



ISSN: 2230-9926

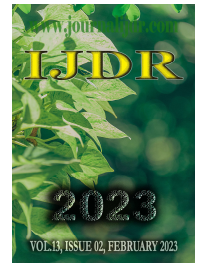
Available online at <http://www.journalijdr.com>

# IJDR

International Journal of Development Research

Vol. 13, Issue, 02, pp. 61813-61815, February, 2023

<https://doi.org/10.37118/ijdr.26346.02.2023>



RESEARCH ARTICLE

OPEN ACCESS

## LINGUISTICA DE CORPUS APLICADO A ARTIGOS CIENTÍFICOS DE QUÍMICA

Diego Sampaio Amariz\*<sup>1</sup> and Marcus Vinicius Carvalho Guelpeli<sup>2</sup>

<sup>1</sup>Mestrando do Programa de Pós-Graduação em Educação – Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM; <sup>2</sup>Professor do Programa de Pós-Graduação em Educação – Universidade Federal dos Vales do Jequitinhonha e Mucuri - UFVJM

### ARTICLE INFO

#### Article History:

Received 11<sup>th</sup> January, 2023  
Received in revised form  
25<sup>th</sup> January, 2023  
Accepted 10<sup>th</sup> February, 2023  
Published online 28<sup>th</sup> February, 2023

#### KeyWords:

Corpus; Química;  
Artigos; Estatística.

#### \*Corresponding author:

Diego Sampaio Amariz

### ABSTRACT

O interesse em aprender os conceitos químicos está conectado à concepção de que os conhecimentos abordados permitem um olhar de um mundo mais articulado e menos fracionado. Contribuindo para que o cidadão se veja como participante de um mundo em constante transformação. A construção de um corpus linguístico apresenta fatores que podem certamente auxiliar pesquisadores a obter e organizar informações para criar sua própria base de textos que auxiliem no processo de Mineração de Textos. Pretende-se apresentar o processo de construção de um corpus constituído de artigos sobre conceitos químicos e apresentar a metodologia de pesquisa utilizada para construir e organizar esse corpus. A partir das dificuldades encontradas no ensino de Química, a proposta desta pesquisa é oferecer aos pesquisadores, estudantes e profissionais na construção de sua própria base de textos - corpus que possam ser úteis dentro de determinados estudos e pesquisas específicas.

Copyright©2023, Diego Sampaio Amariz and Marcus Vinicius Carvalho Guelpeli. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Diego Sampaio Amariz and Marcus Vinicius Carvalho Guelpeli. 2023. "Linguística de corpus aplicado a artigos científicos de química", *International Journal of Development Research*, 13, (02), 61813-61815.

## INTRODUCTION

A Química é composta pelos conhecimentos fundamentais para o ser humano em seus distintos aspectos. Por meio desse conhecimento, é possível entender as perspectivas relacionadas à cidadania (MENESES e NUÑEZ, 2018, p.176). O interesse em aprender seus conceitos está conectado à concepção de que "Os conhecimentos abordados em Química permitem um olhar de mundo mais articulado e menos fracionado. Contribuindo para que o cidadão se veja como participante de um mundo em constante transformação" (PCNEM, 2000, p. 32). A química é uma disciplina obrigatória na educação básica. "No Brasil, essa ciência foi introduzida regularmente a partir de 1931" (LIMA, 2013). São muitas circunstâncias que dificultam o processo de aprendizagem dos discentes frente a disciplina de química. Este trabalho apresenta a construção de um *corpus*, composto por artigos relacionados aos conteúdos de química abordados no Ensino Médio, coletados da revista eletrônica Química Nova na Escola entre os anos de 1978 e 2022. A expressão *corpus* segundo (Sardinha, 2004), "representa um conjunto de dados lingüísticos textuais que foram coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua". A partir de 1990, os corpora passam a ter papel relevante nas pesquisas lingüísticas, uma vez que data dessa época o início das contribuições

advindas da Computação e da Lingüística Computacional" de acordo com (Aluísio & Almeida, 2006). Desse modo, segundo (Sardinha, 2000) "por meio do *corpus*, podem-se observar aspectos morfológicos, sintáticos, semânticos, discursivos, entre outros, relevantes para uma pesquisa lingüística. Podem-se ainda descobrir fatos novos, não perceptíveis pela intuição". Segundo (Aranha, 2006) "coleta é a etapa que tem como objetivo formar uma base de dados textual, conhecida na literatura como *Corpus* ou *Corpora*". Pode se dar de várias maneiras, porém todas necessitam de grande esforço, a fim de se conseguir material de qualidade e que sirva de matéria-prima para a aquisição de conhecimento. Portanto, a etapa de recuperação e coleta de dados tem como função a criação de uma base de dados textual chamada *corpus* ou *corpora* que servirá de base para aplicar as técnicas de Mineração de Textos. A formação do *corpus* sobre os conteúdos de Química abordados no Ensino Médio, auxiliará pesquisas no descobrimento de informações úteis sobre o tema. Uma vez que, são muitas circunstâncias que dificultam o processo de aprendizagem dos discentes frente a disciplina de química.

## METODOLOGIA

A pesquisa pretende gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos, logo, é classificada, quanto à natureza, como aplicada. Quanto aos objetivos e

procedimentos, trata-se de uma pesquisa exploratória, bibliográfica, qualitativa e quantitativa. A escolha do *corpus* para esse trabalho teve a duração de três meses, janeiro a março de 2022, foi definida após a leitura de artigos científicos em português. Foram observados os critérios de gratuidade, possibilidade de cópias das bases de dados e a classificação dos mesmos. Após a leitura de artigos científicos relacionados aos conteúdos de química da revista Química Nova na Escola, <http://qnesc.sbq.org.br/edicao.php>, essa revista foi a escolhida para essa pesquisa, porque aborda trabalhos relacionados a educação em química. Observa-se uma similaridade entre algumas palavras dentre os conteúdos da disciplina de Química. Este estudo visa a criação de um *corpus* relacionado aos conteúdos abordados nessa disciplina no Ensino Médio. Foram coletados os artigos científicos, onde estão armazenadas informações de trabalhos entre os anos de 1995 a 2022. Os artigos foram distribuídos em três classes, Química Inorgânica, Físico-Química e Química Orgânica, conforme a Figura 1.

| Nome                   | Data de modificação | Tipo              | Tamanho |
|------------------------|---------------------|-------------------|---------|
| 1 - Química Inorgânica | 04/12/2021 20:24    | Pasta de arquivos |         |
| 2 - Físico-Química     | 04/12/2021 20:38    | Pasta de arquivos |         |
| 3 - Química Orgânica   | 05/12/2021 19:47    | Pasta de arquivos |         |

Fonte: Autoria Própria, 2022.

Figura 1. Distribuição em classes dos artigos de Química da Revista Química Nova e Química Nova na Escola

Foram coletados 230 artigos científicos, entre os anos de 1995 e 2022, disponíveis no idioma português da revista Química Nova na Escola. A coleta teve como objetivo formar uma base de dados textuais, denominada “*Corpus*”.

## RESULTADOS E DISCUSSÕES

O *corpus* foi convertido do formato PDF para o formato TXT para o processamento computacional, nessa conversão foram retiradas as imagens, gráficos, tabelas, números de páginas e todas as anotações que não faziam parte do corpo do texto. Os arquivos foram renomeados seguindo uma ordem sequencial iniciando em 1 e terminando em 230 e foram colocados em uma única pasta denominada *Corpus* de química. Foi utilizado um software de análise chamado Get Finecount 2.6 para a contagem das palavras, conforme Figura 2.

| File name | Only rec. | Words | % of nu... | Charact... | Charact... | Redu... | Char... | Lines  | Sentences | Pages | File path                  | Words (...) | Charact... | Lines (R... |
|-----------|-----------|-------|------------|------------|------------|---------|---------|--------|-----------|-------|----------------------------|-------------|------------|-------------|
| 1.txt     | 4609      | 5901  | 18,51%     | 2601       | 2990       | 386     | 2012    | 594,36 | 1497      | 25,24 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 2.txt     | 5182      | 6187  | 18,09%     | 2845       | 3438       | 275     | 1843    | 625,42 | 1592      | 23,59 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 3.txt     | 4509      | 4981  | 7,62%      | 26140      | 33931      | 36      | 983     | 562,39 | 1066      | 30,21 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 4.txt     | 5506      | 5872  | 6,23%      | 29630      | 35438      | 60      | 792     | 644,33 | 937       | 34,61 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 5.txt     | 5431      | 5738  | 5,22%      | 31468      | 37092      | 39      | 767     | 674,40 | 1122      | 36,22 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 6.txt     | 3418      | 3220  | 22,08%     | 20623      | 24963      | 939     | 1826    | 138,39 | 1538      | 28,38 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 7.txt     | 3233      | 3846  | 15,94%     | 19069      | 22799      | 217     | 1588    | 414,53 | 1187      | 22,26 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 8.txt     | 7343      | 7885  | 6,87%      | 41992      | 49702      | 216     | 1651    | 905,13 | 1476      | 48,62 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 9.txt     | 3850      | 4330  | 11,03%     | 20588      | 25380      | 496     | 985     | 461,09 | 1086      | 24,77 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 10.txt    | 7655      | 8707  | 9,79%      | 45976      | 54920      | 69      | 2669    | 985,49 | 2440      | 52,82 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 11.txt    | 6372      | 7715  | 17,41%     | 37032      | 44556      | 300     | 2646    | 810,11 | 1979      | 43,51 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 12.txt    | 5328      | 5749  | 7,32%      | 28811      | 34723      | 269     | 900     | 631,33 | 1183      | 33,91 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 13.txt    | 4044      | 5054  | 20,22%     | 24472      | 29586      | 278     | 2747    | 539,75 | 2196      | 28,99 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 14.txt    | 4077      | 4950  | 17,64%     | 23970      | 28799      | 273     | 2200    | 523,62 | 1565      | 28,12 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 15.txt    | 3747      | 4347  | 13,80%     | 21480      | 25763      | 193     | 1232    | 468,42 | 1115      | 25,16 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 16.txt    | 3141      | 3255  | 6,39%      | 18585      | 21916      | 96      | 463     | 396,47 | 600       | 21,40 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 17.txt    | 3857      | 4206  | 10,22%     | 22044      | 27201      | 164     | 999     | 494,56 | 933       | 26,56 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 18.txt    | 3731      | 4367  | 14,56%     | 20030      | 25157      | 264     | 1331    | 457,40 | 1128      | 24,89 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 19.txt    | 3415      | 3812  | 10,41%     | 19437      | 23167      | 101     | 956     | 421,22 | 737       | 22,62 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 20.txt    | 4277      | 4826  | 11,39%     | 25568      | 30244      | 108     | 1345    | 549,89 | 1288      | 29,54 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 21.txt    | 3747      | 4366  | 14,18%     | 19030      | 22793      | 132     | 1194    | 414,42 | 1095      | 22,26 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 22.txt    | 3165      | 3489  | 9,29%      | 18262      | 21759      | 125     | 703     | 395,62 | 593       | 21,25 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 23.txt    | 1956      | 2068  | 5,09%      | 9945       | 11966      | 19      | 424     | 217,56 | 350       | 11,69 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 24.txt    | 3444      | 3940  | 11,32%     | 18507      | 22373      | 122     | 892     | 440,78 | 828       | 21,05 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 25.txt    | 3710      | 4197  | 11,69%     | 20494      | 24661      | 112     | 1276    | 446,39 | 920       | 24,08 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 26.txt    | 3259      | 3682  | 11,49%     | 18324      | 21891      | 69      | 1045    | 398,02 | 814       | 21,38 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 27.txt    | 4578      | 5467  | 16,11%     | 26521      | 31631      | 195     | 1787    | 976,11 | 1533      | 36,89 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 28.txt    | 2883      | 3195  | 10,29%     | 15983      | 19553      | 77      | 718     | 344,60 | 605       | 18,51 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 29.txt    | 3373      | 3832  | 11,98%     | 18798      | 22629      | 190     | 923     | 411,44 | 827       | 22,10 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 30.txt    | 3216      | 3573  | 9,99%      | 17661      | 21088      | 105     | 727     | 383,42 | 680       | 20,59 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 31.txt    | 2871      | 3110  | 14,12%     | 14934      | 17859      | 94      | 894     | 324,62 | 699       | 17,44 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 32.txt    | 3028      | 3387  | 10,29%     | 16455      | 19698      | 47      | 814     | 380,15 | 662       | 19,24 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 33.txt    | 2477      | 2883  | 14,08%     | 14152      | 16940      | 89      | 964     | 398,00 | 673       | 16,54 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |
| 34.txt    | 6769      | 8116  | 16,60%     | 40067      | 48009      | 411     | 2874    | 872,89 | 2485      | 46,88 | C:\Users\Diego\Desktop\... | N/A         | N/A        | N/A         |

Fonte: Autoria Própria, 2022.

Figura 2. Contagem de palavras através do programa Finecount

Foram realizados cálculos de amplitude, média aritmética, desvio padrão e coeficiente de variação para os 230 artigos científicos coletados. Esses resultados estão evidenciados na Tabela 1.

Tabela 1. Cálculos realizados no *Corpus* de Química contendo 230 artigos

| Tabela Estatística       |                |
|--------------------------|----------------|
| Corpus                   | Quantidade     |
| Média Aritmética         | 4.405 palavras |
| Desvio Padrão            | 1.659 palavras |
| Amplitude                | 8.992 palavras |
| Coefficiente de variação | 37,67%         |

Fonte: Autoria Própria, 2022.

Com esses resultados o *corpus* contendo 230 artigos apresenta-se heterogêneo em relação a distribuição da palavras. Com isso o mesmo foi selecionado e 120 artigos foram separados em uma nova pasta denominada *Corpus* química 120 e foram estudados para buscar uma homogeneidade. O resultado desse estudo está descrito abaixo. No *Corpus* química 120 foi realizado o cálculo da amplitude, por meio da quantidade de palavras existentes nos artigos. A amplitude é uma medida de dispersão que determina o grau de variação dos números, essa medida é determinada pela diferença entre o valor máximo encontrado e o valor mínimo, conforme a equação matemática abaixo.

$$R = X_{\text{máximo}} - X_{\text{mínimo}}$$

Onde X máximo é o valor máximo encontrado e X mínimo é o valor mínimo encontrado. O artigo que contém a maior quantidade de palavras apresentou 5.566 e o artigo que apresentou a menor quantidade de palavras apresentou 2.428 palavras. A amplitude encontrada foi de 3.138 palavras, conforme Tabela 2.

Tabela 2. Cálculos realizados no *Corpus* de química 120

| Tabela Estatística       |                |
|--------------------------|----------------|
| Corpus                   | Quantidade     |
| Média Aritmética         | 3.730 palavras |
| Desvio Padrão            | 800 palavras   |
| Amplitude                | 3.138 palavras |
| Coefficiente de variação | 21,45%         |

Fonte: Autoria Própria, 2022.

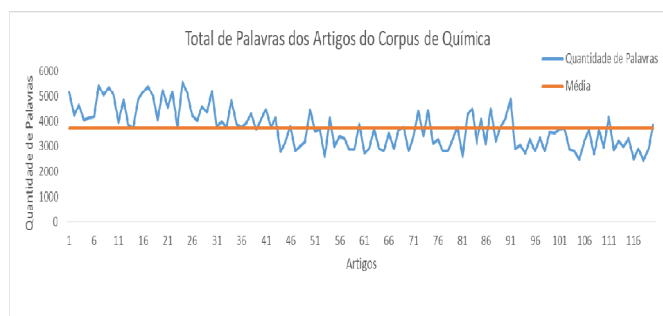
Também foi realizado o cálculo da média das palavras existente nos artigos desse *corpus*. A Média faz parte dos conceitos da Estatística. No que concerne à média, esta pode ser aritmética (simples ou ponderada), geométrica, harmônica, quadrática, cúbica ou biquadrática. Tratando especificamente da média aritmética, esta é considerada como “o conceito mais básico da Estatística e da ciência experimental, é também o mais utilizado na vida cotidiana das pessoas” (MAGINA; CAZORLA; GITIRANA; GUIMARÃES, 2010, p. 61-62). A equação utilizada para o cálculo da média está descrito abaixo.

$$\bar{X} = \frac{\sum X_i}{n}$$

Onde, Xi é a quantidade total de palavras existentes em todos os artigos e n a quantidade de artigos selecionados. Realizado esse cálculo observou-se que a média de palavras para os artigos foi 3.730 palavras, conforme Tabela 2. A quantidade de palavras em relação à média está evidenciada na Figura 3. Esse gráfico evidencia a dispersão da quantidade de palavras em relação à média. Com o valor da média obtido pode-se calcular o desvio padrão existente nesse *corpus*. O desvio padrão é o protótipo das medidas de dispersão em virtude de suas propriedades matemáticas e de seu uso na teoria da amostragem” (OLIVEIRA, 2017, p. 8). É uma medida de dispersão, que indica o quanto o conjunto de dados é uniforme. Conforme equação matemática abaixo.

$$S = \sqrt{\frac{\sum (X - \bar{X})^2}{n}}$$

Onde,  $X$  é o valor individual de cada palavra de cada artigo, é a média dos dados obtidos e  $n$  número total de artigos o corpus. Conforme a Tabela 2 o desvio padrão para os dados foram 800 palavras. Com os dados obtidos da média e do desvio padrão pode-se calcular o coeficiente de variação dos dados. Experimentos confiáveis requerem a avaliação dos resultados pela verificação da precisão deles próprios, que pode ser realizada pelos valores dos coeficientes de variação, CV, (NESI et al., 2010; STORCK et al., 2011). Conforme Steel et al. (1997), o CV permite a comparação de resultados de diferentes experimentos, envolvendo uma mesma variável ou espécie, permitindo, assim, quantificar a precisão de suas pesquisas.



Fonte: Autoria Própria, 2022.

**Figura 3. Dispersão da quantidade de palavras em relação à média do Corpus**

O CV é uma medida importante sobre a variabilidade dos resultados experimentais, podendo ser útil na definição do número de repetições do ensaio, necessário para detectar uma diferença entre médias de tratamentos com uma dada probabilidade (PIMENTEL-GOMES, 2009; NESI et al., 2010). De acordo com Storck et al. (2011), a distribuição de CV possibilita estabelecer faixas de valores que orientam os pesquisadores sobre a validade de seus experimentos. Dessa forma, podemos dizer que o coeficiente de variação é uma forma de expressar a variabilidade dos dados excluindo a influência da ordem de grandeza da variável. O coeficiente de variação é igual ao desvio-padrão dividido pela média aritmética, multiplicado por 100% (LEVINE et al., 2014).

$$CV = \frac{S}{X} \times 100$$

Onde,  $S$  é o desvio padrão e é a média dos dados obtidos

Como o coeficiente de variação analisa a dispersão em termos relativos, ele será dado em %. Quanto menor for o valor do coeficiente de variação, mais homogêneos serão os dados, ou seja, menor será a dispersão em torno da média (LEVINE et al., 2014). De uma forma geral, se o CV for menor ou igual a 15% o resultado apresenta uma baixa dispersão dos dados, dados homogêneos. Se o cálculo dos dados ficarem entre 15 e 30% os mesmos apresentam uma média dispersão. E se for maior que 30% apresentam uma alta dispersão, dados heterogêneos (LEVINE et al., 2014). Portanto, ao se calcular o coeficiente de variação dos dados do *Corpus* de química 120 obteve-se 21,45%, conforme Tabela 2, o que indica uma média homogeneidade dos artigos que compõem esse *corpus*. Entretanto esse valor obtido está mais próximo de 15% do que de 30% o que indica uma maior tendência a homogeneidade do que uma tendência para a heterogeneidade.

## CONCLUSÃO

Este artigo apresentou o desenvolvimento de um *Corpus* aplicado a Química, expondo os meios para realizar a estatística de um *corpus*, por meio de artigos que abordam conceitos químicos a fim de descobrir informações, tendências e padrões sobre o assunto estudado.

Este *corpus* tem como objetivo viabilizar estudos da influência destes textos dentro da Química, por meio das técnicas de Mineração de Textos - MT, como a clusterização e a sumarização. Conforme Carrilho Jr. (2007, p. 12), o principal objetivo de se minerar textos é “descobrir conhecimento novo e inovador a partir de massas de texto livre, isto é, na formatural que conhece-se e lida-se diariamente, agregando valor comercial a empresas e organizações”. Dessa forma, uma grande quantidade de textos não estruturados pode ser transformada em informação útil, a partir da aplicação de técnicas de MT como sumarização, clusterização e classificação, dentre outras (ARANHA E PASSOS, 2006). Então a utilização da Mineração de Textos torna-se viável para analisar o conteúdo gerado nesse *corpus*, possibilitando descobrir informação desconhecida e relevante que pode auxiliar no aprimoramento dos conceitos abordados pela Química.

## REFERÊNCIAS

- Aluísio, S. M., & Almeida, B. G. M. 2006. O que é e como se constrói um corpus? Lições aprendidas na compilação de vários corpora para pesquisa linguística. *Calidoscópico*, 4(3), 156-178.
- Aranha, C.; Passos, E. (2006) A Tecnologia de Mineração de Textos. Lab. ICA Elétrica PUC-Rio. *RESI-Revista Eletrônica de Sistemas de Informação*, Nº2.
- ARAUJO, Amanda Caroline Ferreira; DE OLIVEIRA FÉLIX, Maria Elisabeth; DA SILVA, Gilberlândio Nunes. RELATO DAS DIFICULDADES EM APRENDER QUÍMICA DE ALUNOS DA EDUCAÇÃO BÁSICA DE UMA ESCOLA PÚBLICA DE CAMPINA GRANDE.
- BIZERRA, Ayla Márcia Cordeiro et al.. Dificuldades e motivações no ensino de química: uma análise da perspectiva docente. VI CONEDU - Vol 1... Campina Grande: Realize Editora, 2020. p. 1406-1420. Disponível em: <<https://editorarealize.com.br/artigo/visualizar/65351>>. Acesso em: 05/11/2022 16:16
- CAVALCANTE, Marcelo; DA COSTA, João Gomes. Considerações sobre planejamento experimental e adequabilidade do uso de testes estatísticos em Ciências Agrárias. *Diversitas Journal*, v. 6, n. 4, p. 3706-3723, 2021.
- CRUZ, Luanna Azevedo. Modelo para recuperação de informação em repositórios institucionais utilizando a técnica de sumarização a partir da seleção de atributos do Cassiopeia. 2019.
- DOS REIS BRAGA, Fabiane. Extração semiautomática de taxonomia para domínios especializados usando técnicas de mineração de textos. *Ciência da Informação*, v. 45, n. 3, 2016.
- GUELPELI, Marcus VC; BRANCO, Antonio Horta; GARCIA, Ana Cristina B. Cassiopeia: A model based on summarization and clusterization used for knowledge discovery in textual bases. In: 2009 International Conference on Natural Language Processing and Knowledge Engineering. IEEE, 2009. p. 1-8.
- GUELPELI, MARCUS VINICIUS CARVALHO. Cassiopeia: Um modelo de agrupamento de textos baseado em sumarização. Niterói: Tese (Doutorado em Computação) - Univerisade Federal Fluminense, 2012.
- MAGINA, Sandra; FONSECA, Sônia. A APRENDIZAGEM DA MÉDIA ARITMÉTICA SIMPLES A PARTIR DE MATERIAIS DIDÁTICOS DISTINTOS: uma comparação entre duas propostas de ensino Learning simple arithmetic average using different teaching materials: a comparison between two educational proposals. *Em Teia| Revista de Educação Matemática e Tecnológica Iberoamericana*, v. 7, n. 1, 2016.
- OLIVEIRA, Cassius Gomes et al. Desvio padrão e imprecisão de leitura: Paquímetro. *Caderno de Graduação-Ciências Exatas e Tecnológicas-UNIT-SERGIPE*, v. 5, n. 3, p. 27-27, 2019.
- SCHMILDT, Edilson Romais et al. Coeficiente de variação como medida da precisão em experimentos de alface. *Revista Agro@mbiente On-line*, v. 11, n. 4, p. 290-295, 2017.