração, Ciências Contábeis, Computação e Engenharia. Após reunir todos estes artigos, ocorreu a separação e quantificação dos trabalhos por ano e área, a fim de se obter a percepção sobre a atual situação do número de publicações em cada área do conhecimento citada.

Adiante, após uma comparação entre os dados quantificados das revistas, buscou-se um aprofundamento a respeito da área de computação. Para isso, foi utilizado softwares como Sobek, que é uma ferramenta de mineração de textos que como resultado gera a demonstração de grafos, Voyant Tools, que possui diversas ferramentas para mineração e análise de textos. Para a utilização dos aplicativos mencionados, sucedeu-se a preparação dos artigos, no qual retirou-se dos mesmos as referências, rodapé, título, tabelas, figuras, autores e periódicos em que foram publicados, para que ocorra a correta análise do corpo do texto.

Ademais, foi feita a reunião de autores, da área de computação, contidos nas referências dos artigos analisados, para verificar em quais literaturas os autores estão buscando seu referencial teórico. Todos os passos encontram-se sintetizados na Figura 2:

Figura 2: Etapas do estudo



Fonte: Dados primários (2018)

4. RESULTADOS e DISCUSSÃO

Nos anos de 2014 a 2017, a Revista do Centro de Ciências da Economia e Informática e a Revista da Jornada da Pós-Graduação e Pesquisa – Congrega Urcamp publicaram, juntas, um total de 73 artigos, conforme Figura 3.

Figura 3: Quantificação das publicações nas revistas

<u>CCEI</u>	Administração	Ciências Contábeis	Computação	Engenharia
2014	6	2	2	0
2015	3	0	5	0
2016	3	1	3	0
2017	6	0	6	0
JORNADA	Administração	Ciências Contábeis	Computação	Engenharia
2014	3	0	2	2
2015	0	0	0	0
2016	10	0	1	0
2017	7	0	6	5

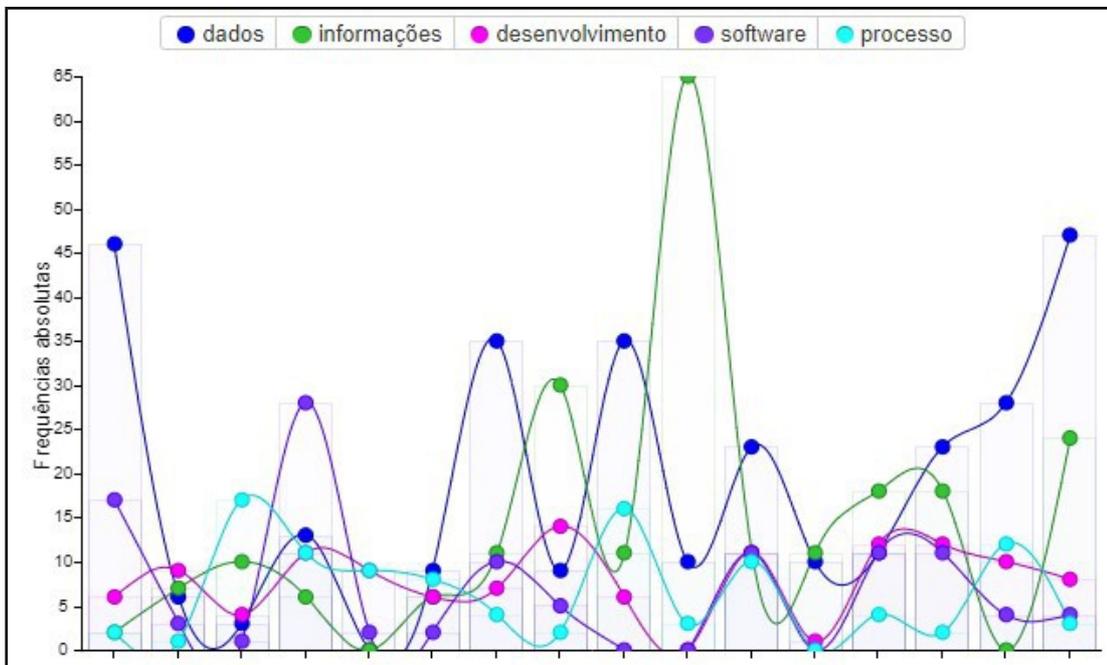
Fonte: Dados primários (2018)

Destes, na revista do CCEI, a computação possui um percentual de 43,24% do total de artigos. Enquanto na Jornada, apresenta 25% de um total de 36 trabalhos publicados. Isto demonstra que, comparado às outras áreas de comum publicação das duas revistas, a computação possui um papel de destaque, ficando atrás apenas da área de administração. Destaca-se o fato de que no ano de 2015 não houve edição da Jornada, evidenciando um número alto de publicações, pela administração, em 2016.

Utilizando as etapas do processo de Descoberta de Conhecimento em Textos (DCT), após ter reunido os artigos dos últimos 04 anos das duas revistas, separadamente, e posterior preparação dos mesmos para a utilização dos softwares, foi utilizada as ferramentas Cirrus, Terms e Trends do software Voyant Tools. E, posteriormente, a utilização do software Sobek.

Foram obtidos os seguintes resultados: primeiramente, a respeito dos 16 artigos publicados na Revista do Centro de Ciências da Economia e Informática nos anos de 2014 a 2017:

Figura 4: Frequência de palavras mencionadas nos artigos da Revista do CCEI, ferramenta *Trends* (Voyant Tools).

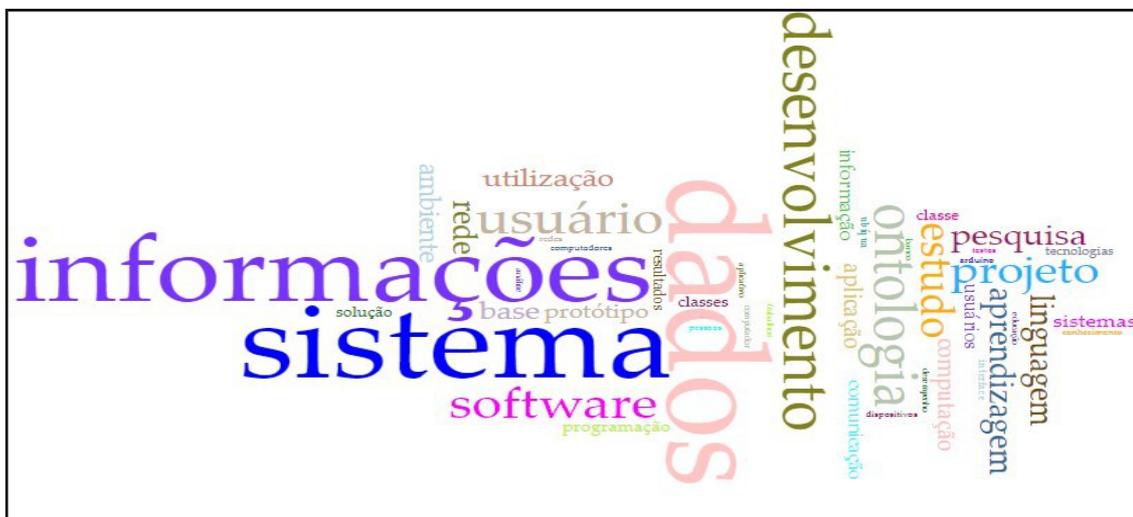


Fonte: Dados primários (2018)

A figura 3 representa a análise dos 16 artigos difundidos pela Revista do CCEI dos anos de 2014 a 2017, essa ferramenta possibilita a união dos artigos e verifica a frequência das palavras comuns em ambas publicações, além disso, demonstra essa repetição através de um gráfico que é dividido em segmentos no número de documentos inseridos na ferramenta.

Através deste é possível observar que em todos os artigos houve uma preocupação com os dados, bastante consideração com as informações excetuando-se os artigos 5 e 15. Ademais, percebe-se um interesse notável pela área de desenvolvimento, softwares e pelos processos que envolvem as áreas da computação.

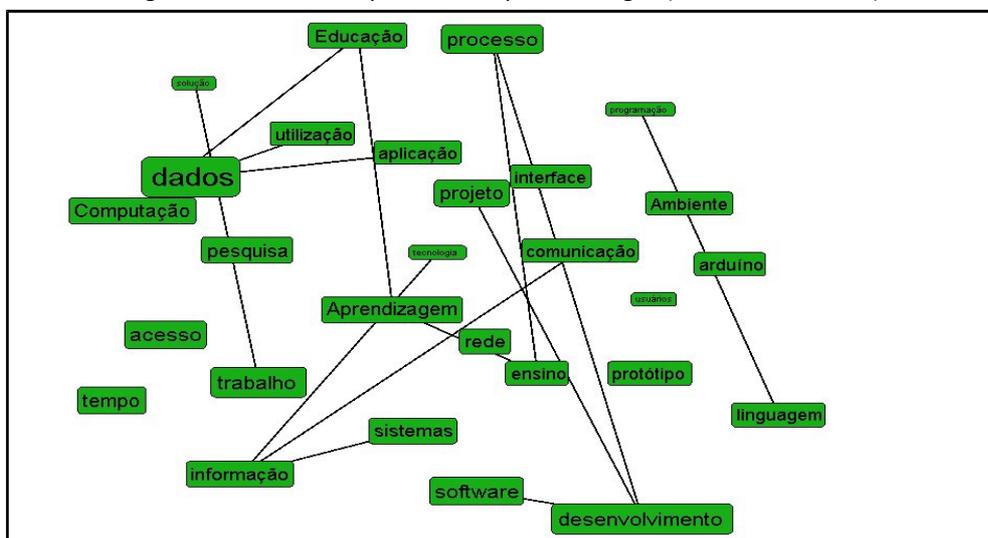
Figura 5: Nuvem de palavras feita a partir do corpo dos artigos, ferramenta Cirrus (Voyant Tools).



Fonte: Dados primários (2018)

A Figura 4 reforça graficamente o comportamento detectado em outras análises desse estudo. Alguns termos merecem destaque porque não estão diretamente associados a área de computação e podem apresentar tendências: ontologia, projeto, pesquisa e aprendizagem. Ainda, há de se acrescentar alguns termos que merecem destaque: protótipo, desenvolvimento e aplicação, demonstrando uma tendência ao desenvolvimento de aplicações nestes estudos.

Figura 6: Grafo feito a partir do corpo dos artigos (Ferramenta Sobek)



Fonte: Dados primários (2018)

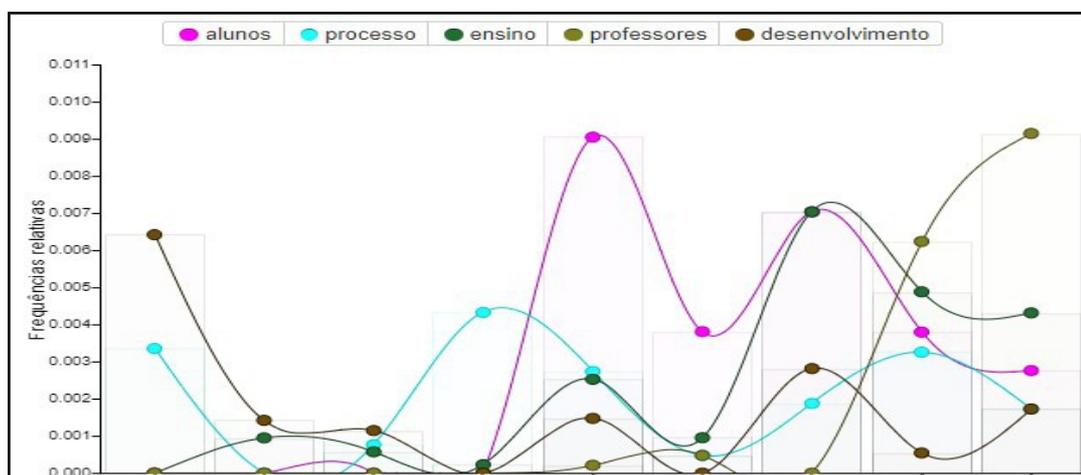
Por fim, através da ferramenta Sobek, representada pela Figura 5, é possível visualizar um grafo do mesmo conjunto de artigos, que evidencia as ligações entre termos. Além de

confirmar a tendência das figuras 3 e 4, o grafo revela que a palavra dados, mais citada pelos autores, tem ligações com utilização, aplicação e educação, que por sua vez, liga-se à palavra aprendizagem.

Como mostrado acima, os termos mais utilizados nos artigos publicados na revista do CCEI, foram dados e sistema, que aparecem nas 04 formas de verificação de texto. Outra importante observação é que também aparecem palavras na área educacional, como ensino e educação. Ou seja, há uma preocupação por parte dos desenvolvedores em pesquisar, projetar e desenvolver sistemas, a partir de diversos dados, voltados para essa área, tornando-a cada vez mais tecnológica.

Por fim, os 9 artigos da Revista da Jornada da Pós-Graduação e Pesquisa – Congrega Urcamp passaram pelo mesmo processo e geraram o seguinte resultado:

Figura 7: Frequência de palavras mencionadas nos artigos da Revista da Jornada, ferramenta *Trends (Voyant Tools)*.



Fonte: Dados primários (2018)

Novamente, passa-se o conjunto de artigos pela ferramenta *Trends*. Embora estejam contidas palavras da área de computação nos artigos da Revista da Jornada de Pós-Graduação, diferentemente do esperado, as publicações demonstraram uma preocupação maior com relação à área educacional, com grande incidência de palavras como “alunos”, “ensino” e “professores”, o que representa um interesse por parte dos escritores da área de computação em projetar e desenvolver softwares para a área educacional.

Figura 8: Nuvem de palavras feita a partir do corpo dos artigos

“tempo” e “professores”, como os que mais aparecem no corpo dos artigos reunidos, igualmente como mostrado na Figura 8.

Figura 10: Relação dos autores mais citados nas referências dos artigos publicados em ambas revistas.

Ordem	Autores	Contagem	País de Origem
1	MORAN, J.M.	9	Brasil
2	DIAS, F.L.	4	Brasil
3	BARANIUK, Richard G.	3	Canadá
4	KIPPER, L.M.	3	Brasil
5	LACERDA, D.P.	3	Brasil
6	LIBRELOTTO, Giovanni R.	3	Brasil
7	NIELSEN, J.	3	Dinamarca
8	PRESSMAN, R.S.	3	USA
9	SOMMERVILLE, Ian.	3	Inglaterra
10	BEVAN, N.	2	Inglaterra

Fonte: Dados primários (2018)

A Figura 8 demonstra os autores mais citados nos artigos de computação, tanto na Revista do Centro de Ciências da Economia e Informática (CCEI) quanto na Revista da Jornada de Pós-graduação e Pesquisa, destaca-se a grande presença de autores brasileiros, presentes em 05 posições de 10, isto revela a pouca busca pouco por autores internacionais. Por outro lado, confirmando as hipóteses anteriores, o autor mais citado pelos autores dos artigos deste estudo é da área educacional. Ou seja, além de projetar e desenvolver ferramentas tecnológicas para a área educacional, os autores basearam-se em um escritor de outra área para relatar a importância da tecnologia na educação descrita por ele. Excetuando os autores da área da computação, destaca-se ainda: LACERDA, D.P, da área de Engenharia de Produção.

5. CONSIDERAÇÕES FINAIS

Conforme os dados apresentados acima, os 25 artigos publicados nas duas revistas, demonstraram uma grande preocupação com dados e informações, bem como um interesse em desenvolver para área educacional. Através de referências nacionais e internacionais, atingiram um grande número de publicações em comparação com as outras áreas do conhecimento abrangidas pelas revistas. A utilização de ferramentas de mineração de texto possibilitou uma maior eficácia no alcance do objetivo, pois as mesmas demonstram quantitativamente e graficamente informações sobre o conjunto de artigos inseridos para análise.

Tendo isso em vista, acredita-se que há um aumento no número de publicações na área da computação, incentivados, provavelmente, pela área educacional. Esse dado revela que há um interesse por parte dos desenvolvedores, não só em realizar seus projetos, mas também divulgá-los, através de revistas científicas, para que a comunidade fique por dentro das novas tecnologias e desenvolvimentos para a região.

6. AGRADECIMENTOS

Este trabalho contou com financiamento da URCAMP através da Pró-Reitoria de Ensino, por meio de bolsa de Iniciação Científica.

7. REFERÊNCIAS

BARION, E. C. N.; Lago, D. Mineração de Texto. Revista de Ciências Exatas e Tecnologia, v. III, p. 123-140, Dezembro 2008.

BROOKSHEAR, J.G. Ciência da Computação: Uma Visão Abrangente. Porto Alegre: Artmed Editora, 2005.

CUNHA, Leo. (1997). Publicações científicas por meio eletrônico: critérios, cuidados, vantagens e desvantagens. **Perspect. cienc. inf.**, Belo Horizonte, v. 2, n. 1, p. 77 - 92, jan./jun.1997.

KEIM et al., "*Information Visualization and Visual Data Mining*", **IEEE Transactions on Visualization and Computer Graphics**, vol. 8, No. 1, Jan.-Mar. 2002, pp. 1-8.

KERLINGER, F. N. Metodologia da pesquisa em ciências sociais: um tratamento conceitual. São Paulo: EPU/EDUSP, 1979.

KRASSMANN, Aliane L., HERPICH, Fabricio., BERCHT, Magda., CAZELLA, Sílvio C. Analyzing trends in academic papers about ubiquitous virtual worlds in education using text mining. **International Journal for Innovation Education and Research**. Vol:-5, No-04, p. 167-180, 2017.

ROMERO, C; VENTURA, S. Educacional data mining: a survey from 1995 to 2005. **Expert systems with applications**. v. 33, n. 11, p. 135-146, 2007.

Rev. CCFI – URCAMP, V.24, n39, 2019

SANTOS, Raimundo Nonato Macedo dos. Produção científica: por que medir? o que medir? Revista Digital de Biblioteconomia e Ciência da Informação, 2003, vol. 1, n. 1.

SCHIESSI, Marcelo; BRASCHER, Marisa. Descoberta de Conhecimento em Texto aplicada a um Sistema de Atendimento ao Consumidor Knowledge Discovery in Text applied to a Help Desk System. **RICI: Revista Ibero-americana de Ciência da Informação**, v. 4, n. 2, p. 94-110, ago./dez, 2011.

SLONE, D. J. (2009). *Visualizing qualitative information*. **The Qualitative Report**, 14(3), 488-497. Disponível em: <<http://www.nova.edu/ssss/QR/QR14-3/slone.pdf>>. Acesso em: 29/10/2018.

TAN, A. H. Text Mining: The state of the art and the challenges , Proceedings of the Pacific-Asia. **Conference on Knowledge Discovery and Data Mining (PAKDD)**, Beijing - China, 26-28 April, 1999.

TEIXEIRA, E. (2011). A Análise de Dados na pesquisa Científica: importância e desafios em estudos organizacionais. *Desenvolvimento Em Questão*, 1(2), 177-201. <https://doi.org/10.21527/2237-6453.2003.2.177201>

VIJAYARANI, S., ILAMATHI, J., NITHYA. **Preprocessing Techniques for Text Mining - An Overview**. *International Journal of Computer Science & Communication Networks*, Vol 5(1), 7-16, 2015.